Construction of a Simian Immunodeficiency Virus Vaccine from Cynomolgus Macaque Cytomegalovirus

by

Justen Norman Hoffman Russell

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Immunology
University of Toronto

Construction of a SIV Vaccine from Cynomolgus Macaque Cytomegalovirus

Justen Norman Hoffman Russell
Doctor of Philosophy
Department of Immunology
University of Toronto
2018

Abstract

The current HIV pandemic has killed more than 35 million people since 1985, and infects a further 30 million globally. Vaccines based on herpesvirus vectors, in particular cytomegalovirus, generate unique, robust immune responses that may overcome many of the typical obstacles to an effective preventative HIV vaccine. Recent studies using a cytomegalovirus based vaccine in the rhesus macaque monkey model have protected close to half of all vaccinated monkeys from SIV, but questions remain as to the mechanism of protection and whether similar results will be seen in humans against HIV.

These questions may be answerable through the use of the cynomolgus macaque monkey model, and a minimally passaged cytomegalovirus vector that more closely resembles human cytomegalovirus. To this end a novel cytomegalovirus was isolated from a Mauritian cynomolgus macaque monkey. This virus was sequenced in full and phylogenetically characterized alongside other primate cytomegaloviruses and their hosts.

To generate a vaccine construct the newly isolated and minimally passaged cynomolgus macaque cytomegalovirus was first cloned as a bacterial artificial chromosome (BAC) to facilitate the preservation of the virus without further accumulation of attenuations from *in vitro* passaging. After sequencing the BAC, and comparison of the growth kinetics of the viruses derived from the BAC to determine if there was an observed *in* vitro loss of fitness, the BAC was used to construct a vaccine against SIV.

This new cynomolgus macaque cytomegalovirus based SIV vaccine for use in cynomolgus macaques retains wildtype growth kinetics *in vitro*. It is the first cytomegalovirus based SIV vaccine for use in the cynomolgus macaque model and the first cytomegalovirus used in vaccine studies that purposely has been generated with a focus on retention of viral fitness. This is a valuable molecular tool that will permit the further understanding of herpesviruses as vaccine vectors and their potential application in humans.

Acknowledgements

First and foremost I want to thank Dr. Kelly S. MacDonald, my supervisor on this project. You were the vision behind much of what was done here, and the voice that encouraged me to chase my own passions in research. You were the voice of reason when nothing seemed to work, and could always be counted on to match my euphoria when it came, even when the most exciting results were sent as they emerged in the small hours of the morning. I value your mentorship on this project as much as I value the tangential in-depth conversations we had around the role of grand-parenting in elephant troops, or the difficulty of strapping newton meters to baby sea turtles. Wherever I end up after this I will recall fondly my time in the MacDonald lab.

Kathryn, we were only dating when we came together to Toronto each chasing our own science. We lived separately at first so we could find ourselves, but I am so glad we came together eventually so we could find us. Somehow, in the middle of this PhD, we fit in three moves, two masters' degrees (both yours) and a wedding – making me wonder if perhaps I took way too long. You supported me when I needed it, distracted me when you could, encouraged me with your every smile, and made these years the happiest I have had yet. None of this would exist without you. I cannot wait to see where we end up next, but I know it will be fantastic because it will be with you.

To my parents– you gave me an analytical mind and an argumentative spirit. I am sure anyone who has met me is unsure how you tolerated a young me, but at least some of them seem thankful that you did. You certainly did something right. Thank you for only ever being a phone call away.

To Catia, my co-conspirator in science, we started these degrees together and you will not be far behind me in finishing. We were not always in the same city, but we were always in the same boat. You were my co-reveller and co-commiserate in the MacDonald lab. Thank you for your collaborations, guidance, companionship and lessons in Portuguese.

To Meika Richmond, Jackie Chan, Cristina Csordas, David Willer, Joseph Antony, Angie Marsh, Kamran Haq, and Aruna Ambagala – my lab mates and partners in science. You were all more than colleagues. It was a pleasure to work with, and get to know each of you.

To Vineet Joag, Angela Zhou, Jordan Schwartz, and all the other students in the Department of Immunology thank you for welcoming me to Toronto with open arms. I will miss our boardgames, and island trips, coffee breaks and holiday parties. It has been a pleasure to be in each of your proximities.

The Department of Immunology has been wonderful throughout my time, and I am thankful to all the wonderful people within it. I am proud that this work might join the shelves within it, alongside the works of other accomplished scientists because of the pride and professionalism the people there place in their work. Thank you for admitting me and accepting me among your ranks.

To the University of Toronto, that ancient institution with its sheer stone facades, I will remember my time here fondly. I will remember the opportunities and resources so generously bestowed upon its scholars. I will remember the feuding unions that so shocked a naïve Alberta kid, and the inner tube water polo that kept me fit. Thank you.

To the people of the JC Wilt infectious diseases research centre and the whole Canadian National Microbiology Laboratory my brief time in Winnipeg has been amazing because of you. Thank you.

To the HIV Vaccine Conference 2013– I wrote you a letter without hope and expectation, but with a new student's brashness and impulse. I asked for a scholarship to visit your conference in Barcelona because I had no funding, or right to be there on my own. Somehow you accepted my reasoning, that as a new student, young an impressionable with nothing yet to present, I could benefit the most from attending, I asked for the chance to network and learn, and you gave it to me. I have conversations still based on talks I saw there, in my first year as a graduate student – and have collaborated on projects that exist only because you took a risk and helped me attend. A part of this thesis is dedicated to that conference and the unknown reviewers who accepted my pitiful application. Thank you.

Table of Contents

List of Common Abbreviations

aa	Amino Acids
AAV	Adeno-associated virus
ADCC	Antibody Dependant Cellular Cytotoxicity
AD5	Adenovirus Type 5
AIDS	Acquired Immune Deficiency Syndrome
APC	Antigen Presenting Cell
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3
ART	Antiretroviral Therapy
ATCC	American Type Culture Collection
BAC	Bacterial Artificial Chromosome
BaCMV	Baboon Cytomegalovirus
BCG	Bacilli Calmette-Guerin
BNAb	Broadly neutralizing antibody
bp	Base pairs
BST-2	bone marrow stromal antigen-2
Cas-9	CRISPR associated protein 9
CCMV	Chimpanzee Cytomegalovirus
CCR5	CC Chemokine Receptor 5
CE	Common Era
CMV	Cytomegalovirus
COX-2	Cyclooxygenase-2
CFU	Colony Forming Units
CIHR	Canadian Institutes for Health Research
CPE	Cytopathic Effect
CRISPR	Clustered regularly interspaced short palindromic repeats
CTL	Cytotoxic T Lymphocyte
CXCL5	CXC Chemokine Ligand 5
CXCR4	CXC Chemokine Receptor 4
CyCMV	Cynomolgus Macaque Cytomegalovirus

DC-SIGN	Dendritic cell-specific Intercellular adhesion molecule-3-grabbing non-integrin
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
dsDNA	Double Stranded DNA
eGFP	Enhanced Green Fluorescent Protein
Env	HIV Envelope Protein
ELISA	Enzyme Linked Immunosorbent Assay
ELISPOT	Enzyme Linked ImmunoSpot Assay
FBS	Foetal Bovine Serum
Fc	Fragment, Crystallisable
Gag	Group Specific Antigen
gB	Glycoprotein B
gH	Glycoprotein H
gL	Glycoprotein L
GMCMV	Green Monkey (Aotine) Cytomegalovirus
gO	Glycoprotein O
GPF	Gag-Pol Fusion Gene
GPCMV	Guinea Pig CMV
GTR	Generalized Time Reversible
HCMV	Human Cytomegalovirus
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
HSV-1	Herpes Simplex Virus-1
HVTN	HIV Vaccine Trials Network
IE	Immediate Early
IF	UNAIDS New Investment Framework
IFE	UNAIDS Investment Framework Enhanced
IFN	Interferon
IFNγ	Interferon-gamma
IFTM	Interferon Induced Transmembrane

IL	Interleukin
IRL	Internal Repeat Long
IRS	Internal Repeat Short
IV	Intra Venous
kDa	KiloDalton
KSV	Kaposi's Sarcoma-associated Herpesvirus
kya	Thousand Years Ago
LB	Luria-Bertani Broth
MCMV	Murine Cytomegalovirus
MEM	Minimum Essential Media
MHC	Major Histocompatibility Complex
MPI	Message Passing Interface
MRC-5	Medical Research Council Cell Strain 5
MSF-T	Mellow Skin Fibroblasts – Immortalized cynomolgus macaque fibroblast line
MSM	Men who have Sex with Men
mya	Million Years Ago
NCBI	National Center for Biotechnology Information
Nef	Negative Regulatory Factor
NF-κB	Nuclear Factor Kappa-light-chain-enhancer of Activated B Cells
NKC	Natural Killer Cell
NTR	NeTaRev (Nef-Tat-Rev) Fusion Gene
OHTN	Ontario HIV Treatment Network
OMCMV	Owl Monkey Cytomegalovirus
ORF	Open Reading Frame
OriLyt	Origin of Lytic Replication
PBMC	Peripheral Blood Mononuclear Cells
PFU	Plaque Forming Units
Pol	DNA Polymerase Gene
PrEP	Pre-exposure Prophylaxis
PTGS2	Prostaglandin-Endoperoxide Synthase-2

PRR	Pattern Recognition Receptor
qPCR	Quantitative PCR
RCMV	Rat Cytomegalovirus
RhCMV	Rhesus Macaque CMV
RNA	Ribonucleic Acid
rpm	Rotations per Minute
RRE	Rev Response Element
RRV	Rhesus Rhadinovirus
SAMHD1	Sterile Alpha Motif Domain and HD Domain-containing Protein 1
SHIV	Simian-Human Immunodeficiency Virus
SIV	Simian Immunodeficiency Virus
SIVcpz	Chimpanzee SIV
SIVmac	Macaque SIV
SIVsm	Sooty Mangabey SIV
SMCMV	Squirrel Monkey CMV
SNP	Single Nucleotide Polymorphism
sps	Substitutions per Site
ssRNA	Single Stranded Ribonucleic Acid
stHIV	Simian Tropic HIV
TAP	Transporter Associated with Antigen Processing
Tat	Trans-activator of Transcription
TBS	Tris Buffered Saline
TBST	TBS + Tween 20
Telo-RF	Telomerised Rhesus Macaque Fibroblast Cell Line
TB	Tuberculosis
T_{EM}	Effector Memory T Cell
T_{RM}	Tissue Resident Memory T Cell
TLR	Toll-like Receptor
TNFRSF14	Tumor Necrosis Factor Receptor Superfamily Member 14
TNFα	Tumor Necrosis Factor-Alpha
TRIM5α	Tripartite Motif-containing Protein 5

TRL	Terminal Repeat Long Region
TRS	Terminal Repeat Short Region
UL	Unique Long Region
UNAIDS	Joint United Nations Programme on HIV/AIDS
US	Unique Short Region
USA	United States of America
USD	United States Dollars
Vif	Viral Infectivity Factor
Vpr	Viral Protein R
Vpu	Viral Protein U
Vpx	Viral Protein X
VZV	Varicella Zoster Virus
YDNA	Y Chromosomal DNA
Δnef	Nef Deficient

List of Figures

List of Tables

List of Manuscripts arising from this Thesis

1. Marsh A. K., Ambagala A. P., Perciani CT**, Hoffman Russell J. N**., Chan J. K., Janes M., Antony J. M., Pilon R., Sandstrom P., Willer D. O., MacDonald K. S. 2015. Examining the species-specificity of rhesus macaque cytomegalovirus (RhCMV) in cynomolgus macaques. PLoS One 10:e0121339.

2. **Hoffman Russell JN,** Marsh AK, Willer DO, Ambagala AP, Dzamba, M, Chan JK, Pilon R, Fournier J, Brudno, M. Antony JM, Sandstrom P, Evans BJ, MacDonald KS. 2016. A novel strain of cynomolgus macaque cytomegalovirus: implications for host-virus co-evolution. BMC Genomics 17:277 DOI: 10.1186/s12864-016-2588-3

3. **Hoffman Russell JN**, Marsh AK, Ambagala AP, Chan JK, Csordas FC, Richmond M, Antony JM, Willer DO and MacDonald KS. 2017 (Submission Pending). Vector Engineering of novel cynomolgus cytomegalovirus based SIV vaccine constructs.

Chapter One: Introduction

1.1 Vaccines against Human Immunodeficiency Virus

1.1.1 HIV

1.1.1.1 HIV Virology

Human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS), is a retrovirus of the lentivirus family [1-3]. It is an enveloped virus 100-120nm in diameter containing two copies of a single stranded ribonucleic acid (ssRNA) genome in a conical viral capsid [4-7]. The genome contains 9 open reading frames (ORFs) that encode just under 20 viral proteins [4, 5, 8]. Following viral entry into the target cells, CD4+ cells, the virus uncoats [4]. Like other retroviruses the ssRNA genome is translated into a cDNA copy that is then trafficked into the cell nucleus and integrated into the host genome [4, 9, 10]. Integrated HIV can remain in the host genome untranscribed as the cell returns to a resting state, becomes dormant, activates, differentiates, or even proliferates [11-16]. Latently infected cells may not produce any viral proteins and are undetectable to the host immune system before activation, though low-level viral transcription is observed in a portion of latently infected cells [12, 15, 16]. Some integrated HIV genomes will be completely non-functional and never produce virus or viral genes [12]. Alternatively, HIV ORFs can activate immediately upon integration or at a later time point to produce viral proteins [11, 12, 14, 17]. Newly produced virus can bud from the cell surface within 24-48 hours of entry into a target cell [11, 18-20].

The HIV envelope (Env) structural protein mediates viral entry. It is a highly variable protein and can vary in amino acid sequence by as much as 35% between viral clades, and 20% within viral clades [3]. The *env* gene is translated as a single polyprotein, gp160, which is subsequently spliced into the glycoproteins gp120 and gp41 [11]. The transmembrane gp41, and glycosylated gp120 associate and move to the cell surface as a trimer of protein heterodimers necessary for mediating viral entry into target cells [11].

The Env trimer binds to the CD4 receptor of CD4+ cells to initiate viral entry [11]. Following a conformational change, Env binds with a second co-receptor, commonly CC chemokine receptor 5 (CCR5) [18, 21-26]. Alternatively, in CXC Chemokine receptor 4 (CXCR4) or dual tropic HIV,

CXCR4 may be used as a co-receptor [18, 25-28]. Transmitted founder virus is typically CCR5 tropic, becoming dual or CXCR4 tropic only in later infection following mutation [18, 21, 22, 26, 29]. Timothy Ray Brown, known as the Berlin patient, was infected with CCR5 tropic HIV when he received a bone marrow transplant for acute myeloid leukemia from a CC chemokine receptor (CCR5)-Δ32 HIV resistant donor. His reconstituted CD4+ cell population was no longer susceptible to viral entry from Env utilizing CCR5 co-receptor. Though he remains susceptible to CXCR4 tropic HIV, viremia has not returned years following his treatment [30].

Following Env binding of the co-receptor, the viral envelope fuses with the cell lipid membrane releasing the viral capsid and tegument proteins into the cytoplasm. As the Env trimer is the only exposed HIV protein on the virion surface it is the primary target for most antibody based vaccine attempts (see Section 1.1.3). Antibodies against HIV occur in natural HIV infection and are a key component of the host immune response against HIV but are typically non-neutralizing and do not lead to the elimination of HIV from the host (see section 1.1.2). As Env is only produced late in the viral lifecycle in the HIV infected cells, cytotoxic T lymphocyte (CTL) responses against Env are thought to be of limited value in controlling the virus [31].

In some cases vaccine induced immune responses against Env have increased rather than decreased susceptibility to HIV [32, 33]. Viral infection can be aided by the recruitment of activated CD4+ immune cells, which are target of the virus, to sites of infection [32, 33]. Viral entry into these immune cells may be further aided by the binding of the host cell Fc (Fragment, crystallisable) receptor to antibody or complement opsonised virions, attaching the virus to its potential target [32, 33]. Integrins or their ligands expressed on activated immune cells may similarly enhance cell infection rates as they interact with integrins or ligands retained on the HIV viral envelope [34-37]. In macaques a varicella-zoster virus (VZV) based Env vaccine that activated CD4+ T cells without inducing neutralizing antibody production or CD8+ T cell responses resulted in enhanced susceptibility to SIV in vaccine recipients, demonstrating that a non-productive but highly immunogenic vaccine actually increases mucosal susceptibility to SIV [32, 38].

HIV encodes for two other structural proteins, Group-specific Antigen (Gag) and DNA (deoxyribonucleic acid) Polymerase (Pol). The HIV virion is given structure by the proteins of the *gag* gene. *Gag* is translated as a single polyprotein, pr55, that is digested into smaller proteins by the HIV protease [4]. The nucleocapsid protein p7 covalently binds and protects the ssRNA genome [11]. The viral capsid is formed from p24 and HIV is anchored to the lipid envelope by matrix protein p17 [11].

Gag is a tempting target for a HIV T cell mediated vaccine as it is made early in the HIV infection cycle and is highly conserved. It contains the most highly conserved epitopes – across strains, clades

and time – in the HIV genome [39]. Escape mutations to avoid CTL responses have the highest fitness costs in the *gag*, *pol* and *nef* genes and revert the most frequently [40]. Gag is present on the surface major histocompatibility complex (MHC) -I of infected cells within two hours of viral entry, well before Negative Regulatory Factor (Nef) mediated downregulation of MHC-I can begin 6-12 hours after viral entry [31].

The *pol* gene is transcribed out of frame as part of a single large ORF with *gag* [41, 42]. It is translated only after ribosomal slippage occurs at a UUUA sequence located along the mRNA to produce a -1 frameshift in the translation [41]. As a result Gag-Pol pre-protein is produced at a ratio to Gag alone between 1:8 [41] and 1:20 [42]. The digested Pol proteins are encapsulated in the HIV virion and distributed into newly infected cells [11].

Pol pre-protein is digested into Protease, which is responsible for digesting Pol and other HIV pre- or poly proteins, and the proteins Integrase, Reverse transcriptase, and RNAse H which are essential to the retroviral aspects of the HIV lifecycle [11]. Following fusion of the viral and cell membrane, and uncoating of the viral capsid, Reverse transcriptase converts the two ssRNA HIV genome copies into a cDNA copy [11]. Frequent switching of which ribonucleic acid (RNA) genome is utilized as the template during copying introduces errors and contributes to the high mutation rate [4, 43]. RNAse H digests the RNA copy of the HIV genome, and Integrase catalyzes the integration of the genome into the host cell [4, 11]. Both Reverse transcriptase and Integrase are key targets of antiretroviral therapy (ART) [11]. Reverse transcriptase has a high error rate, and is partially responsible for the high mutation rate of HIV [11] (See section 1.1.2). Host proteins, including A3 cytidine deaminases, further contribute to establish a mutation rate of one incorrect base pair every 1500-4000 DNA bases [44]. This high error rate can be both beneficial and problematic for viral fitness. It drives adaptation to immune or chemical interventions, but also produces non-functional virus [4, 44]. Replication incompetent or non-functional virus is difficult to distinguish from latent virus by PCR based methods, often confounding estimates of the size of the latent viral reservoir [45].

Similar to Gag, Pol is expressed early after viral entry and found on surface MHC-I within 2 hours of viral entry making them both natural targets for CD8+ T cell mediated vaccine approaches [42]. Functional regions of the *pol* gene are well conserved across HIV clades, meaning a single vaccine may be sufficient to target most strains of HIV [46]. In studies with SIV, CD8+ T cells specific for Pol were able to eliminate infected cells early after viral entry, prior to the expression of other proteins, or the production of virus [42].

Beyond the structural proteins, HIV-1 encodes two regulatory proteins, Rev and Trans-activator of Transcription (Tat), and four accessory proteins. SIV and HIV-2 encode a fifth accessory protein, Viral Protein X (Vpx) but most do not encode Viral Protein U (Vpu). The regulatory proteins Tat and

Rev regulate gene expression, both supressing select host genes, and activating viral genes. Tat is an intracellular transcription activator to promote HIV gene expression, and a secreted protein [47-50]. It is produced and secreted very early after viral entry [51]. Secreted Tat binds surface receptors on surrounding cells with immune modulatory effect [52-54] and to upregulate CCR5 and CXCR4 expression [55-57]. Endocytosis of Tat into surrounding nerve cells causes excitotoxicity, neuronal degradation, and neuronal cell death [58].

Rev is a regulatory protein responsible for the shift from early to late gene production [51, 59]. Genes bearing the Rev response element (RRE), like *gag* and *env*, exhibit codon optimization dissimilar to host genes, with a low G/C content and a high frequency of U/A in the final letter of each codon [60, 61]. The odd codon usage, and small compatible tRNA pool, slows translation of these genes in the early phase of HIV replication [60]. As it builds up in the nucleus, Rev binds the RRE on HIV mRNAs to facilitate export to the cytoplasm, and improve the rate of translation [59-62]. Rev is essential for viral replication and virion formation will not occur in the absence of this role [51, 59-62]

Similar to Pol, Tat and Rev gene products are relatively highly conserved across clades [51, 63]. This makes them ideal vaccine targets. Antibody responses generated against Tat react across subtypes [51, 64]. Tat and Rev specific CTL responses are inversely correlated with progression to AIDS in HIV infected individuals [65, 66]. Tat specific antibodies are capable of sequestering secreted Tat, preventing neurotoxicity and its immune modulatory effect [51, 67, 68]. They are only observed in 10-20% of HIV infected individuals but are correlated with non-progression to AIDS [51, 64, 68]. Rev has not been used alone as a vaccine target but is frequently included in candidate vaccines alongside other genes [51].

The accessory proteins Nef, Viral Infectivity Factor (Vif), Viral Protein R (Vpr), and in HIV-1: Vpu or in HIV-2 and SIV: Vpx, are not essential for productive HIV infection *in vitro* [69, 70]. *In vivo*, however, the accessory proteins play key roles in viral infection [69, 71-73].

Nef is translated as a 27 kiloDalton (kDa) protein with a myristoylation sequence to target it to the cell membrane [74, 75]. A smaller non-myristoylated form of Nef also produced [69, 76]. Nef is packaged into the HIV virion [69, 77]. Upon entry into the cell this virion trafficked Nef remains with the viral core or forming ribonucleoprotein complex, while newly synthesized Nef localizes throughout the cytoplasm, nucleus, and plasma membrane [69, 77-79]. The *nef* gene is not expressed as early in the infection cycle as *gag* or *pol* (see above), but Nef is present within 12 hours after viral entry [42].

Nef interacts with a number of proteins to influence cell activation, to increase virion production rates, and to prevent apoptosis [69, 80]. Secreted Nef similarly affects surrounding non-infected cells,

and may stimulate viral replication in latently infected cells [69, 81]. Through multiple mechanisms cell membrane associated Nef alters endosome trafficking pathways to downregulate surface expression of cell proteins including CD4, MHC-I, MHC-II, CD28, and TNF, and to upregulate expression of DC-SIGN (Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin) [69, 82-84]. MHC-I and MHC-II downregulation interferes with the immune detection of infected cells, while DC-SIGN upregulation improves the clustering of infected dendritic cells with HIV susceptible T cells to improve cell-to-cell transmission [69, 82, 85]. CD4 downregulation limits subsequent superinfection of HIV infected cells and curtails the binding of surface expressed Env to CD4 on infected cells [69, 84]. Together both dramatically improve the spread of HIV within a host [69, 84, 86].

There is above-average conservation across the *nef* gene between clades, marking it a potential vaccine target [46, 75]. However, most HIV patients develop both T cell and antibody responses against Nef and this natural immune response does correlate with protection from disease progression [46, 69].

HIV and SIV lacking Nef are replication competent, but produce lower viral yields *in vitro* and impaired pathogenicity *in vivo* [69, 87]. A number of HIV infected long-term non-progressing patients have been found with mutations in the *nef* gene [69, 88, 89]. Macaques infected with *nef*-deleted SIV (Δnef) attenuated vaccine candidates progressed towards simian AIDS at a slower rate and were protected from infection with closely related strains of SIV [51, 69, 90-94] (See Section 1.1.2). Vaccination with Δnef provided one of the highest levels of protection against further challenge with SIV to date [90-93].

The *vpu* gene is present in HIV-1 and chimpanzee SIV (SIVcpz) -related SIV strains, but not HIV-2 or sooty mangabey SIV (SIVsm) related SIV strains including macaque SIV (SIVmac) [69, 72]. The protein is dispensable, particularly *in vitro*, but improves viral pathogenicity *in vivo* [69, 72]. Vpu is not trafficked in the virion, and is instead produced in infected cells where it targets intracellular CD4 and facilitates virion budding [69, 72]. CD4 binds to the Env precursor in the endoplasmic reticulum of infected cells, preventing the export of either to the cell surface [69, 72]. Vpu binds the Env precursor-CD4 complex to target CD4 for proteosomal degradation and free Env for surface expression [69, 72].

Vpu is also important in the budding of completed HIV virions from the cell [69, 72]. Vpu binds a number of cell factors; in tissue culture the effect of Vpu on viral budding is dependent on the host cell phenotype [72]. In SIV and HIV-2, most of which lack *vpu*, an additional domain of Env fills a similar role [72]. In SIVmac this function is also aided by Nef [72].

Vif is produced late in HIV and SIV replication and packaged into the virion [4, 69, 95-97]. Vif was initially thought extraneous as it is unnecessary in many common cell culture lines, all of which were later found to lack APOBEC3 (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3) [4, 69]. APOBEC3 is a cytidine deaminase which causes hypermutation of the pre-integration HIV cDNA genome [73, 95, 98]. Vif promotes the proteosomal degradation of APOBEC3 enzymes to protect the HIV or SIV genome from excessive mutagenesis [4, 44]. Vif is essential for HIV fitness throughout replication in CD4+ T cells and macrophages *in vivo* [69]. A3 proteins lead to a substantial decrease in the fitness of HIV when not countered by Vif [44, 98]

Vpr is a highly conserved protein in HIV and SIV, and a key target of host CTL responses [99]. Proportional to coding space, Vpr contains the greatest number of epitopes highly susceptible to reversion after escape mutation [40]. It is also a key virulence factor. Vpr is packaged in the virion and secreted Vpr can cross cell membranes to act in surrounding uninfected cells [69] Virion packaged Vpr is associated into the HIV/SIV pre-integration complex which traffics the viral cDNA genome into the cell nucleus [69, 99]. Independently Vpr drives apoptosis in CD4+ T cells and neurons [69, 99]. This acts to supresses the host immune response [69, 99]. CTL responses and antibody titres are higher in SIV infection with Vpr deleted strains [69, 99].

In infected and surrounding cells Vpr acts through multiple redundant pathways to arrest the cell cycle in the G2 phase [69, 99]. In this phase host factors bias gene transcription and translation towards viral genes [69]. Host genes which utilize the 5’ m7GpppN cap for translation initiation are supressed while the cell is in G2, preventing extraneous gene translation [69]. Meanwhile transcription from the viral LTR promoter is marginally increased, and polyadenylation of viral genes is improved [69]. This leads to an overall increase in virus gene translation [69, 99].

In some SIV strains the functions of Vpr are reduced and instead accomplished by Vpx, a related gene in SIVsm related strains of SIV and HIV-2 [69, 100]. Vpx is incorporated into the virion, and localizes in the nucleus. Vpx is incorporated into the pre-integration complex and contributes to nuclear import [100]. Vpx is also key in protecting the virus from SAMHD1 (Sterile alpha motif domain and HD domain-containing protein 1), a restriction factor expressed in myeloid cells that restricts HIV replication despite viral entry [73, 98]. Vpx is highly variable in a species specific manner having evolved rapidly alongside SAMHD1 to protect the virus in its natural host [73]. As such, and because it is not in HIV-1, Vpx is not a potential vaccine target in HIV vaccines or their models.

Restriction factors play a role in limiting HIV and SIV to their native hosts, and can be confounding factors in studies using animal models of HIV. Among the key restriction factors are the IFN induced transmembrane (IFTM) proteins IFTM1, IFITM2 and IFITM3, which are expressed in

IFN stimulated cells [98]. IFTM proteins are incorporated into the virus envelope as HIV buds from infected cells and interfere with subsequent fusion of the virus to target cells, impeding viral infection [98]. Tripartite motif-containing protein 5 alpha (TRIM5α) also impede cellular infection by preventing the uncoating of the viral core after it enters the cytoplasm, leading to proteosomal degradation rather than infection [4, 73].

Tetherin, also known as bone marrow stromal antigen-2 (BST-2), binds to and links the HIV matrix protein to the cell cytoskeleton preventing viral budding [69, 72, 73].

Schlafin 11 prevents viral mediated alteration of the cellular tRNA pool, potentially through the sequestering of tRNA [60]. In the absence of Schlafin 11 tRNA pools are altered in infected cells to advantage the expression of viral genes over host genes [60, 101].

HIV and SIV have evolved alongside these restriction factors and have specific mechanisms to circumvent them in their natural host [73, 98]. When virus is transmitted to a distant host, such as when humans are exposed to SIV or non-human primates to HIV, the virus is often unable to establish infection or spread because of these or other restriction factors [98].

HIV is a result of the cases where the virus was able to adapt and overcome the species barriers. Early outbreaks of HIV in the 1920's in Cameroon and the 1950's in Kinshasa are thought to have begun in hunters handling the carcases of SIV infected animals [102]. HIV-1, transmitted to humans from such encounters with SIV infected apes, is the predominant form of HIV globally. There are 4 groups of HIV-1, groups M, N, O and P [18]. Group M, N and O result from separate historical transmissions of HIV to humans from chimpanzees. Group P was transmitted to humans from Gorillas [18]. The majority of HIV infections globally are derived from the 1980's group M pandemic and can be classified into 9 subtypes, A, B, C, D, F, G, H, J, and K as well as multiple circulating recombinant forms [18].

HIV-2, which occurs mainly in western Africa, originated through transmission to humans from sooty mangabeys [18, 103]. HIV-2 results in the same symptoms as HIV-1, but with a prolonged asymptomatic phase and slower disease progression as the virus is less adapted to human hosts [18, 103]. It also has a lower transmission rate between humans than HIV-1 [18].

1.1.1.2 HIV Pathogenesis

In all but the rarest of cases HIV results in lifelong infection that can be divided into three general phases. All patients will experience an initially acute phase, followed by a chronic phase that, without treatment, lasts an average of 8-10 years [104, 105]. Without treatment HIV replicates in, and gradually depletes, host CD4+ T cell leading to AIDS [106-109]. HIV infection establishes a latent viral reservoir with a long half-life [110-112]. Though HIV can be treated and controlled with ART to

below the limit of detectability for plasma viral load, it cannot be eliminated and viremia returns if ART ceases [110, 111].

Acute HIV infection is further classified into an eclipse phase of approximately ten days, before HIV RNA is detectable in the plasma, followed by five Fiebig stages [29, 113] (**Figure 1-1**). A Sixth Fiebig stage marks the end of acute HIV and the commencement of chronic HIV infection. The Fiebig stages are defined by the detection of landmark molecules in patient plasma after infection [113]. Viral RNA can be first found in the plasma during Fiebig I, the GAG p24 antigen during Fiebig II, and HIV targeting IgM during Fiebig III [113]. By Fiebig IV indeterminate HIV protein bands can be detected by western blot while determinant protein bands to at least two of p24, gp41 and gp120/gp160 but not p31 can be detected by western blot during Fiebig V [113].

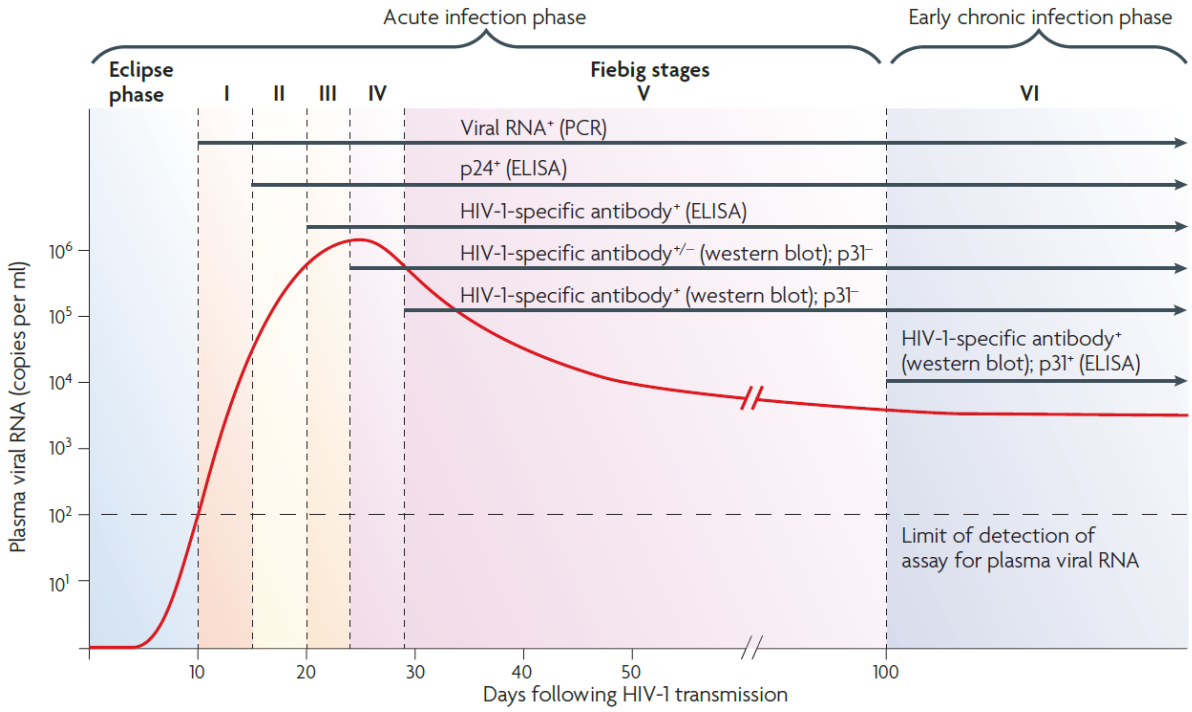


Figure 1-1: Early stages of HIV infection. Red line plots trend for plasma viral load following infection while Fiebig stages are indicated above. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology [29], Copyright 2010 modified with permission from [114] copyright 2008 National Academy of Sciences, USA.

During the eclipse phase HIV infection is established. In sexual transmission multiple exposures are usually required for infection to occur [29]. With simian immunodeficiency virus (SIV), a non-human primate model of HIV, infection can be initiated with cell bound or free virus, but successfully crossing the epithelial barrier is a rare event for HIV and SIV in either form [29, 115]. A small

number of CD4+ cells are initially infected with a small number of viruses, or in 80% of cases - a single founder virus [18, 29, 114, 116, 117]. The initial founder strain is typically CCR5 tropic, not CXCR4 or dual tropic, regardless of method of transmission [18, 21, 22, 26, 29]. The virus replicates locally, infecting CCR5+ CD4+ T cells as well as CD4+ monocytes and macrophages [29, 118].

The innate immune response to HIV begins early. Signalling molecules can be found in the blood during the eclipse period [29]. Pro-inflammatory cytokines including interleukin (IL)-1, IL-18, IL-22, tumor necrosis factor alpha (TNFα), and interferons (IFNs) are released in response to stimulation of pattern recognition receptors (PRRs) in infected and nearby cells [29, 98]. These recruit innate and adaptive immune cells to the area as well as promote the expression of a number of anti-viral genes, including many restriction factors that limit viral infection of, or replication in, host cells [98]. Natural killer cells (NKCs) in particular seek out viral infected cells early, and contribute to the milieu of antiviral cytokines and chemokines [29].

Within 10 days following sexual transmission HIV reaches the draining lymph nodes, where it spreads among CCR5+ CD4+ T cells [29]. Dendritic cells and B cells bind to, and help spread, the virus to activated T cells within the lymph nodes [29, 119, 120]. The virus disseminates and replicates rapidly, spreading throughout the host lymphoid tissue at the end of the eclipse phase and the start of Fiebig I, where viral RNA can be detected the plasma [29].

The initial adaptive immune response to HIV comes from CD8+ T cells [29, 121, 122]. CD8+ T cell responses mount through Fiebig I/II between day 10 and day 20 and drive the early selection of mutations in the virus [29, 121, 122]. Antibodies against HIV are slower to emerge. The viral load rises rapidly during Fiebig I and II peaking in Feibig III/IV [29]. With HIV-1 and SIV, 20% of CD4+ T cells in the gut-associated lymphoid tissue become infected within the first three weeks [29, 123-125].

The first antibodies against Env are seen around 18-23 days after initial infection [29]. Peak plasma viral load occurs shortly afterwards at 21 to 28 days after initial infection [29]. At this time there may be more than a million RNA copies of the HIV genome per milliliter of blood [29].

During this time the proportion and number of CD4+ CCR5+ T cells decrease dramatically in the sigmoid colon from 11×10^6 cells/g of tissue (69.1% of CD3+ T cells) in HIV uninfected individuals, to 6.9×10^6 cells/g of tissue (67.3% of CD3+ T cells) in Fiebig I/II, 1.2×10^6 cells/g of tissue (35.5% of CD3+ T cells) in Fiebig III and 8×10^5 cells/g of tissue (17.4% of CD3+ T cells) in Fiebig IV [126]. Proportions of peripheral blood CD4+ cells and mucosal Th17 cells, an important antibacterial CD4+ T cell line in the mocosa, fall dramatically between Fiebig I/II and Fiebig II [126]. These cell populations begin to replenish slowly if ART commences later in infection, but never fully recover in

number or function [126]. CD4+ T cells, including crucially those specific for HIV, are rapidly depleted. [126, 127]

After peak viral load the adaptive immune response pushes the viral load down throughout Fiebig V over the course of 12-20 weeks to a relatively stable set-point viral load [29]. Within the first 80 days of infection more than 80% of CD4+ T cells in the gut may be depleted and – due to a lack of CD4+ T cell help, up to 50% of B cell germinal centers [29, 123-125, 128]

The early events and the extent of lymphoid CD4+ cell depletion in in Fiebig I-IV will determine rate of disease progression afterwards [127]. The gradual decline of CD4+ T cells throughout Fiebig VI and the eventual onset of AIDS can be prevented through the use of ART. However, chronic immune inflammation will be maintained throughout infection with HIV-1, even if treated with ART [29, 129].

Neutralizing antibodies against HIV are rarely seen before 12 weeks, and typically develop around 3 months after infection [18, 29]. This is after immune control of the virus and the set-point viral load have already been established. Once they appear they begin driving the selection of Env mutations [18, 29]. Antibodies that appear within the first 3-6 months are typically specific for the infecting HIV strain and target the V1 and V2 region of HIV Env [94, 130, 131]. Broadly neutralizing antibodies (BNAbs) develop in only a portion of HIV infected individuals and appear late in infection [121, 132-135]. These target regions of Env like the CD4+ binding site [136, 137], V1V2 glycan [138-140], V3 [137, 141] or membrane proximal region [142, 143] which are better conserved across strains and clades of HIV [132, 133, 135]. Antibodies against these regions must adapt to self-mimicry and tight conformational requirements. As a result they often have levels of self-reactivity and abnormally long constant regions that require extensive somatic hyper mutation to achieve [133, 136, 144-150].

1.1.1.3 HIV Epidemiology

As of 2016, an estimated 35 million people have died from AIDS related causes since the beginning of the current epidemic in 1981 [108]. Roughly 30 million people are currently living with HIV [106-109]; a further 2 million people are newly infected each year [107-109]. This rate of infection has remained stable since 2010 [107, 108].

HIV is transmitted through contact with the bodily fluids of infected individuals, such as blood, semen, vaginal fluids, rectal fluids, or breast milk [151, 152]. Globally most transmission occurs through sexual contact [151]. Drug use is another key source of transmission, either directly through the sharing of needles, or indirectly through the promotion of risky sexual behaviour [18]. Though not a major topic of discussion in this thesis, due to the circumvention of protective mucosal barriers

infection through blood-to-blood contact, as with injection drug users, blood transfusions, or accidental needle pricks, does result in a larger number of transmitted founder viruses [153]. Transmission risks are higher per exposure by this route and range from 95:100-1:50, but are responsible for only a small proportion of HIV cases worldwide when compared to sexual transmission [154, 155]. Interestingly, though more viruses are transmitted in cases of injection drug use, there are still selection processes that result in a smaller number of founder viruses than would be expected from the transmission [153]. Additionally, though CXCR4 and dual tropic viruses may be transmitted, founder viruses are typically CCR5 tropic [153].

Vertical transmission of HIV from mother to child can occur through the placentae, in the transmission of fluids at birth, or through breast milk resulting in infection of roughly 30%-45% of exposed infants [18, 156]. Fortunately vertical transmission can be effectively prevented through the use ART [18, 156].

Sexual transmission rates per exposure event are low compared to other routes of infection and range, for heterosexual transmission, from 1:200-1:2000 for women to 1:700-1:3000 for men [154, 157, 158]. Anal intercourse has a higher risk of transmission – 1:20-1:300 per exposure event – but is responsible for a lower proportion of overall transmissions globally than vaginal intercourse [154, 158, 159]. Women are more vulnerable to infection in heterosexual transmission then men due to a variety of biological and social factors [154, 160]. Despite this, there is no overall gender bias in the global burden of HIV [109]. In the Americas, Western Europe, and Central Europe half of new infection are among men who have sex with men (MSM) [18, 109].

Once infection has occurred HIV cannot be cleared (See Section 1.1.1.2). The development and availability of ART has led to profound increases in the life expectancy of people living with HIV [161]. However, even when treated early in infection, HIV infected individuals on ART have higher systemic inflammation, and are at higher risk of all age-related diseases except cancer and severe neurocognitive diseases, despite life expectancies that are approaching that of the overall average [161, 162]. Even with ART there is never complete recovery from damage done early in HIV infection, though early interventions can preserve key immune cell populations and substantially benefit disease prognosis [126, 127, 163] (See Section 1.1.1.2). Only if ART treatment is initiated prior to and not after Fiebig III can CD4+ T cell levels, including Th17 cell populations, be maintained [126]. Importantly, early intervention can prevent the depletion of HIV specific CD4+ T cell populations leading to better control of the virus throughout infection [127]. On a global scale it is not feasible to diagnose all or even most HIV cases this early, but it points to the effect of early events and the benefit of early intervention.

Even in the chronic stage of HIV there is significant benefit to beginning ART as early as possible [163]. However, in 2015 only half of those infected with HIV had access to ART [109, 164]. AIDS related deaths have decreased globally from their peak in 2005 of 1.7-2 million per year, but in 2015 one million people died of AIDS related diseases [107-109, 164].

In 2011 the Joint United Nations Programme on HIV/AIDS (UNAIDS) proposed the New Investment Framework (IF) to reduce the global burden of HIV, calling for interventions targeting key populations, the promotion of condom use, behavioural changes in high risk population, voluntary medical male circumcision, investment into preventing mother to child transmissions, and increased treatment, care and support for those already living with HIV [165, 166]. Costs for the IF were estimated at an initial requirement of $22-24 billion United States Dollars (USD) annually [167]. As the global burden of HIV decreased these costs would decrease and the IF would become cost effective compared to the previous status quo in the following decades [167]. By 2014 however, global funding had reached only $20.2 billion USD annually [166, 167].

In 2013 the IF was replaced by the Investment Framework Enhanced (IFE). Treatment guidelines were updated alongside acknowledgement of new uses for antiretroviral drugs, and a call for investment into a HIV vaccine [166, 168]. The IFE formally acknowledged the merit of treatment as prevention, noting that ART treated individuals with undetectable viral counts are unlikely to transmit HIV, and called for more comprehensive use of ART for this purpose [166, 168]. The IFE also promoted pre-exposure prophylaxis (PrEP), the use of antiretroviral drugs in HIV negative persons from at-risk populations to protect against infection with HIV [160, 166, 168, 169]. Multiple trials of PrEP have demonstrated protective efficacy, showing prevention of HIV acquisition is closely linked to adherence to PrEP regimens in males [169-172]. However, PrEP efficacy is less consistent in females [173, 174], possibly due to interactions between PrEP and the vaginal microbiota [175, 176]. Women on PrEP with a non-lactobacillus dominant microbiota, and particularly those with higher levels of *Gardinerella vaginalis*, are less protected from HIV acquisition compared to controls than lactobacillus dominant women are [175]. *G. vaginalis* and other bacteria metabolize tenofovir reducing the uptake and effectiveness of the drug [175]. Without addressing microbial constrains, PrEP will be of limited value in a substantial portion of the population.

While mathematical modeling has suggested that PrEP in conjugation with other interventions could prevent 3 million new infections in sub-Saharan Africa over 10 years, this would require significant investment in PrEP availability alongside other prevention measures [160, 169]. To meet the IFE, global funding would need to expand to $31 billion USD each year without investment in PrEP and $50 billion USD with sufficient investment in PrEP included [166, 168]. Given the current US political climate it is clear that these targets are unlikely to be met.

In 2014 UNAIDS proposed the 90-90-90 campaign for 2020, calling for 90% of people living with HIV to know their HIV status, for 90% of those diagnosed with HIV to receive ART, and for 90% of ART recipients to achieve viral suppression [168]. If met, these ambitious goals are expected to reduce new HIV acquisitions per year to slightly over half a million per year by 2070, meaningfully decreasing but not eliminating the burden of HIV [168].

A vaccine against HIV has been sought as an ideal intervention against the epidemic. Successful vaccines are often thought about as all or nothing prophylactic measures that will prevent infection in a sterilizing manner. However, an effective vaccine against HIV can work via a number of mechanisms to prevent the harm of the virus or reduce transmission (**Figure 1-2**). A vaccine which provides sterilizing immunity (**Figure 1-2A**) by preventing the initial infection of HIV or eliminating the infection would halt disease progression and prevent transmission of the virus. Such a vaccine has proven difficult to develop (See Section 1.1.2 and Section 1.1.3.1), but would not need to be 100% effective to substantially alter the course of the HIV epidemic.

In the best case scenario of a successful 90-90-90 campaign a partially efficacious vaccine with 70% efficacy at preventing HIV acquisition introduced by 2027 would reduce the expected yearly infection rate by 2070 from 547 000 to 122 000 [16]. However, the yearly rate of new HIV infections has not fallen since 2010 [107, 108] and global funding levels have not yet been achieved that meet the IF, IFE, or 90-90-90 campaign [166-168]. In the reality that the IF is only 50% met, or remains unmet, the proportional economic and humanitarian benefit of a sterilizing vaccine against HIV is dramatically increased [166, 168].

As noted above, early events in HIV dictate the course of infection. Any treatment that is initiated after HIV is detectable will be unable to reverse damage done in the early Fiebig stages of infection [126] (See Section 1.1.1.2). PrEP, which is initiated prior to infection, may not be effective or logistically achievable in large subsets of the population [166-168, 175, 176], but a non-preventative vaccine that accelerated the immune response can produce substantial benefits to disease prognosis. If key immune responses involving CD8+ T cells or antibodies occur earlier they will limit the peak viral load (**Figure 1-2B**). Due to the rapid rate of HIV expansion in Fiebig I/II even a slight acceleration of the immune response may decrease the peak viral load by several logs (**Figure 1-2B**). A lower peak viral load affects the spread of the virus throughout the tissue, the size of the viral reservoir, and the set-point viral load [177, 178]. By curtailing the spread of the virus key populations of HIV susceptible immune cells, such as HIV specific CD4+ T cells and Th17 cells in the gut mucosa, can be preserved [126, 127]. Such a vaccine, though unable to eliminate infection, may prevent many of the HIV associated complication that cannot be recovered from despite ART [126, 127, 163].

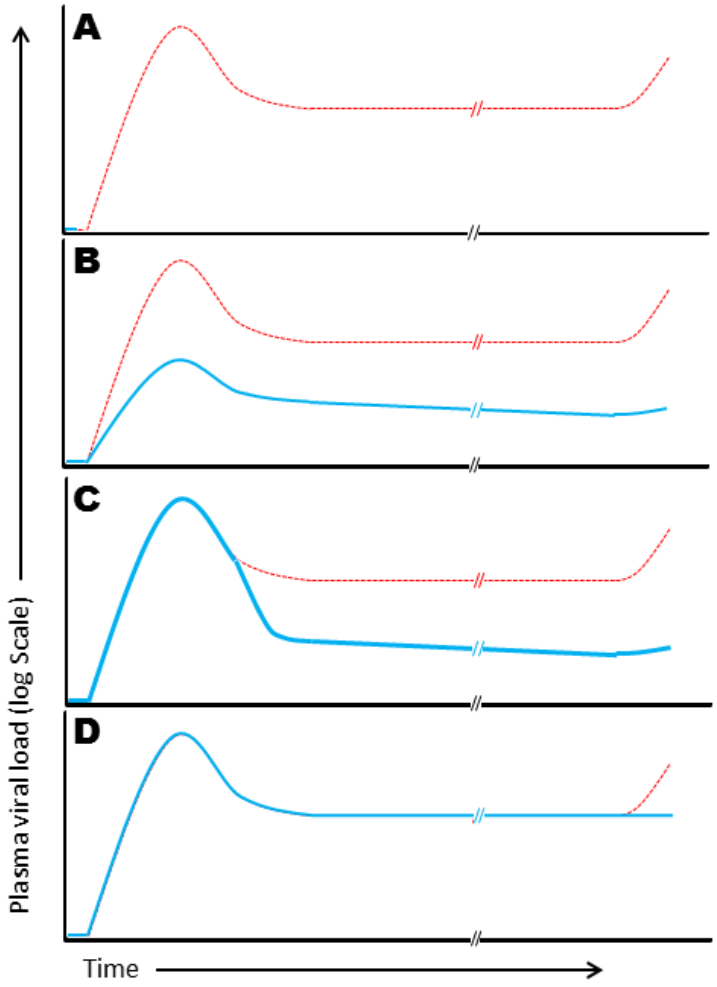


Figure 1-2: Protective mechanisms of a vaccine against HIV. An impactful vaccine against HIV can act by: (A) preventing initial infection with HIV or eliminating the founder virus before it can spread beyond a local infection; (B) lowering peak viremia; (C) lowering the set-point viral load throughout infection; or (D) preventing disease progression. Plasma viral load without intervention shown with red dotted line; plasma viral load with intervention shown in blue.

A vaccine which lowers the peak viral load, or independently lowers the set-point viral load (**Figure 1-2C**) will further impact the spread of HIV. Viral load is directly associated with HIV transmission rates [157, 179] and disease progression [177]. The highest rates of HIV transmission occur during the first three months of infection while viral load is still high. At this point HIV is twenty-seven times more likely to spread per sexual exposure than during chronic infection [158]. By

contrast, during late stage infection and AIDS, when the viral load is again increasing, the transmission rates per sexual exposure are only seven times higher than during chronic HIV infection [158]. A vaccine that stops disease progression will prevent disease morbidity and spread (**Figure 1-2D**). By lowering viral load, especially if it can be lowered below the point of detection, further HIV transmissions can be prevented, impacting the epidemic in the same manner as a preventative vaccine.

In all scenarios the development of a vaccine of any modality is the only measure that allows a sustainable approach to manage the HIV epidemic. The benefits of a vaccine against HIV are irreplaceable in the face of the current epidemic [168].

1.1.2 Obstacles to a HIV Vaccine

The creation of a sterilizing vaccine against HIV (**Figure 1-1A**) is complicated by the absence of confirmed natural clearance of the virus in humans or any model [180, 181]. There have been several human cases with extended periods of apparent HIV clearance, where the virus remained temporarily below detectable levels [182-184]. Additionally, in the cytomegalovirus (CMV) vaccinated macaque model some SIV infected animals have remained undetectable following transient CD8+ T cell depletion despite CD8+ T cells mediating control of the virus [185]. However, even studying these unusual cases, it has been impossible to completely decipher the factors involved in protective immunity against HIV.

No currently existing vaccine for any virus completely prevents infection, with the possible exception of rabies [180, 186]. Instead they prepare the immune system to control or eliminate infection while it is still localized. With HIV there appears to be a narrow window of opportunity for sterilizing immunity, when the virus can be completely eliminated. After vaginal inoculation it has been shown that SIV replicates locally and spreads to the draining lymph nodes within 4 to 5 days; after intrarectal inoculation virus can be detected in the draining lymph nodes within 4 hours [187]. Instances where ART was started within hours of exposure have been shown to decrease the infection risk by 81% [188]. In macaques the success of post-exposure prophylaxis depends on both time between exposure and commencement of ART and the duration of ART [189, 190]. Together these indicate that abortive infection is theoretically possible.

Ongoing infection with HIV/SIV may not lead to protective immunity from further HIV exposures. Instead subsequent exposure to a sufficiently diverse strain of HIV, even one of the same clade, can lead to superinfection with both strains [191-194]. Despite this there is evidence of immune protection against HIV and SIV infection.

Studies of HIV exposed uninfected individuals, particularly repeatedly exposed sex workers, have informed our understanding of natural susceptibility to HIV. A small subset (approximately 5%) of the Majengo female sex worker cohort in Nairobi, Kenya who were found to be HIV seronegative despite years of high-risk sexual activity and a high number of estimated unprotected HIV exposures were examined exhaustively for demographic, behavioural, genetic and immunologic risk factors [195-197]. Partial protection was found to be associated with a number of factors, including the presence of the class I human leukocyte antigen (HLA) A2/6802 supertype and the DRB1*01 allele among the sex workers [197].

Interestingly, HIV acquisition rates among the women in the cohort were highest in the first two year of high risk activity and after that the HIV acquisition rate decreased significantly [195]. After periods without regular exposure, such as if the sex workers took a break, there was an increased risk of infection upon resumption of sex work suggesting that exposure to HIV induced protection [195]. Protection was correlated with mucosal CD8+ and CD4+ T cell responses and HIV specific mucosal IgA, but not with plasma IgG or IgA [196]. A decrease in HIV specific CTL responses coincided with increased acquisition rates after periods without regular exposure [195]. That exposure led to short lived partially protective immune responses in the women provides hope that a sterilizing HIV vaccine may be possible.

However, to be successful a HIV vaccine does not need to be sterilizing (See Section 1.1.1.3; **Figure 1-2**). A non-sterilizing vaccine that maintains an immune response to fully supress viral replication despite a maintained viral reservoir would prevent the development of disease in currently infected individuals and the transmission of virus [198, 199] (**Figure 1-2C**). If a vaccine can lower the peak viral load, lower the set-point viral load, or halt the progression of disease it will reduce the transmission rate of HIV, benefit those who are currently infected, and will have a profound impact on the HIV epidemic [198] (See Section 1.1.1.3).

An effective, preventative HIV vaccine will need to overcome the high diversity and variability of the HIV virus. Similar to influenza, one vaccine antigen will not protect from the antigens of all HIV groups and subtypes [149]. Different HIV subtypes were originally geographically isolated [18]. Due to increased rates of global movement geographically disparate subtypes and recombinant forms are increasingly found at higher rates everywhere [18]. It is therefore no longer feasible to design different vaccines for different regions, instead antigen design in any potential HIV vaccine must accommodate for differences between clades and subtypes.

On top of high population diversity there is a rapid rate of adaptation within the host [44, 121, 200]. HIV has a high rate of replication, high rate of mutation, and high level of functional plasticity that allows HIV to mutate and rapidly escape immune control [44, 121, 200].

There are preserved sites on all the HIV proteins that incur higher fitness costs when mutated [40]. Selective pressures from an immune response will drive mutations in all HIV proteins, but mutations at these sites frequently revert upon removal of those pressures, such as when the virus is transmitted to a new host [40]. These sites are ideal vaccine targets, as they are well conserved between strains and clades. Gag, Pol and Nef contained the most preserved sites, and are ideal targets for T cell responses but due to where they are expressed cannot be the target of neutralizing antibodies [40] (See Section 1.1.1.1). By contrast, Env has few highly conserved sites visible to antibodies. HIV Env, and in particular Gp41, has the highest rate of mutation of any HIV protein [44].

The HIV virion is enveloped in the host's phospholipid bilayer, and coated in host proteins [4, 34]. Host surface molecules, including integrins and their ligands, aid in viral attachment to target cells, but, as self-molecules, cannot be targeted by the host immune response [34-37]. Env is the only HIV protein that is a viable target for neutralizing antibodies. As a result, neutralizing antibodies are produced during natural HIV infection but they are typically only narrowly neutralizing against a specific clonal lineage of HIV at a specific time [121, 201, 202]. Rather than controlling infection these neutralizing antibodies drive the viral swarm to adopt escape mutations [121, 201-203].

There is further difficulty in generating an effective antibody response against the conserved sites of Env. Glycosylation and a folded conformation protect the CD4 binding sites from antibodies [121, 149]. The Gp41 neutralization site of HIV Env mimics human antigens and is further protected from the development of binding antibodies by self-tolerance controls [149].

There are conserved non-neutralizing immune dominant epitopes on Env that attract antibody responses [121, 149]. Decoy immunogenic sites that are not exposed in the functional Env trimer are presented on shed Env and non-functional Env monomers are present on the virion [121, 149, 204]. Antibody targeted against these decoy sites do not neutralize the virus, though they may aid in opsonization. Additionally there are few functional Env trimers on the surface per virion [205, 206]. In the face of these obstacles the majority of antibodies developed throughout HIV infection will be transiently effective and only 20% of HIV infected individuals will develop antibodies at any measureable level that target conserved neutralization sites and are broadly neutralizing across multiple strains [121, 132-135] (See 1.1.1.2). This likely contributes to the degree of HIV superinfection discussed above. Further, BNAb that do develop are rarely effective in more than one clade of HIV [139, 207].

Although broadly neutralizing antibodies and bispecific antibodies have been able to show a synergy in terms of neutralizing capacity [139, 141, 207-209] and to demonstrate resistance to challenge *in vivo* [141, 149, 210-213], efforts to induce these monoclonal antibody combinations following vaccination have been limited primarily to transgenic mouse and small animal models [207,

214, 215]. No vaccine attempt has successfully elicited BNAb against HIV/SIV in vaccine recipients to date, a putative correlate of protection in most currently licenced vaccines for infectious diseases [149] (See Section 1.1.3.1). To truly succeed in preventing multiple subtypes of HIV without escape mutations a potential neutralizing antibody based HIV vaccine would need to induce development of multiple BNAbs [149]. Antibodies that react against the broad range of encountered HIV lineages should be mathematically favoured in antibody maturation [121]. Correspondingly, BNAb should be more likely to emerge in individuals infected with multiple viral founder strains [121]. However, the key neutralization sites on HIV are, as described above, protected from typical antibody responses and BNAb require extensive somatic hypermutation to target them [133, 136, 147-150]. By screening large numbers of HIV seropositive individuals for BNAbs, a small number have been identified and cloned that are highly effective, even across clades HIV [139, 141, 207]. This painstaking work, and evolutionary analysis chronicling BNAb development has yet to pay off in a multivalent BNAb vaccine, but offers a potential direction for vaccine research [121, 148, 207].

The CD4+ cells that are activated and recruited to sites of infection and inflammation, lymphocytes and macrophages, are targets cells of HIV [18]. Vaccine attempts that recruit activated CD4+ cells to the site of infection without other protective measures, such as CD8+ T cell responses or antibody production, risk increasing rather than decreasing the susceptibility of vaccine recipients to infection [32, 38, 216]. As discussed in Section 1.1.1.1, this increased susceptibility was observed in macaques vaccinated with a VZV based vaccine that activated CD4+ T cells without generating protective antibody or CD8+ T cell responses [32, 38]. Fortunately not all CD4+ T cells are equally susceptible to HIV infection [217]. Expression of integrins or their ligands, which may interact with Env or host proteins on the HIV virion surface, increase the susceptibility of CD4+ T cells to HIV infection [34-37] (See Section 1.1.1.1). Activated CD4+ T cells primed against some viruses are more susceptible to HIV than others, contraindicating the use of some viral vectors in HIV vaccine design [217] (see section 1.1.3.1).

Another obstacle to HIV vaccine development is the latent viral reservoir. It effectively limits efforts to prevent HIV to mucosal approaches at the primary site of infection, typically the genital tract and primary lymph nodes. Once the gut immune reservoir becomes involved, immune depletion occurs and a latent pool of virus becomes established (See Section 1.1.1.2).

It is difficult to estimate the size of the HIV latent viral reservoir [45] but any vaccine-induced immune response that does not completely prevent infection in a sterilizing manner must outlast the latent reservoir to prevent viral resurgence. The viral reservoir is established within days of initial infection [112] and wanes with an estimated half-life of 4.6 months to 44.2 months with ART [110, 111]. Stopping ART prior to complete clearance of the latent viral reservoir leads to viral rebound

[110, 111]. If vaccine-induced immune responses against HIV wane, or immune suppression occurs for any of a number of reasons, viral rebound will reoccur as it does in the macaque SIV vaccine model when anti-CD8 antibodies are administered following CD8+ T cell suppression of viremia [185, 218]. Any vaccination that controls disease, or eliminates infection after the establishment of the latent viral reservoir must maintain enough antigenic load to lead to a vigorous long-lived CD8+ T cell immune response. It is not known whether immunotherapy alone in the absence of ART can do this, or if immune exhaustion will eventually occur.

1.1.3 HIV Vaccine Strategies

Since the discovery of HIV as the causative agent of AIDS, there has been any number of proposed vaccine strategies. To test these, numerous non-human primate trials and a small number of human trials have taken place (See Section 1.1.3.1).

Although antibodies, specifically IgG or IgA, are the predominant mechanism of protection in most vaccines [219], T cell responses play an important role in the smallpox and yellow fever vaccines where CTL responses are correlated with protection [220]. Only in a few vaccines, the bacilli Calmette-Guerin (BCG) tuberculosis vaccine and VZV_{Oka} vaccine for chickenpox and zoster, is protection mediated by cellular immune responses [219]. Not surprisingly, the first HIV vaccines attempts were simple Env subunit-based vaccines. Single Env subunits were used to generate anti-Env antibody responses with no success towards preventing HIV acquisition [221] (See Section 1.1.3.1). In retrospect, these vaccines did not generate anything near conformational-intact antigen and simply induced more of the same non-functional decoy Env monomers that were being blebbed off T cells during HIV infection (See Section 1.1.2).

Some might argue that no vaccine has ever completely prevented viral infection (See 1.1.2), and that any attempt to do so would require abundant neutralizing IgG and IgA at the portals of entry [180, 181, 186]. Env evolves quickly and while there are highly conserved sites, they are heavily glycosylated or conformationally masked (See 1.1.2). So a successful humoral immune response will require multiple BNAbs that target key conserved sites across all potential founder viruses, and will require a sufficiently immunogenic antigen achieve sufficient antibody titres [149].

Potentially vaccine-achievable concentrations of infused BNAbs can protect rhesus macaques from vaginal and intrarectal challenge with Simian/Human Immunodeficiency Virus (SHIV) [141, 149, 211, 212]. In pigtailed macaques sterilizing protection from SHIV is achieved against high dose (75ID50) $SHIV_{DH12}$ when given 100 times the 50% inhibitory concentration of antibody 24 hours before challenge [213]. It is unknown, however, how to induce the development of potent BNAbs with a vaccine. There has been recent focus on creating a better vaccine antigen to mimic the Env

trimer with the non-neutralizing immunodominant epitopes removed or hidden [222, 223]. To generate a BNAb response it is likely that the vaccine will require multiple diverse antigens [121]. The most potent BNAbs have undergone extensive somatic hypermutation, have levels of auto-reactivity, long CDR3 loops, and other features of secondary mutation that make them difficult to induce [133, 136, 144-150]. Further, to develop the known BNAbs specific initial germ line alleles in naïve B cells are required [149].

Recently the development of some BNAbs has been studied to determine intermediate, non-broadly neutralizing antibodies along the BNAb developmental lineage [148]. It may be possible to induce BNAb production through a stepwise series of vaccine boosts with antigens designed to produce successive intermediate antibodies [133, 149]. If this is possible, it will still be necessary to ensure the induced BNAb of the proper antibody subtype are produced, and at a high enough titre to protect from a potential HIV exposure years after vaccination.

Producing known BNAbs as transgenes, rather than attempting to induce novel responses in each individual, has been suggested an alternative approach. Cloned BNAbs with the capability to neutralize more than 90% of circulating HIV strains can currently be synthesised in large quantities by hybridomas or in plant cells [94, 141, 211, 212]. These can be transfused to protect from infection [94, 141, 211, 212]. Regular antibody infusions are not practical as a long-term preventative approach, but BNAb production can be induced directly rather than developed as an immune response. Adeno-associated virus (AAV) has been used to deliver the DNA to produce BNAb into muscle tissue [94, 224, 225]. In rhesus macaques this approach has resulted in circulating anti-SIV BNAb levels as high as 270ug/mL blood [225]. In most animals high levels of host-produced antibodies target the induced antibodies, limiting the anti-SIV BNAb titres achieved [225, 226]. In an attempt to limit host responses against the transgene antibodies, rhesus rhadinovirus (RRV) – a simian herpesvirus related to Kaposi's sarcoma-associated herpesvirus (KSV) in humans – was modified to encode two anti-SIV BNAbs for production in virally infected B cells [226]. Infection with the modified RRV induced more modest antibody levels in macaques of 200-300 ng/mL blood, but still resulted in notable anti-antibody responses [226]. There are further technical and ethical complications of the transgene approach, as the transgenic BNAb production cannot be stopped and there is risk of immunogenicity against the vector, the transgenic BNAbs, or the tissue producing them [94].

Attenuating viruses to produce vaccines has resulted in several of most effective and impactful vaccines worldwide, including vaccines against smallpox, rabies, and yellow fever [220, 227]. Traditionally, attenuated viral vaccines were produced by passaging the viruses repeatedly in cell culture to yield a culture adapted virus that, as a consequence, had lost many of the

virulence genes necessary to produce disease in the host [227]. In the modern era where the entire viral genome is known and the functions of all viral genes are understood, this attenuation process can be aided through modern molecular techniques. Since HIV is a small virus with a simple genetic structure compared to smallpox or yellow fever, it initially appeared to be a good candidate for targeted molecular attenuation. Efforts targeted *nef* after a partial ΔNef wildtype strain was shown to be associated with HIV non-progression in a cohort of blood transfusion recipients in Sidney Australia (all of whom later progressed) [88, 89].

Although impressive results demonstrating protection were initially seen in macaques vaccinated with attenuated ΔNef SIV, there were concerns regarding disease progression using the existing constructs and the potential for pathogenicity resulting from the vaccine [90-93]. When the level of attenuation was increased to address these concerns, only half of the macaques vaccinated with SIVmac239Δ3 were protected from challenge with the related SIVmac251 swarm [90-93]. Furthermore, this protection was short lived and did not extend to other strains of SIV [90-93]. However this trial, and other similar attenuated SIV vaccine trials, were among the first to demonstrate that sterilizing immunity was not required to protect from SIV disease [92].

Safety has remained the biggest roadblock to attenuated HIV vaccines. Vaccine efficacy against future challenge in these attenuated trials was inversely proportional to the level of attenuation [91, 92]. Furthermore, there always remained the possibility of the attenuations reverting and the vaccine strain becoming more virulent. The virus is capable of repairing altered genes, possibly from random mutation and selective pressure, or possibly through recombination with endogenous retroviral DNA in the host genome [11, 40, 44, 92]. Thus it is currently accepted that there remains an unacceptably high risk of attenuated vaccine virus reverting to a virulent state and causing disease [92]. This was demonstrated in studies including a longer-term follow-up of animals receiving attenuated SIV vaccines, where the macaques eventually developed symptoms from even the more attenuated vaccine strain viruses. More than half of SIVmac239Δ3 vaccinated macaques developed symptomatic T cell depletion by 6.8 years post vaccination, and 18% had progressed to simian AIDS [228].

A safer alternative is available in the use of viral vectors to deliver vaccine antigens. First human and, more recently, non-human primate adenoviruses have been studied as potential HIV vaccine vectors. Over the past twenty years a variety of poxviruses have been examined in-depth, first with Modified Vaccinia Ankara, and then the slightly less virulent strain of New York Vaccinia [229]. Most recently attenuated Canarypox has been utilized as a vector in multiple Human HIV vaccine trials (See section 1.1.3.1).

The poxviruses produce robust cellular responses in both the CD4+ and CD8+ T cell compartments [230-232]. Poxvirus generated humoral responses, in contrast, can be weak and are best generated by pairing a poxvirus boost with a paired vaccine prime [229, 233, 234]

In general vector based approaches focus on generating cellular immune responses in the form of both CD8+ and CD4+ effector T cells. CD8+ effector T cell responses are traditionally thought of as cytolytic, but increasingly it has been accepted that robust CD8+ effector T cells are multifunctional [235-237]. They secrete a variety of cytokines, and more than one lytic molecule [235-237]. CTL are the earliest adaptive immune responders against HIV in natural infection of the submucosa (See Section 1.1.1.2). The timing and size of the initial wave of CD8+ effector T cells is thought to alter the ultimate size of the viral reservoir and to determine the extent of CD4+ cell depletion in the gut [122, 238-241]. A rapid onset of robust CD8+ T cell responses early after infection is associated in animal models with a lower set-point viral load, less bacterial translocation throughout infection, and less inflammation throughout disease [126, 127, 177, 178, 238, 240, 241]. CD8+ T cell responses are active throughout the mucosal tissue of the body, especially in the gut, but other than those found at the site of primary infection they probably appear too late to prevent the establishment of infection and entrenchment of a substantial latent viral reservoir [148, 180]. CD8+ T cells target virally infected cells that are actively producing viral proteins, but not latent cells [148]. In most HIV infected individuals CD8+ T cell responses are unable to clear infection or fully supress viral replication by the time they occur. Despite this, strong CTL responses are still correlated with improved disease outcome. Long-term non-progressors have more active CD8+ T cell responses than HIV progressors [242, 243].

Pre-existing CD8+ T cell responses, alongside CD4+ T cell responses, are correlated with protection from infection in a repeatedly exposed, high risk population of sex workers that remained seronegative for HIV [195] (See section 1.1.2). A robust, immediate CTL response at the portal of entry may be capable of preventing HIV infection before it can be established. This idea led to the development of effector memory T cell (T_{EM}) based vaccines. CD8+ and CD4+ T_{EM}, alongside tissue resident memory T cells (T_{RM}) which are often not differentiated from T_{EM} in non-human primate vaccine models (See Section 1.2.2.2), act immediately without secondary signals or trafficking to lymph nodes [185, 187]. A maintained population of circulating T_{EM} and T_{RM} located at portals of entry in the mucosa are hypothesized to pre-empt the establishment of a substantial latent viral reservoir and immediately control infection to mitigate the acute phase of HIV replication, minimize the HIV set-point viral load, and potentially clear virus before the establishment of a large latent viral reservoir though they do not protect in a sterilizing manner [185, 187].

Frequently reactivating, lifelong infection with herpesviruses maintains robust T_{EM} and T_{RM} populations (See Section 1.2.2.2) that make them ideal as vaccine vectors for a T_{EM}/T_{RM} based HIV vaccine approach (See Section 1.3.2). A number of herpesviruses have been explored as possible vaccine candidates, including CMV, VZV, KSV and Herpes Simplex Virus (HSV), but the most promising data to date is from CMV and VZV vectored vaccines. Roughly 50% of male rhesus macaques vaccinated with a CMV based SIV vaccine controlled the virus after they were infected when challenged intrarectally [185, 244]. Similarly, 9 out of 16 female rhesus macaques (and none of 18 controls) infected through an intravaginal challenge with SIV after vaccination with the same vaccine controlled the virus [187]. Further 2 out of 6 rhesus macaques given low dose *intra venous* (IV) challenge controlled viremia [187]. In each case all macaques were infected with SIV following challenge before a portion of the vaccinated macaques controlled the infection at plasma viral loads below detectable levels with occasional spikes in viral load [244]. Protection of the macaques was not associated with MHC haplotype or TRIM5α, but potentially mediated by robust CD8+ T_{EM}/T_{RM} biased responses against SIV [187, 200, 244]. Due to mutations in the CMV vector all vaccinated macaques generated CD8+ T cells responses to MHC-II and MHC-E presented epitopes, which it has been suggested may also be involved in some manner in the protection seen in half of the vaccinated animals [200, 245] (See Section 1.3.2). Interestingly, despite the inclusion of Env in the vaccine the animals had no detectable antibody against it, a fact the investigators were unable to explain.

These protected macaques exhibited transient increases in SIV viral load to just above the limit of detection, which the authors termed "viral blips". The SIV viral load occasionally jumped to low detectable levels for a single test before returning to undetectable throughout the first year [244]. After 30 weeks these "viral blips" decreased in frequency and were not observed after 52 weeks in the male macaques or after 70 weeks in the female macaques [187, 244]. SIV infection spread to multiple tissues and SIV infection was present in the gut mucosa, bone marrow, lymph nodes, and cervix of female macaques for the first 9 weeks then undetectable afterwards [187]. Later using anti-CD8 antibodies, CD8+ T cells were briefly depleted in protected macaques without SIV viral rebound leading to the possible conclusion that the macaques did not have an accessible viral reservoir and were thus "cured" [185, 187] (see section 1.3.2).

The CMV vector-based vaccine provided protection regardless of route of exposure – vaginal, intrarectal, or IV – indicating that the maintained presence of T_{EM}/T_{RM} populations may be sufficient to protect from HIV/SIV without antibody help [185]. However, why only half of vaccinated animals were protected remains unknown – as do the exact correlates of protection from these CMV based SIV vaccines [185, 244]. A number of unusual features of the vaccine vector, rhesus macaque CMV

(RhCMV) 68-1, may be involved in protection (see section 1.3.2.1) but these effects applied equally to protected and non-protected vaccine recipients [200, 245].

It is likely that ultimately a successful HIV vaccine will use both CTL and antibody mediated immunity, although not necessarily with neutralizing antibody. The most successful human vaccine trial, RV144, combined two individually ineffective vaccines, one a CD8+ T cell based approach and one an antibody based approach [246]. The development of better T_{EM}/T_{RM} based vaccines, alongside independent BNAb based approaches, may generate the synergistic approach that ends the HIV epidemic.

1.1.3.1 Human HIV Vaccine Trials

Only six vaccine phase IIb (looking for efficacy greater than 0%) or phase III (looking for efficacy greater than 30%) efficacy trials targeting HIV have been published: AIDSVAX B/B in 2005, AIDSVAX B/E in 2006, HIV Vaccine trials Network (HVTN) 502 (Step study) in 2008, HVTN 503 (Phambili) in 2011, RV144 (Thai Trial) in 2012, and HTVN 505 in 2013 [46, 221].

AIDSVAX B/B, which ran 1998-2003, and AIDSVAX B/E which ran 1999-2003, attempted to induce a humoral response against recombinant gp120 Env glycoprotein [221]. The limited antibody responses developed were to portions of Env not typically exposed in the functional Env trimer and the vaccines were not protective [221].

HVTN 502, HVTN 503, and HVTN 505 all used a human adenovirus type 5 (Ad5) vectored vaccine. Despite high levels of protection observed against SHIV86.9P in vaccinated rhesus macaques, protection in macaques against the more virulent SIVmac239 was less consistent [46, 247-252]. The human AD5 vaccines ultimately failed to protect against HIV acquisition [46, 253]. HVTN 502 ran as a Phase IIb vaccine efficacy study in 18-45 year old high risk populations across the Americas and Australia from 2004-2007 [46]. Vaccine recipients were given 3 replication deficient AD5 vectors expressing HIV Gag, Pol, or Nef respectively in a 1:1:1 ratio with boosts at week 4 and week 26 [46]. The three antigens were chosen due to their high level of conservation across clades, favourable expression kinetics for vaccination, and their targeting during natural HIV infection [46] (See Section 1.1.1.1). As the mechanism of protection in AD5 vectored vaccines is T cell derived and antibody responses were not expected against the antigens, the highly variable env was not included [46].

Despite immunogenicity in the phase 1 trial, and efficacy in the SHIV challenged macaque model HVTN 502 was not successful [46, 247]. The vaccine induced Interferon-gamma (IFNγ) elispot responses, but did not prevent HIV acquisition or lower set-point viral load after infection [32, 46]. Instead a greater number of infections were observed among a subset of vaccine recipients then those

receiving the placebo [32]. Self-reported risk behaviour was not different between vaccinated and unvaccinated individuals [32]. There were too few infections to examine risk factors in women, but risk of HIV infection in vaccinated men was increased if the men were uncircumcised, AD5 seropositive prior to vaccination, or both [32, 46]. There were no vaccine recipients at increased risk as compared to non-recipients who were AD5seronegative prior to vaccination and circumcised [32, 46].

The correlation between previous AD5 seropositivity and increased risk of infection following vaccination may partially explain the difference between the macaque model and human trial. While human AD5is not a natural pathogen of macaques, ⅔ of North Americans, and a higher rate of South Americans and South Asians, are seropositive [46]. Human anamnestic expansion in response to the vaccine vector in previously seropositive vaccine recipients recruited more target cells susceptible to HIV infection, something that would not have happened in macaque's who are by definition naïve to human AD5 [32, 46]. A subsequent trial pre-exposing macaques to host-range mutant AD5 prior to vaccination with an AD5 vectored SIV vaccine replicated increased susceptibility to penile SIVmac251 exposure [254]. Possibly because natural adenovirus infects the mucosa, activated AD5 specific CD4+ T cells have since been found to be more susceptible to HIV infection than T cells specific to many other viral infections [217]. AD5 specific T cells adopt a Th17 like pro-inflammatory cytokine profile that increases the AD5 specific CD4+ cells' susceptibility to HIV, and increased α4β7 expression that traffics the cells to sites of vulnerability [217]. Those cells may further interact with ligands on the virion surface to aid viral binding [34-37, 217]. By contrast CMV specific CD4+ T cells are less susceptible than most CD4+ T cells to HIV infection [217]. Infection risk was observed to decrease among AD5 seropositive vaccine recipients gradually over 18 months post vaccination, potentially as vaccine induced HIV specific CD4+ T cells target cells were no longer as prevalent at viral portals of entry [32].

As a result, the other AD5 trial, HVTN 503, which began recruiting in 2007, was cut off prematurely as the HVTN 502 trial ended. This trial used the same vector and appeared to show no decreased risk of if infection among vaccines, though these results are confounded by the premature end to the trial before the enrolment phase was completed [253].

HVTN 503 used a DNA prime recombinant AD5 vaccine in high-risk circumcised men and male-to-female transgender women with AD5 neutralizing antibody titres less than 1:18 [253]. The DNA prime, delivered by bioinjector in the deltoid at weeks 0, 2 and 8, consisted of 6 circular double stranded DNA (dsDNA) plasmids in a 1:1:1:1:1:1 ratio: HIV clade B gag, pol, nef and env, HIV clade A env and HIV clade C env [253]. This was followed at week 24 with an intramuscular vaccination with replication defective recombinant AD5 vectored Clade B gag-pol fusion, clade A env, clade B

env, and clade C env in a 3:1:1:1 ratio respectively [253]. The AD5 vector differed from that used in the previous trials. A number of deletions to the AD5 genome removed key immunogenic sites to reduce vector immunogenicity [253].

There was no increased risk of HIV among vaccine recipients compared to controls, and strong immune responses against HIV [253]. CD4+ and CD8+ T cell responses were similar to other AD5 vectored HIV vaccine immunogenicity trials [253]. IgG and IgA were generated against HIV gp140. IgG binding the V1-V2 loop of gp140 however, a correlate of protection in the RV144 trial (See below), was at lower levels than in that vaccine [253]. Gp140 binding IgA, which was higher in HVTN 503 than in RV144, was similarly correlated with a lack of protection in RV144 [246, 253].

At the HVTN 503 midpoint analysis 27 vaccine recipients and 21 controls had been infected with HIV, demonstrating no statistical difference in protection between the groups [253]. There was no difference between set-point viral loads in the HIV infected of either group and the trial was cancelled due to lack of protective efficacy [253]. As a result human AD5 may no longer be seen a viable option for development as a vaccine vector against HIV.

The human phase III trial, RV144 is generally considered the most successful human HIV vaccine trial to date. The vaccine was given to 18-30 year old HIV seronegative men and women in Thailand. Vaccine recipients were vaccinated and boosted 3 times with ALVAC-HIV, a replication deficient canarypox containing env, gag, and pol, and two boosts with the AIDSVAX B/E vaccine that had previously failed to protect when used alone [255] (See Above). The trial reported a protective efficacy of 31.2% over 42 months, with efficacy dropping afterwards over time [246, 255]. Accusations of mishandling or misinterpretation of the statistical analysis of the trial led to some conflict. Gilbert et al. argued that when proper Bayesian analysis was conducted there remained only a 78% chance the vaccine had protective efficacy against HIV, well shy of the standard threshold of statistical significance [221].

CD4+ and CD8+ T cell effector responses were not correlated with protection from HIV [246]. Instead, protection was observed to be correlated with IgG antibody binding to the V1-V2 loop of Env [246]. Genetic analysis of HIV in infected vaccines showed 89% of breakthrough infection came from strains with mutations in the V2 loop [256]. IgG avidity, neutralization capacity, and level of antibody dependant cellular cytotoxicity (ADCC) were not statically different between protected vaccine recipients and unprotected vaccine recipients [246]. Vaccine recipients did not develop BNAb, but a vaccine boost after 6-8 years did boost antibody production and lead to increased antibody somatic hyper-mutation and recombination [147].

Protection in RV144 was inversely correlated with titres of Env binding IgA in the serum [246]. Vaccine recipients with higher Env binding IgA titres were at higher risk of HIV infection, though

increased IgG or IgA levels in unvaccinated controls were not correlated with protection or risk [246]. Neither IgA nor IgG titres correlated with altered disease progression following infection [246].

No vaccine trial has been wholly successful, but numerous conclusions can be drawn from the HIV vaccine trials to date. The choice of vaccine vector is of prime importance. Pre-existing immune responses to the vector or simply the nature of the immune response against the vector can predispose recipients to increased risks of infection, as with human AD5 [217, 254]. CMV, which by contrast provokes immune cells less susceptible to HIV may be a better choice although there is currently no human data, and the mechanism of protection is not understood suggesting caution [217]. Sterilizing immunity and neutralizing antibody responses are not necessary for protection against HIV, which is important knowledge: non-neutralizing antibody responses were protective in the RV144 trial [246]. As seen by increased susceptibility to HIV of AD5 CD4+ cells and correlation of higher titre of Env specific IgA with risk of infection, the nature of the vaccine-induced immune response is of more importance than the magnitude [217, 246]. A successful HIV vaccine will likely include a synergistic humoral and cellular immune response, which may come from optimizing strategies in both arenas.

1.2 Cytomegalovirus

Primates have evolved alongside their microbiome and virome. Viruses of Herpesviridae, a family of viruses with large dsDNA genomes, are among the oldest constituents of the mammalian virome. Herpesviruses were present in primates 70 million years ago (mya) [257] and have since evolved and co-speciated with their mammalian hosts.

Unlike newly emerged viruses, highly prevalent herpesviruses like CMV have an evolved detent with their host immune system [258]. The relationship between CMV and its host can be seen as symbiotic and part of the natural virome.

CMV is a member of the betaherpesvirus subfamily of herpesviridae. The virus is asymptomatic in normal infection and integral in shaping the host immune response [259]. There are differences in gene regulation and immune activation between CMV seropositive and seronegative individuals at all ages [259-262]. Chronic CMV infection alters the host immune profile and promotes inflammation [259, 260, 262]. These changes generate a protective phenotype that protected mice against subsequent challenge with a variety of diseases, including *Yersinia pestis*, influenza, and the retrovirus friend virus [259, 262].

Human CMV (HCMV) prevalence ranges geographically from 45% in parts of North America and Europe to near 100% prevalence in parts of the developing world [263]. In the developed world 0.7% of infants are born with HCMV [264]. By age 11, a third of all children in the United States of America (USA) are HCMV seropositive [265-267], and by age 50 seroprevalence has risen to 60%

[263, 266, 267]. Seroprevalence is correlated with ethnicity. Rates of HCMV seroprevalence in the USA are lowest among Caucasians and highest in Black and Hispanic populations [263, 265-267].

CMV is spread vertically through placental transfer in utero, or by breast milk [265, 268-271]. Most HCMV seropositive women shed HCMV in breast milk after delivery [265, 269, 271]. Roughly 10% of infants are HCMV seropositive at 1 year of age primarily due to this shedding [265]. Horizontal spread of CMV occurs through close contact and bodily fluids. HCMV is shed from epithelial surfaces in bodily secretions at up to 10^5 infectious particles/mL of fluid [265, 272]. Infected infants shed HCMV in their urine and saliva at high levels for the first 4 years of life, with the volume of shed virus decreasing over time after 6 months until eventually ceasing in early childhood [265, 272]. Countries like Sweden, Japan, and Taiwan with widespread use of daycare centers have higher population HCMV seroprevalence levels, counter to the typical trend of low HCMV rates in developed countries [263].

HCMV is commonly acquired prior to adulthood. Later in life HCMV transmission can occur sexually, through genital secretions, or medically through blood products and solid organ transplants, presenting health risks for recipients [265, 273, 274]. Prior infection with CMV is not protective against superinfection at subsequent exposures; multiple strains can concurrently infect one individual or alternatively recombine [275-277].

1.2.1 CMV Virology

CMV is a large, enveloped virus with around 200 distinct genes [278-282], and multiple surface glycoproteins [283, 284]. Around 60 viral proteins and greater than 70 host cellular proteins are carried in the virion tegument between the envelope and viral capsid to be released into infected cells upon entry [284, 285]. An icosahedral capsid of 200nm encapsulates a single copy linear dsDNA genome of 200-250 kbp, the largest of any known herpesvirus [286]. During genome replication the virus circularizes and replicates as a rolling concatemer [283, 287].

The genome is organised into long and short segments, each with flanking repeat regions (**Figure 1-3**). The long region contains a terminal repeat long (TRL), a unique long (UL), and an internal repeat long (IRL) region. The short region is comprised of an internal repeat short (IRS), a unique short (US), and a terminal repeat short (TRS) region. The IRL and IRS regions are absent in some primate CMVs, notably the sequenced African green monkey CMV (GMCMV), RhCMV and cynomolgus macaque CMV (CyCMV) viral genomes [278, 288-290].

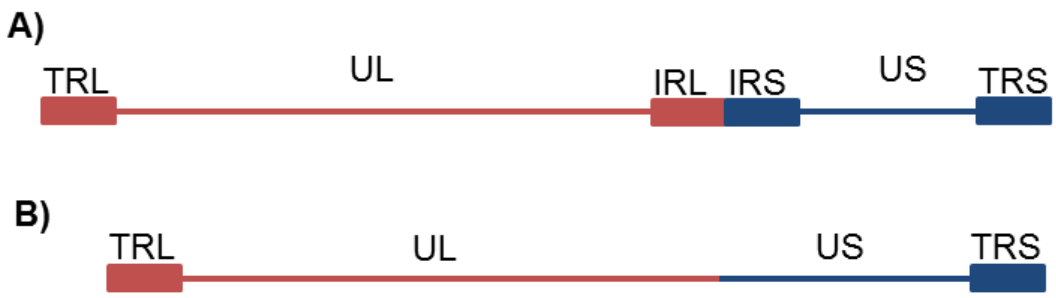


Figure 1-3: Cytomegalovirus genome organization with (A) and without (B) IRL and IRS regions.

1.2.1.1 CMV Replication Cycle

CMV strains are highly species specific and rarely, if ever, transmit between even closely related species, a possible consequence of lengthy co-evolution and co-speciation with their hosts (See Section 1.4.4) [268, 278, 291-294]. Within a host CMV are quite promiscuous and infect most cell types: endoderm, ectoderm or mesoderm [295]. CMV replicates in peripheral blood mononuclear cells (PBMCs) of bone marrow, the connective tissue of organs and in the smooth muscles of the gastrointestinal tract [295-298]. Lymphocytes and polymorphonuclear leukocytes are the only notable cell types that are thought to be non-permissive to CMV replication, though both cell types are abortively infected and play a role in viral dissemination [295, 299-302].

Entry into different cell types is mediated by a number of glycoproteins on the surface of the enveloped CMV virion [283, 303, 304]. CMV viral proteins UL128, UL130 and UL131 compete with glycoprotein O (gO) to complex with the heterodimer of glycoprotein H (gH) and glycoprotein L (gL) in the endoplasmic reticulum of infected cells during viral replication [283, 303]. Both the gH/gL/gO heterotrimer and gH/gL/UL128/UL130/UL131 pentameric complex are expressed on the virion surface and are responsible for receptor binding and cell entry [303, 305]. The pentameric complex is essential for entry into endothelial and epithelial cells through endocytosis or macropinocytosis of the virion into an endosome and pH dependant membrane fusion [303, 305]. The gH/gL heterodimer with or without gO on the virion surface is necessary for infecting fibroblast cells [304, 305]. Glycoprotein B (gB), a highly conserved envelope glycoprotein across all alpha-, beta- and gamma- herpesviruses, mediates membrane fusion for CMV to enter the cell [304, 305].

Upon membrane fusion a number of tegument proteins are released into the cytoplasm of the newly infected cell, some of which lead to the trafficking of the viral capsid into the nucleus using the host microtubule system [283, 306, 307]. The linear viral genome circularizes in the nucleus, where viral genes are transcribed into mRNA and the viral genome replicates [274, 283]. Regulation of temporal gene expression, for immediate early (IE), early, early-late, late, or latency associated genes

is initially controlled by virion transported tegument proteins, and later CMV expressed proteins [274, 283, 308]. A number of CMV genes are expressed throughout latent infection, though most genes are supressed [309-311]. During lytic infection viral mRNA is transported from the nucleus and translated by host machinery [274, 283].

Viral capsid proteins are produced as late genes and assembled in the nucleus [274, 283]. Viral DNA-containing capsids are trafficked from the nucleus to the viral assembly complex in the cytoplasm where they bud, along with tegument proteins, into intracellular vesicles coated internally with virion glycoproteins to become vesicle contained enveloped virions [283, 284]. The vesicles are then trafficked to the cell surface and infectious particles are released from the cell to complete the viral lifecycle [283].

1.2.2 CMV Pathology

1.2.2.1 CMV Infection

CMV results in lifelong infection with continuous low-level viral replication [309, 312]. CMV can replicate in most cell lines, and disseminates throughout the host both by direct cell-to-cell transmission and as free virus [276]. Different glycoprotein complexes on the viral envelope are used for entry into different cell types [303-305, 309] (See Section 1.2.1.1). Non-pathogenic infection with virus production is established in endothelial cells, particularly vascular endothelial cells, and PBMCs [296-298]. The virus establishes a latent reservoir primarily in undifferentiated myeloid cells, other hematopoietic stem cells, and monocytes [298]. Differentiation of latently infected monocytes into macrophages is suspected to be a key trigger of reactivation [308].

In otherwise healthy individuals HCMV is asymptomatic [309, 313]. In the adult population HCMV associated disease is linked to depression of the adaptive immune response in the host, such as with organ transplant recipients or AIDS patients [309, 313]. The induction of cellular immune responses alone against HCMV may be sufficient to prevent virus associated complications following organ transplantation [309]. In congenital infection robust cellular and humoral responses are both required to prevent virus-associated morbidity [309, 313].

HCMV is the most frequent congenital infection worldwide, causing virus associated complications in 30 000 to 40 000 infants in the USA each year [265, 314]. As virus associated complications are only clinically observed in tiny portion of congenitally infected infants, many more infants are infected without observable complications. Neonatal HCMV shedding, traditionally regarded as a marker for congenital infection at birth, is detected in 0.25 to 2% of all newborns worldwide [265, 314].

The HCMV foetal transmission rate is 40% when maternal primary infection occurs [265, 315]. There is still controversy surrounding the degree to which a preconceptual maternal immune responses to CMV limits congenital infection [316]. Up to a third of congenital CMV infections occur in cases where the mother was infected with CMV prior to conception, and HCMV can transmit to the foetus without the need for secondary infection with a new strain of HCMV [265, 315, 316]. However, HCMV maternal transmission rates and outcomes are both linked to whether the mother is naïve to HCMV. Multiple studies have demonstrated that the rate of HCMV transmission to a foetus is at least three-fold lower in CMV re-infected pregnant women then when it is a primary infection, and that the rate of sequelea is even further reduced [316-319]. If symptomatic infection of the foetus or neonate is diagnosed, treatment with antivirals such as ganciclovir can be instituted to reduce the risk of sequelea and prevent further disease progression, but not reverse existing damage [265, 320, 321].

In untreated congenital infection the maternal immune response including maternal IgG, the timing of viral transmission to the placenta and the foetus, and cellular tropism of the transmitted virus are key determinants of foetal outcome [265]. Viral associated complications occur in only 5-13% of congenital infections [264, 265, 322]. Sensineuronal hearing loss, the most common complication, occurs in 30-50% of symptomatic congenital infections while 10% of symptomatic congenital infection result in sensineuronal vision damage [264, 265]. Other complications including symptomatic abnormalities or calcifications in the brain, developmental or cognitive delays, and cerebral palsy are rare [265]. Without more studies incorporating long term follow-up, it is difficult to estimate how many congenitally infected infants born without symptoms will develop sensineuronal damage later in life [264].

Post-natal neurologic infections with HCMV are thought rare, and unheard of in immunocompetent patients [323]. In children or adolescents with impaired immunity neurologic infection can be severe leading to infectious mononucleosis, encephalitis, myelitis, or Guillain-Barre syndrome [265, 324, 325]. In addition to neurologic infection, immune supressed individuals, including AIDS patients, are at risk of CMV retinitis, cochlear infection, and potentially fatal organ disease [265, 323, 326]. The antiviral activity of standard ART therapy in AIDS patients reduces, but does not eliminate, the risk of each of these complications [323].

Unlike many other herpesviruses, CMV is not linked to cancer. The virus lacks any transforming function, and though CMV has been found in malignant tumours in the central nervous system this is due to opportunistic infection of the tumour rather than a causal relationship [327].

As a widely prevalent virus with observed pathogenicity there is great interest in a vaccine to prevent against congenital HCMV infection and HCMV caused disease, whether or not the vaccine

protects against HCMV infection outright [328]. Previous vaccine attempts have most often used the attenuated Towne strain of HCMV, which was developed in the 1970's [329]. Inoculation with HCMV Towne did not prevent infection with subsequent strains of CMV, but did offer protective efficacy against disease in renal transplant patients [328, 330, 331]. In CMV seronegative mothers, HCMV Towne does not prevent the acquisition of further CMV strains [328, 331].

1.2.2.2 Immune Response against CMV

Initial innate immune responses against CMV are triggered by resident PRRs. Toll-like receptor (TLR) -2 on the host cell membrane responds to CMV gB and gH activating NF-κB (Nuclear factor kappa-light-chain-enhancer of activated B cells) [309, 332]. After viral entry, TLR-9 responds to intracellular viral DNA [333]. Pro-inflammatory cytokines attract antigen presenting cells (APC), phagocytes and NKCs to the sites of viral replication [309]. The inflammatory cytokine milieu is partially countered by CMV, which induces production of numerous anti-inflammatory and immune evasion viral genes in infected cells (see below). NKC in particular are key in controlling CMV and are the targets of CMV immune evasion genes like anti-inflammatory IL-10 viral homologues *UL111* and *UL111.5A* [309, 334-336]

A protective post-natal CMV immune response is biased towards a pro-inflammatory Th1 phenotype [259, 309]. In neonates there is a natural immune bias towards a Th2 phenotype with decreased IFNγ expression and increased IL-8 expression [309]. This is detrimental in preventing congenital CMV disease [309]. IL-8 in particular directly stimulates CMV replication [309].

Background expression of IFNγ and TNFα by CD4+ and CD8+ T cells is higher in CMV seropositive individuals than seronegative controls [259, 323]. In general CMV seropositive individuals have higher background cytokine expression levels due to frequent immune stimulation throughout CMV infection [323]. Unsurprisingly, CMV infection has been described as a state of chronic adaptive inflammation [337, 338]. Inflammation in vascular endothelial cells leads to increased blood pressure and arterial stiffness late in life for CMV seropositive individuals [337, 338]. While possibly a case of antagonistic pleiotropy, maintained inflammation may be a calibrated symbiotic benefit of CMV as a part of the natural host virome earlier in life [339]. Mice seropositive for murine CMV (MCMV) fare better then CMV seronegative controls when infected with any of a variety of bacterial and viral pathogens [259, 260] (see Section 4.4).

The adaptive immune response against CMV is particularly potent. CMV elicits persistent effector memory biased CD4+ and CD8+ T cell responses that protect against disease development but are unable to clear the virus or prevent superinfection [185]. Unlike in chronic diseases such as HIV or hepatitis C, the adaptive immune response against CMV does not become 'fatigued' [340, 341].

Large populations of functional CD4+ and CD8+ T cells are maintained. In otherwise healthy individuals, 4-5% of all peripheral blood T cell populations [312, 342] and 10% of all peripheral blood memory T cell populations [312, 323] are specific for CMV. Inflationary CMV specific CD4+ T cell populations are largely polyfunctional, producing IFNγ, TNFα, IL-2 and CCL4 [185, 323].

The initial CD8+ T cell response against CMV comes from activated naïve T cells homing to the lymph nodes [309, 343]. In established CMV infection a sustained inflationary T cell mediated response is established predominantly by CCR7- CD27- CD28- CD45RA+ T cells [309, 343]. Inflationary CD8+ T cells produce IFNγ, TNFα, and CCL4, but not IL-2 [185]. They are not terminally differentiated, expand vigorously upon stimulation, and are effector cells capable of lysing target cells with granzyme B or perforin [185, 309].

A single population view is still reflected in much of the non-human primate CMV literature, where non-expression of CCR7, CD27 and CD28 are used to identify a rapidly responding cohort of "T_{EM}" cells in both the peripheral blood and the mucosa [185, 187, 200, 244, 344-346]. However, it has since become evident that two separate effector-memory populations are involved in the CMV immune response – one in circulation, T_{EM}, and the second, T_{RM}, are resident in the peripheral tissue and mucosa. The discovery that CCR7, which the CMV specific inflationary T cell populations lack, is involved in cellular exit from peripheral tissue through lymph vessels challenged the notion that a single population of T_{EM} could recirculate through both blood and tissue [347, 348]. Subsequent experiments in mice confirm that T_{EM} do not readily egress from the blood into non-lymphoid tissue, and that the T_{RM} remain in the peripheral non-lymphoid tissue where they differentiate and do not circulate [349-352]. T_{RM} populations have been identified at the portals of viral entry including skin, lung and intestinal mucosa of humans [351, 353] as well as genital mucosa in mice [354]. These cells have some proliferative capacity but T_{RM} populations are augmented throughout CMV infection by effector T cell populations which migrate into the peripheral tissue or mucosa and differentiate into effector-memory cells [355, 356]. Like T_{EM}, T_{RM} have effector functionality and can rapidly respond rapidly to antigens [350, 355]. T_{RM} are defined by their localization and lack of systemic circulation. Multiple markers have been identified to differentiate CD8+ T_{RM} from CD8+ T_{EM}, although none encompass all T_{RM} in all tissues. T_{RM} have distinct, tissue specific transcriptional profiles; mucosal CD8+ T_{RM} can be differentiated from circulating CD8+ T_{EM} by the expression of CD69, CD103 and lower expression of CX3CR1, but this is not true of all T_{RM} in all tissues [349-351, 357].

T_{RM} cells have been identified as playing a key role in controlling several viral infections in humans and mice, including influenza and HSV type-1 and type-2 [354, 358-362]. Like circulating T_{EM}, T_{RM} populations are maintained in CMV seropositive individuals and are believed to play a key role controlling the virus within a variety of tissues and at the mucosa [356, 363]. To date the relative

contributions of T_{RM} and T_{EM} have not been well differentiated in non-human primate models of CMV. As new reagents and techniques become available to researchers to better differentiate the cell types this will change.

Additionally there is a small maintained, circulating population of CD4+ $CD8^{DIM}$ MHC-II restricted cells that respond to CMV [323, 364]. These cells have a memory phenotype, are cytotoxic effector cells, and contribute a disproportionately high proportion of the CMV specific CD4+ T cell cytokine response [364]The immune response to CMV is broad: CD4+ and CD8+ T cell responses target 70% of HCMV genes [312]. The immune response to each gene is mostly proportional to the coding space of the gene with two notable exceptions [312]. Perhaps not surprisingly, IE genes are more immunogenic than would be expected by their coding space and latency associated proteins are less immunogenic, but still targeted by CMV specific T cells [309, 312].

In addition to a cellular response, there is a substantial humoral immune response maintained by the host against CMV. Antibody is produced against the surface glycoproteins – predominantly against gB, gH and the gH/gL multimer [309, 365]. Antibody against the gH/gL complex is neutralizing and prevents the infection of specific cell types through cell-free virus [309, 365, 366]. Non-neutralizing antibodies are also produced against tegument proteins and some key non-structural proteins like IE-1 [309, 366]. ADCC plays a key role in controlling CMV disease in immune compromised individuals [367].

The CMV genome encodes numerous immune modulatory genes to protect viral infected cells from the host immune response. Two CMV encoded homologues of the anti-inflammatory cytokine IL-10, UL111 and UL111.5A, are secreted from CMV infected cells [335, 336, 368]. In addition to immune suppression these genes downregulate the expression of MHC-I and MHC-II [368].

Cell surface MHC-I expression is further supressed by a multitude of viral proteins. US2 redirects the MHC-I alpha chain for proteosomal degradation [369-373], while US11 redirects the MHC-I heavy chain for degradation [200, 334, 369, 370, 373, 374]. US6 prevents completion of the MHC-I complex by preventing the Transporter associated with antigen processing (TAP) protein from translocating peptide fragments into the endoplasmic reticulum to be loaded on the MHC complex [334, 369, 373, 375]. US3 and US10 delay trafficking of completed MHC-I complexes out of the endoplasmic reticulum towards the cell surface [274, 334, 372, 373, 376, 377]. US2 and US3 further interfere with MHC-II presentation by inducing MHC-II mislocation and degradation [334, 373, 376, 378]. In old-world primate CMV, the protein O22 (Rh178) acts before any of the other MHC-I repressing proteins to prevent the completion of translation of the MHC-I heavy chain [275, 379, 380]. These proteins are key to CMV superinfection, and isolates of CMV lacking MHC-I downregulatory capacity are unable to superinfect CMV seropositive hosts [275].

Substantial MHC-I downregulation would make target cells vulnerable to NKC mediated cytotoxicity if NKC activity were not also supressed [334, 381, 382]. In addition, apoptosis of infected cells is prevented by UL36, UL37, and UL38 [288, 334, 383, 384].

In RhCMV 68-1, which is fibroblast adapted and lacks the anti-apoptotic UL36, it has been suggested that the combined effect leads to unorthodox antigen presentation as a means to control CMV infection [200, 245]. In the model, non-canonical MHC-II antigen presentation to CD8+T cells has been observed along with an increased reliance on MHC-E mediated antigen presentation, which is unregulated to counter NKC cytotoxicity [245, 385]. However this is not observed in HCMV, even fibroblast adapted HCMV intended to mimic RhCMV 68-1, and may be a unique trait of CMV in rhesus macaques [386].

A level of inflammation is maintained throughout CMV infection. Neutrophil chemo-attraction is mediated by the viral CXCL5 chemokine like secreted proteins UL146 and UL147 [274]. The inflammatory cyclooxygenase-2 (COX-2) is required for viral replication in vitro [298]. Some strains of CMV encode their own viral COX-2 homologue [278, 298, 387]. In all CMV, COX-2 expression is induced by the viral capsid bound chemokine receptor like viral protein US28 [387]. CMV is not a passive observer, or un-invested pathogen, but is involved directly in modulating the immune system that controls it. CMV and the host immune system have evolved alongside each other in a carefully negotiated detent.

1.3 Cytomegalovirus as a Vaccine Vector

CMV has been considered as a potential vaccine vector due to several characteristics of infection. The potency of the sustained robust immune response against CMV in CMV seropositive individuals is of key interest. If a fraction of the 10% of memory T cells maintained against CMV were to be directed at antigens of a desired target pathogen it could provide reliable protection from disease or infection [312, 323]. The ability of CMV to superinfect seropositive individuals further increases the usefulness of a CMV vector which, unlike other vaccine vectors, can be utilized in all individuals whether previously seropositive or seronegative CMV is of particular interest as a vaccine vector against pathogens where a CD8+ T_{EM} or T_{RM} response has the potential to protect from disease. CD8+ T_{EM} response was correlated with protection in the attenuated virus vaccines against yellow fever and smallpox [220]. CD4+ and CD8+ T_{RM} have been correlated in mice with broad vaccine induced protection against influenza [361, 362] There is thus reason to believe that T_{EM} and T_{RM} responses induced by a CMV vaccine vector may similarly contribute to protection from other pathogens. This has provided part of the rationale for the recent interest in employing CMV vectors in novel vaccine approaches against HIV/SIV [185, 388], tuberculosis (TB) [388], Hepatitis C [388], and a variety of

cancers [389]. CMV is a potent immune modulator with a lifelong effect on its host. The reactivating nature of infection may induce the robust protection required to redefine current epidemics.

1.3.1 CMV Based Vaccines

CMV has been used as a vaccine vector for candidate vaccines in the SIV model of HIV (See 1.3.2), TB [390], Ebola [391-393], and HSV type-1 [394]. Other additional targets are potentially under development.

A MCMV based TB vaccine was designed by inserting a codon optimized TB antigen 85 complex – consisting of three key proteins in *mycobacterium tuberculosis* infection – into MCMV for testing in mice [390]. The MCMV model provides a cost effect setting for early proof of concept work, but should be evaluated carefully due to substantial differences between MCMV and HCMV, as well as well characterized differences between human and mouse immune systems. The MCMV strain used in this study lacked several MHC-I downregulating proteins to prevent immune suppression of the vaccine construct, ultimately altering the nature of the CMV immune response observed [390]. MCMV-antigen 85 infected and replicated in BALB/c mice; antigen 85 RNA was constitutively expressed [390]. Despite detection of the vaccine in the liver, spleen and salivary gland the immune response generated was small [390]. T cell responses to antigen 85 epitopes could not be detected prior to an adenovirus-antigen 85 boost 4 weeks after infection [390]. The vaccine and boost did not protect from pulmonary TB when the mice were challenge at five weeks, but pulmonary TB colony forming units (CFU) were reduced in vaccinated mice [390]. The protection came mostly from NKC, with a small memory CD4+ T cell compartment, rather than the inflationary T cell response expected with CMV vectored vaccinations, possibly as a result of the adenovirus vectored boost rather than the CMV based vaccination [390].

CMV based vaccines targeting Ebola were more promising and demonstrated potentially protective immune responses [391-393, 395]. A MCMV Ebola vaccine was generated by fusing an 11 amino acid T cell epitope from the nucleoprotein of Zaire Ebola virus to the C terminus of the MCMV *IE-2* gene. The construct induced strong IFNγ+ and TNFα+ CTL responses against the Ebola virus nucleoprotein in C57/BL6 and 129S1/SvlmJ/Cr mice [392]. The T cell immune response inflated over time and adopted a T_{EM} phenotype [392]. All vaccinated mice were protected following challenge with Ebola virus while unvaccinated controls succumbed to the virus [392, 393]. Due to the absence of detectable antibody against Ebola and the presence of only a single Ebola CTL epitope in the vaccine the CMV induced CD8+ response can be credited as protective [393]. There is controversy over the validity of the murine Ebola model, as adult mice are not naturally susceptible to Ebola infection. Despite differences from Ebola in primates, the mouse adapted Ebola virus used in

these studies is lethal after intraperitoneal inoculation [396-398]. Separated from the challenge efficacy the immune responses generated by the CMV vaccine vector were still promising.

A further CMV based Ebola vaccine was tested in a small study using the Rhesus macaque model. The vaccine construct expressed a full length, codon optimized Ebola glycoprotein under the control of the *RhUL83* (*Rh112*) gene promoter in RhCMV 68-1 [391]. Macaques were vaccinated and boosted at twelve weeks, then challenged with a lethal dose of infectious Ebola virus [391]. Three of four vaccinated animals, and neither of two unvaccinated controls, survived the challenge [391]. Protection was not sterilizing, two of the three protected animals developed fever but fully recovered [391], a result that is difficult to interpret into solid conclusions due to the small size of the study.

Protection against Ebola in the study was not correlated with T_{EM} responses. In fact the vaccinated animals lacked detectable CD4+ or CD8+ T cell responses to the Ebola glycoprotein before challenge [391]. Ebola glycoprotein targeting non-neutralizing IgG was abundant in vaccinated animals, but the study was too small to establish any correlation or lack there-of between IgG responses and protection [391]. However, the apparent protection does demonstrate that the construct warrants further attention. The interesting lack of T cell responses against Ebola glycoprotein can potentially be attributed to timing of protein expression. Expression using the RhUL83 promoter, contrary to the constitutive protein expression used in other CMV based vaccines, expresses the Ebola glycoprotein concurrent with peak expression of MHC downregulating viral proteins, which are under similar promoters [391]. MHC presentation of Ebola glycoprotein may therefore have been supressed and a robust memory T cell response unable to develop [391]. Why IgG was produced against Ebola glycoprotein, when it is not observed in other CMV based vaccines (See Section 1.3.2), is worth investigating in further studies.

Expression characteristics of the vaccine transgene also proved important in shaping the immune phenotype in a MCMV based HSV-1 vaccine [394]. Several putative MCMV based HSV-1 vaccines were designed by fusing a single immunodominant MHC-I restricted epitope of the HSV-1 *gB* to the C terminus of the *IE-2* gene or *M45* (partially homologous to HCMV *UL45*) gene [394]. The *IE-2* gene is expressed most abundantly immediately after infection and at low levels during latency [394]. *M45* is expressed as an early gene, similar to *RhUL83*, concurrent with most MHC downregulatory proteins [391, 394].

Both vaccines protected mice from challenge with HSV-1 [394]. Expression of the antigen with early kinetics under the M45 promoter induced a CD8+ T cell response within 12 hours of inoculation that waned after 14 days [394]. The inflammatory response from mice inoculated with the IE expressed antigen under the *IE-2* promoter was initially milder. Early expressed M45 fused antigen had a higher percentage of CD8+ T cells targeting the antigen (15% against 5% for the IE-2 based

vaccine), and a higher percentage of antigen specific CD8+ T cells producing IFNγ (30.5% versus 19.9%). However the CD8+ T cell response to the IE expressed antigen IE-2 continued to expand over 60 to 120 days and was maintained until follow-up ceased at 180 days post inoculation [394]. At 180 days ~5% of the CD8+ T cells in mice given the *M45* fused vaccine targeted HSV-1 epitopes, compared to ~35% of CD8+ T cells in mice given the *IE-2* fused vaccine [394]. A larger portion of the CD8+ T cells in mice given the *IE-2* fused vaccine retained effector functionality [394]. This alongside the findings of Marzi et al. with their RhCMV 68-1 vectored Ebola Vaccine, supports the idea that timing of antigen expression in a CMV vector vaccine is key to develop the desired sustained CD8+ inflationary T cell phenotype. Antigen expression should be IE, and continuous at low levels throughout viral latency to develop robust inflationary T cell responses [391, 394].

1.3.2 CMV Based HIV Vaccines

Hansen et al. developed a RhCMV based SIV vaccine candidate to model a HCMV based vaccine against HIV in the rhesus macaque model [185]. RhCMV 68-1 was prepared to express SIVmac239 derived transgenes [185]. The RhCMV 68-1 vectored SIV vaccine failed to protect any animals from SIV acquisition but led to control of the virus down to below the level of detection in roughly half of all vaccinated animals through multiple trials (see 1.1.3) [185, 187, 200, 244]. The remaining vaccinated animals had peak and set-point viral loads identical to negative control animals and rapidly progressive disease [185, 187, 200, 244]. In the studies RhCMV 68-1 was modified to express one of five SIV protein constructs: a Rev, Tat, and Nef fusion protein, a Gag protein, an Env protein or one of two inactivated Pol proteins [185, 187, 200, 244]. All constructs except Env were designed with an *EF1α* promoter [185, 187, 200, 244]. The Env protein, modified to lack the RRE, was designed under the control of the CMV *gH* promoter [185, 187, 200, 244].

RhCMV 68-1 vector-induced SIV specific T cells exhibited the functional profile of natural CMV specific T cells [185]. CD4+ T cells were polyfunctional; a large SIV specific CD8+ T_{EM} pool was established [185] (See 1.2.2.2). Following vaccination with the RhCMV 68-1 vectored SIV vaccines, 1.5% of memory CD4+ T cells and 2.0% of memory CD8+ T cells in the blood of vaccinated animals were specific for SIV antigens [185]. The vaccine generated CD8+ T cell responses were broad: responses to the SIV Gag protein targeted three times as many epitopes as occur with a conventional SIV-induced Gag response [200, 245]. Substantial CD8+ T cell responses were observed against the Env transgene, despite its expression under the *gH* promoter, rather than the constitutive *EF1α* promoter used to promote the other transgenes [185, 244]. It was not presented, however, whether responses against Env exhibited a T_{EM} phenotype [185, 200, 244]. A lack of CD8+ T_{EM} responses were observed in similarly non-IE expressed transgenes in other CMV vectored vaccines (See Section

1.3.1). It is thought that CD8+ T cell responses against Env are of limited value in a vaccine as HIV/SIV Env is highly variable and produced only late in infected cells immediately prior to virus production [31], so the contribution of Env to protection, whether T_{EM} or not, was likely minimal. Alongside cellular responses, the vaccines induced only weak, non-neutralizing antibody to SIV targets and none against Env [185, 200].

Interestingly the CD8+ T cell response was not limited to MHC-I epitopes [200, 245]. Instead T_{EM} targeted MHC-I, MHC-II and MHC-E restricted epitopes [200, 245] (See Section 1.3.2.1). There were no CD8+ T cell responses to the typical immunodominant epitopes on Gag and Tat, amino acids 181-189 and 28-35 respectively [200]. These epitopes had been present in the vaccines, and the group of macaques primed with AD5 vaccination prior to the RhCMV 68-1 vaccines developed responses to these epitopes that were not boosted by the subsequent RhCMV 68-1 vaccine [200, 244]. The failure to target typical immune dominant epitopes was identified to be due to action of the MHC-I downregulatory CMV protein US11 [200] (See Section 1.2.2.2). Deletion of US11 restored immune targeting of conventional vaccine induced SIV epitopes while maintaining the breadth of alternative targets but impaired the ability for RhCMV to superinfect CMV seropositive macaques [200].

The exact correlates of protection from vaccination in these studies with the RhCMV vectored vaccine are unknown. Protection was mediated by CCR7- CD28- inflationary T cell responses, but these responses were observed in both protected and unprotected vaccinated animals [185, 187, 200, 244]. Similarly, CD8+ T cells targeting MHC-I, MHC-II and MHC-E restricted epitopes were seen in all vaccinated animals, and there were no quantitative differences found in the magnitude or qualities of the immune responses between the SIV controlling and non-controlling animals after exhaustive examination [200, 245]. Furthermore, no known genetic phenotype was involved in protection; there was no significant difference in MHC type and TRIM5α alleles between groups [244].

The T_{EM} response, though non-sterilizing, was able to effectively control and eliminate SIV infection in the half of vaccinated animals that were protected [185, 187, 200, 244]. Protected macaques had lower peak and set-point viral loads, and viremia was quickly reduced to below detectable levels after infection [185, 187, 200, 244]. Vaccine induced T_{EM} responses were long lived and outlasted responses to SIV antigens not included in the vaccines, such as Vif [187]. Though viremia periodically elevated over the first 50-70 weeks, it was subsequently supressed by the immune response [185, 187]. It appears that after this time the macaques had cleared SIV, as transient depletion of the CD8+ T cells using cM-T807 did not result in the return of viremia [185, 187]. In contrast, unprotected animals had a peak and set-point viral load not significantly different from unvaccinated animals and never controlled the virus down to undetectable levels [185, 187, 200, 244].

It may be that as T_{EM} are outside the lymph compartment and have limited expansion potential to match an established viral infection [185].

A lack of antibody response against Env despite its inclusion as a vaccine vector has been observed in multiple herpesviruses based vaccines [185, 244, 399]. The RhCMV 68-1 vectored SIV vaccines reported no notable antibody responses [185, 187]. Regardless, if a method cannot be devised to generate antibody responses through a CMV vector it will still be possible to later pair the vaccine with an Env targeting antibody mediated vaccine as was the case with ALVAC-HIV in RV144 [246].

The protection of half of all vaccinated macaques with CMV-SIV vaccines warrants further investigation of CMV as an SIV vaccine vector. Until the correlates of protection are understood, however, it may be difficult to improve on the results seen thus far. Unfortunately peculiarities of RhCMV 68-1 vector will complicate the search for correlates.

1.3.2.1 Rhesus Macaque CMV 68-1

RhCMV 68-1 has been widely utilized since it was isolated from the urine of a captive Indian rhesus macaque in 1968 [400]. It was not until 2003 that a stable BAC incorporating the full length RhCMV 68-1 genome was developed by Chang et al. [288, 401]. Prior to 2003, the virus underwent many years of adaptation from passaging and selection resulting in known and unknown effects [282, 288, 385, 401]. For instance, the lab attenuation prior to cloning of both the HCMV AD169 strain and the RhCMV 68-1 strain as BACs has resulted in slower growth kinetics than wildtype CMV and the production of lower viral titres in a single cycle growth curve assay [401].

RhCMV 68-1 is considered fibroblast adapted. The virus is attenuated in the *UL128-UL131* region necessary for entry into endothelial and epithelial cell lines as part of the glycoprotein pentameric complex [200, 303, 305, 309, 379, 385, 402]. However the virus is still capable of infecting a rhesus macaque microvascular endothelial cell line *in vitro* [403]. The genome contains multiple smaller attenuations, and is missing homologues of HCMV *RL13*, *UL36*, *UL119*, *UL118*, and *US14D* present in other wildtype RhCMV strains [379]. Following repair of *UL36* RhCMV 68-1 is also able to infect kidney epithelial cells *in vivo* despite the lack of a pentameric complex [379].

Alongside deletions, RhCMV 68-1 encodes for a number of proteins not present in HCMV, many of unknown function. Of note is Rh10, a homologue of mammalian Prostaglandin-endoperoxide synthase 2 (PTGS2), or COX-2 [288, 298, 387]. *COX-2* expression is induced upon entry into host cells in most CMV strains by the viral protein US28 [387]. A number of CMV strains, including all sequenced RhCMV strains, also encode a *PTGS2* homologue to ensure *COX-2* expression [379].

Despite attenuation RhCMV 68-1 is capable of replication *in vivo*. Reactivating infection with a RhCMV 68-1 vectored SIV vaccine was established in rhesus macaques. Following immunization the macaques developed robust cellular immune responses against the included SIV transgenes, as measured by intracellular cytokine staining [200]. As noted above, vaccinated macaques developed responses to three times more epitopes on SIV antigens than develop in natural SIV infection [200] (See Section 1.3.2). Initially the MHC class restriction and effector cell type was not reported, only that overlapping 15mer libraries of SIV peptides were used to measure TNFα and IFNγ production. It was some time later determined that CD8+ T cell effectors were not targeting conventional MHC-I epitopes [200, 385]. Instead the CD8+ T cells were targeting both MHC-II and MHC-E epitopes [200, 245, 385]. Moreover these responses occurred in both animals that controlled SIV infection and non-controllers, and so could not be considered correlates of protection even when additional studies were completed [200, 245, 385]. This unusual antigen targeting is not associated with the typical anti-CMV immune response; rather it is associated with the use of the particular RhCMV 68-1 strain as a vector [200]. Restoration of the *UL128-UL131* homologous region, as in the RhCMV 68-1.2 construct, results in exclusively MHC-I restricted epitope responses, and the re-excision of *UL128-UL131* restores MHC-E epitope targeting [200, 245]. However, fibroblast adapted HCMV strains lacking a functional *UL128-UL131* region do not generate MHC-II or MHC-E restricted CD8+ T cell responses [386]. It is possible MHC-E targeting is due to some other feature of RhCMV 68-1, is the result of a combination of deleted and altered genes, or that there is some effect from the decreased diversity of MHC-E alleles in humans compared to Rhesus Macaques (2 compared to 25) [404] (See Section 1.4.2.2).

Other reactivating herpesvirus vectored vaccines have achieved levels of protection without notable MHC-II or MHC-E restricted epitope targeting [218, 390, 391, 393]. Instead it is argued that protection is mediated by peripheral T_{EM}/T_{RM} responses, with the exception of one study where protection against Ebola was thought to be mediated by antibody responses [185, 187, 390, 391, 393] (See Section 1.3.2).

The group of Louis Picker and associates who initiated the RhCMV 68-1 vector studies have recently focused on the MHC-II and MHC-E restricted CD8+ T cell immune responses as a potential source of protection [200, 245, 404]. However, to date there have been no efforts to compare the RhCMV 68-1 vector against a less attenuated RhCMV strain, which might induce typical CD8+ T cell effector responses.

The RhCMV 68-1 vectored SIV vaccine model has shown the potential for CMV as a vaccine vector. It has not, however, provided reliable protection in all vaccine recipients. Instead, through a mechanism still not fully understood, only half of vaccinated animals are protected from SIV disease.

Additionally, there is no known way of replicating the odd MHC-II and MHC-E restricted pattern of antigen presentation in a HCMV vaccine [386]. As was observed with AD5 vectored SIV vaccines, protection observed in rhesus macaques does not always correspond to protection in humans [32, 46, 253]. Without knowing the correlates of protection, translating even non-human primate studies from animal surrogate virus models to vaccine efficacy in humans is not predictable.

To better understand the immune mechanisms at play in a complex CMV vectored vaccine, we have proposed the use of the cynomolgus macaque-CMV model. Cynomolgus macaques are a distinct but closely related species to rhesus macaques [405, 406]. They have been routinely used with SIV as a model of HIV in humans. Furthermore they have their own more recently identified and well documented strains of CyCMV with minimal lab adaptation when compared to RhCMV [289, 407]. As a readily available scientific model they represent an excellent opportunity to better understand the role of CMV and immune correlates of protection against SIV/HIV infection with a non-human primate CMV vectored vaccine.

1.4 Macaque Models

1.4.1 Differentiation of Macaque Species

The *macaca* genus, incorporating all macaque monkeys, originated in northeast Africa 7 mya during the late Miocene epoch [408]. The genus spread across Europe and to continental Asia roughly 5.5 mya, diversifying into 22 extant species as they migrated [408, 409]. Today macaques are spread across a range of more than five million square kilometers between the Iberian peninsula, and Japan, encompassing much of southern Europe, Asia and northern Africa [409]. Rhesus macaques (*Macaca mulatta*), cynomolgus macaques (*M. fascicularis*), Formosan rock macaques (*M cyclopsis*) and Japanese macaques (*M. fuscata*) are included in the fascicularis group that evolved from a common ancestor in South-east Asia 2.5 Mya [408-410].

Multiple scenarios have been proposed to explain macaque divergence based on occasionally conflicting evidence from fossil sources [411], macaque morphology [406, 408], protein polymorphisms [412] and MHC haplotyping [413, 414]. Best estimates place the most recent common ancestor of cynomolgus and other fascicularis group macaques in Thailand [415, 416]. From there other macaque species migrated northward into mainland China, Formosa and Japan [408, 417]. Cynomolgus macaques spread throughout peninsular and insular Indochina while rhesus macaques spread between southern China and northern India [416, 418]. The two species began to diverge approximately 1.2 mya [419, 420]. There is still some ongoing introgression between cynomolgus and rhesus macaques in peninsular Indochina where their ranges overlap, though a hard genetic

barrier exists at the Kra isthmus in southern Thailand and Myanmar [416, 421]. Comparisons of mitochondrial DNA and Y chromosomal DNA show a tendency towards male migration and female philopatry both in rhesus and cynomolgus macaques [409, 419, 420, 422]. Y chromosomal DNA (YDNA) shows male cynomolgus macaque migration and genetic exchange continued throughout Indonesia and the Philippines until 400 thousand years ago (kya), likely ending when the islands separated due to Pleistocene glacial events [409, 420, 423].

1.4.2 Macaque Use in Biomedical Research

Cynomolgus and rhesus macaques are key models of biomedical research. They were, respectively, the most and second most procured non-human primates at research facilities in North America between 2011 and 2014 [424]. Macaques and humans shared a common ancestor 25 – 28 mya [425-427]. While not as close of a homologue to humans as chimpanzees, macaques are more readily available, and less costly to work with [428].

There are notable immune differences from the human system that must be accounted for in the macaque model [426]. *In silico* studies have identified 164 human-specific immune related genes, including TLR4 and DEFA4, which have become non-functional pseudogenes in at least one if the macaque species [426]. A further 25 key immune related genes, including a homologue of IL32, are single copy genes susceptible to loss of function mutations [426]. Possibly as a result of this and other factors, results of protection seen in macaque vaccine studies have not always replicated in subsequent human studies (See section 1.1.3.1). It is therefore important to go beyond efficacy observed for an intervention and characterize the underlying mechanisms in any animal model trial.

1.4.2.1 Rhesus Macaques

In the early 2000's rhesus macaques were the most frequently used non-human primate research model by publication number, outcompeting cynomolgus macaques by three to one [424]. They are still a preferred species in much of the USA as disease susceptibility and MHC are well understood [428]. Due to a global moratorium affecting the export of Indian Rhesus macaques in the 1970, the majority of rhesus macaques in North America are captive bred [424]. As a limited commodity they have become expensive to purchase.

Chinese origin rhesus macaques are more readily available to the scientific community, but cannot be used as a direct replacement of Indian rhesus macaques in research [428, 429]. Chinese rhesus macaques are genetically distinct, and respond uniquely to a variety of pathogens, including SIV [426, 428-430]. Diversity between the Indian rhesus macaque and Chinese rhesus macaque subspecies rivals that between rhesus and cynomolgus macaques, and are sufficient to confound

research outcomes [426]. There is a great deal of heterogeneity even within Chinese rhesus macaque populations and the MHC genetics of the subspecies is only now being explored [425]. As of yet it is unknown to what extent RhCMV strains of Chinese rhesus macaques differ from those of other subspecies.

1.4.2.2 Cynomolgus Macaques

Cynomolgus macaques are increasingly the favoured non-human primate research model. They are now the most commonly procured non-human primate in North America, and well used in Europe [424, 428]. There are several populations of cynomolgus macaque available spread across peninsular Indochina and the Indofilipino archipelago. Genetic diversity between macaques across the islands is as extreme as that between Chinese and Indian Rhesus macaques, with some scientists calling for the reclassification of cynomolgus macaques as several distinct species [430].

Sometime between 1598 and 1606 a small founder population of macaques were introduced on the island of Mauritius in the Indian Ocean, possibly on Portuguese trading vessels or pirate ships [431-435]. Due to peninsular YDNA and insular mitochondrial DNA in cynomolgus macaques of Mauritius, also found in Sumatran cynomolgus macaques, Sumatra is assumed as the origin of the Mauritian founder population [436]. This extreme bottleneck on Mauritius has created an abundant and relatively homogenous population of macaques ideal for use in research [430-433, 435]. In particular Mauritian macaques have a lower diversity of TRIM5 genes and MHC haplotypes that may confound vaccine studies [426, 428, 437]. A recent focus on MHC-E restricted CD8+ T cells in the Rhesus Macaque RhCMV 68-1 model has highlighted the importance of Mauritian macaques (See Section 1.3.2.1). There are only two known human MHC-E alleles that narrowly limit the peptide binding diversity [438-440]. In contrast to the 25 MHC-E alleles found in Indian Rhesus macaque populations, up to four of which may be expressed in any one animal, a total of only four MHC-E alleles were found among Mauritian cynomolgus macaques [404]. If the Mauritian macaque population is restricted to most common MHC haplotypes M1, M2 and M3, then similar to humans only two MHC-E haplotypes, with the same peptide binding affinity as human MHC-E, are expressed [404].

Both captive bred and wild caught Mauritian cynomolgus macaques are used in research. Currently there are 30 000 - 40 000 macaques of low heterogeneity on Mauritius [435]. The island is the second largest exporter of live cynomolgus macaques globally, behind only China [441].

Cynomolgus macaques differ from Indian rhesus macaques by 0.4% across the genome including in key immune related genes [426]. TRIM5α, a key restriction factor against retroviruses including SIV, differs between the macaque species at a number of polymorphism sites [426]. In Cynomolgus

macaques, 97.5% have a 6bp deletion resulting in the loss of Thr339 and Phe340, only 50% of Chinese rhesus macaques and 36% of Indian rhesus macaques has the same deletion [426]. Loss of these amino acids is thought to increase susceptibility to retroviral infection [426, 442, 443].

Previous research with cynomolgus macaques was held back by a lower availability of reagents, and genetics that were previously less well characterized than those of rhesus macaques [428]. With increased usage, and the increased development of reagents and resources for use in this model, cynomolgus macaques have come into their own. The recent publication of the complete cynomolgus macaque genome sequence [426] demonstrates the increasing importance of these animals in scientific investigation.

1.4.3 The SIV Vaccine Model of HIV

There is no ideal non-human model of HIV. HIV-1 cannot reproducibly cause human-like disease in any animal model [437, 444-449]. Chimpanzees can be infected with HIV-1, potentially resulting in limited pathogenic infection, but not consistently mimicking disease in humans [437, 446-448, 450]. Furthermore, ethical issues regarding the use of chimpanzees in research firmly prevent the development of this model. Pigtailed macaques (*Macaca nemestrina*), a more distant relative of cynomolgus and rhesus macaques which diverged from the fascicularis group 3.5-5 mya [409], are also susceptible to viral infection and replication by HIV-1, though it does not cause AIDS and the virus is cleared within several weeks [449, 451]. Pigtailed macaque susceptibility to HIV-1 results from the absence of TRIM5α in the species [443, 452]. Instead, Pigtailed macaques encode a TRIMCyp also present in cynomolgus macaques that does not inhibit HIV-1 infection [443, 452]. Replacing the *vif* gene in HIV-1 with *vif* from SIVmac to counteract macaque APOBEC3, another HIV restriction factor (See Section 1.1.1.1), generated a simian Tropic HIV-1 that could establish infection in pigtailed macaques for up to 25 weeks before it was cleared [451]. The pig-tailed macaque HIV-1 model provides an opportunity to test sterilizing immunity in a vaccine, but due to the absence of human-like disease progression and eventual clearance of the virus it does not provide an opportunity to examine alternative forms of vaccine mediated protection (See Section 1.1.1.3).

Instead, to accurately model HIV-1 infection and progression in humans, non-human primate models of SIV have been developed. Unique strains of SIV are endemic to more than 40 species of primates, but result in limited or non-pathogenic infections in most African non-human primate species [437, 450, 453-455]. SIV infected African sooty mangabeys show no signs of disease despite chronic high viral loads [456]. As the viruses progress to AIDS in some Asian non-human primate species, SIV was first discovered and isolated in rhesus and cynomolgus macaques despite the absence of the virus in wild populations [453, 457]. SIV infection was unintentionally established in

captive macaque populations multiple times following the passaging of sooty mangabey tissue into macaques in US primate centers as a part of several unrelated studies [458, 459]. Two sooty mangabey derived SIV lineages, SIVmac and SIVb670/H5, account for the majority of macaque model-SIV research [437, 458, 459]. SIVmne, a third family developed through passaging in pigtail rather than rhesus macaques is considerably less pathogenic [459, 460].

The pathogenic SIVmac251 viral swarm, and highly-pathogenic, neutralization-resistant SIVmac239 viral clone, are derived from the SIVmac lineage after 4 and 7 serial passages in rhesus macaques respectively [437, 458, 459]. They utilize CCR5 as a co-receptor, but replicate better in CCR5+ T cells than macrophages [437, 459, 461]. The SIVb670/H5 lineage includes the commonly used viral swarm SIVsmE660 and related viral clone SIVsmE543-3 [437]. The genetic distance between SIVmac251 and SIVsmE660 is similar to that between HIV-1 isolates of the same clade [437]; the percent similarity between individual proteins of SIVmac239 and SIVsmE543-3 range from 64.1% for Tat to 92.1% for Pol [461]. Unlike the SIVmac viral family, SIVsmE543-3 is a dual tropic virus, and induces early neurologic and pulmonary pathology in infected animals [461].

All four viruses induce human-like disease progression to simian AIDS in infected macaques [437, 459, 461, 462]. The macaque-SIV model of HIV therefore allows examination of both sterilizing and non-sterilizing vaccine approaches (See Section 1.1.1.3). SIVmac251 and SIVmac239 establish similar peak and set-point viral loads in Indian rhesus and pigtail macaques, while infection of cynomolgus macaques results in a lower average peak and set-point viral load [437, 463]. CD4+ cell depletion rates are also similar in rhesus and pigtail macaques, but pigtail macaques progress far more rapidly than rhesus and cynomolgus macaques to AIDS, averaging 42 weeks instead of 70 for rhesus macaques [463]. With SIVsmE543-3 in pigtailed macaques some animals have been observed to progress to AIDS within 14 weeks [461]. This accelerated disease progression is ultimately due to higher natural rates of T cell turnover in pigtail macaques [463].

A third option of modelling HIV-1 infection in non-human primates is through the use of SHIVs. The first generation of SHIV replaced the *env*, *tat*, and *rev* genes of SIVmac239 with Clade B HIV-1 equivalents [464]. The virus could infect and reproduce in macaques, but was non-pathogenic [464, 465]. Subsequent SHIVs included *vpu*, a gene which down-regulates surface CD4 expression and enhances viral budding [69, 72, 465] (See Section 1.1.1.1). The SIV Env, unlike HIV-1 Env, generally contributes to these functions, but Vpu still enhances SIV replication [72]. *Vpu* improved the persistence of SHIV in cynomolgus macaques, but did not immediately augment pathogenicity [465].

SHIV 86.9 constructed expressing dual tropic HIV-1 *env* and *tat*, *vpu* and *rep* replicated at higher levels in rhesus macaques then prior SHIVs constructed using CCR5 tropic Env [466], but it was

repeated passaging in rhesus and pigtailed macaques that produced pathogenic SHIV [467, 468]. Initially $SHIV_{DH12}$/$SHIV_{DH12R}$ required transient CD8+ T cell depletion to establish viral infection, however subsequent passages readily infected and induced AIDS in rhesus macaques [467]. Accumulated mutations in the HIV-1 *env* gene adapted the virus for macaque hosts, and led to CXC4 tropism [467, 469]. This tropism has been problematic, as CXCR4 tropic pathogenic SHIVS targeted naïve CXCR4+ CCR5- T cells, leading to severe immune depletion within weeks of infection [469]. Thinking SHIV 86.9P a robust challenge virus because of this, highly pathogenic macaque passaged SHIV86.9P was utilized in a variety of studies including an AD5 vectored HIV vaccine challenged that led to the failed HVTN 502 clinical trial [247] (See Section 1.1.3.1). AD5 vector HIV vaccinated macaques were protected against SHIV86.9P challenge, but not challenge with SIVmac239 which, it has since been discovered, is far less sensitive to neutralization [46, 247, 251, 252]. As the human trials ultimately failed, the validity of SHIV86.9P as a model of human HIV-1 in most vaccine trials should be carefully evaluated.

There remains continued development of SHIVs. Due to their inclusion of human genes they are susceptible to HIV-1 ART and unlike SIV can test HIV-1 targeting antibodies [437, 459, 464]. The macaque-SHIV model was designed to overcome differences in immune responses to HIV and SIV Env protein that must be accounted for in the macaque-SIV model [464]. Though HIV-1 *env* included in SHIV typically acquire mutations to adapt to their macaque hosts, they remain closer to HIV-1 *env* than SIV *env*. CCR5 tropic SHIV have been developed, better mimicking early transmitted HIV-1 variants, along with SHIV encoding non-clade B HIV-1 *env* [437, 459, 470-473]. They remain an important model of HIV-1 in macaques, but in a vaccine approach focusing on effector CD8+ T cell mediated protection, and examining effect on disease progression, SIVmac239 and SIVmac251 currently present the most robust challenge. In addition, studies utilising SIVmac239 and SIVmac251 can be directly compared with a wide range of existing vaccine trials including the RhCMV 68-1 vectored SIV vaccine in rhesus macaques and human VZV_{Oka} vectored SIV vaccine in the cynomolgus macaque model [187, 200, 244, 251, 252, 288, 345, 474-476]

1.4.4 Macaque Model of CMV

Multiple animal models are utilized for the study of CMV. The most widely used small animal models, mouse and rat, are responsible for the vast majority of CMV based research to date. The murine model in particular is cost effective and quite powerful due to the breadth of well characterized MCMV strains, transgenic mice, and molecular reagents available for use in the model. However, both models are limited in scope due to anatomical differences between humans, mice, and rats, viral differences between HCMV, MCMV and Rat CMV (RCMV) and resulting differences in

their course of infection [268, 293]. The guinea pig model, infected with guinea pig CMV (GPCMV) is well utilized as a small animal model of congenital CMV due to similarities between human and guinea pig placentae not present in mice or rats, and the transmission of CMV across it [268, 293], but remains similarly restricted.

The macaque model of CMV was developed to better understand congenital CMV and CMV related sensineuronal damage [292]. Similar to in small animal models of CMV, the strict host species specificity precludes the use of HCMV in any non-human primate models [268, 292, 293, 477]. Macaque CMV is the most similar animal model of HCMV available – only chimpanzee CMV (CCMV) and CMV of the other great apes, which are unfeasible as laboratory models, are more closely related to HCMV [282, 292, 294]. Similarities also exist in the host pathogen interaction of humans and macaques to their respective CMVs. Macaque CMV grows in the same tissue, establishes latency in the same reservoirs, and produces the same congenital defects, and transmits at similar rates to those observed in humans [268, 477, 478]. Immune responses to CMV in macaques mirror those seen in humans [268, 478]. Macaque rates of seroprevalence to CMV are high: essentially all captive bred rhesus macaques are seropositive at 1 year of age, half at six months [477]. Until the recent implementation of specific pathogen free breeding programs for Indian rhesus macaques, CMV seronegative macaques – including cynomolgus macaques and Chinese rhesus macaques – were not available to researchers [268, 477, 479]. Several RhCMV seronegative breeding groups have been established by removing newborns from dams at birth and raising them separated from infected adults [479]. At present there is a limited supply of RhCMV seronegative macaques, and colonies are maintained at high cost.

Due to the availability of macaques as a model of both SIV and CMV, they are an ideal model to examine a macaque CMV vectored SIV vaccine. Two RhCMV strains have been cultured and characterized, RhCMV 68-1 (See Section 1.3.2.1) and RhCMV 180.92 [282, 480]. Both were isolated from Indian rhesus macaques [282, 480]. CyCMV Ottawa, the only strain of CyCMV isolated and sequenced prior to the work in this thesis, was collected from a 4-year-old Filipino cynomolgus macaque by members of the authors research laboratory [289, 407].

1.4.4.1 Manipulating Macaque CMV

CMV does not accommodate most conventional DNA manipulations [276, 401, 481]. The virus is too large and grows slowly. Instead chemical mutagenesis and lab adaptation have been relied on historically to study the knock-out of particular genes [276]. Lab adaptation through repeated passaging can lead to the loss of genes necessary *in vivo* but not *in vitro*, like immune modulatory

genes [282, 480, 482, 483]. Passaging exclusively on fibroblast cell lines can similarly lead to fibroblast adaptation and attenuation of genes unnecessary in that cell type. This happened with RhCMV 68-1 which lost genes in the pentameric complex [379, 484] (See Section 1.3.2.1). For site directed mutagenesis, or more complex manipulations, homologous recombination has historically been the only viable option with CMVs [276, 401, 481, 485]. Unfortunately, slow growth kinetics and the high rate of genetic adaptation to tissue culture complicate this strategy [485].

Recently CRISPR (Clustered regularly interspaced short palindromic repeats) and CRISPR associated protein 9 (cas-9) systems have been used to edit GPCMV [485]. Specific guide RNA can be prepared to target any unique region of the CMV genome to change, remove, or insert specific DNA sequences [485]. CRISPR/Cas-9 edited CMV improves on homologous recombination in terms of efficiency, but is still limited by viral growth speeds and the risk of recombination events between edited and non-edited viral genomes. Rates of viral attenuation and genomic instability also limit the number of sequential edits possible if multiple changes are desired at different sites in the genome. Additionally, changes which remove essential viral genes to create replication deficient viruses are difficult within a system that edits CMV as a virus, and cannot be done if the gene cannot be expressed by the infected cell line [276].

To preserve the CMV genome in a stable format throughout multiple changes it necessary to clone CMV onto cosmids or as a BAC. To clone CMV onto cosmids the genome is broken up into multiple overlapping fragments which are independently cloned onto smaller plasmids [276]. Each individual plasmid can be replicated in bacterial cells, stored with high fidelity, and manipulated by traditional cloning methodology [276]. Alternatively, as the cost of chemical DNA synthesis falls, entire cosmid sequences can be synthesized exactly as desired. The virus cosmids can be reconstituted as an infectious virus through co-transfection in host cells [276]. Due to the need for replication to produce competent virus cosmid manipulations are limited. Essential genes cannot be knocked out unless they can be expressed in the infected cell line [276]. Additionally unexpected recombination can occur between high copy cosmids within a bacterial cell, or between cosmids in a eukaryotic cell to duplicate regions of DNA, lose regions, or acquire mutations [276, 486]. While revolutionary to the CMV field, cosmids have now largely been replaced by BACs.

BACs are based off the *E. coli* F plasmid, which can hold up to 300kbp of DNA [276, 401, 487]. The BAC backbone is a single copy plasmid [276, 487]. This restricts the possibility of recombination occurring between viral genomes within a bacterial cell [276]. As a result BAC contained viral genomes are exceptionally stable and have been replicated in *E. coli* without changes for over 100 generations [276, 487]. Similar to cosmids, since the DNA is faithfully replicated in bacteria, large volumes of DNA can be easily produced and the viral DNA can be stored in bacterial

glycerol stocks. There is no limit to the number of subsequent manipulations that can be done, and any number of changes can be made to the CMV viral genome once it has been cloned onto a BAC [276].

Edits to BAC DNA are made primarily through homologous recombination with small, editable shuttle vectors – plasmids with regions of homology to the BAC flanking a central region with the desired insertion, deletion, or mutation [276, 481, 488].

A CMV BAC can be reconstituted as infectious virus through transfection into a suitable host cell. Viral genes are transcribed and translated, and viral DNA replicated once the genome is in suitable cell culture. Removal of the BAC DNA from the CMV genome is not necessary to produce infectious virus but is advised. The additional, unnecessary DNA slows viral growth kinetics and decreases viral yield if retained in the viral genome [401, 481]. If retained, BAC DNA in the virus has been linked to instability in the surrounding viral genome [486, 488]. Slower growth kinetics and lower yields can be overcome by removing unnecessary CMV genes where the BAC is inserted to not change the overall CMV genome size [489]. Alternatively the BAC can be removed from the virus after it has been reconstituted to restore viral growth kinetics [481]. BAC DNA can be inserted surrounded by *loxP* so that co expression of Cre in infected host cells will excise the BAC from the virus [481].

There are several limitations to the use of BACs with CMV. The CMV BAC cannot incorporate any genes that will be expressed in bacteria and produce negative effects [485]. The CMV promoters are eukaryotic promoters for expression in mammalian cells, but can contain putative bacterial promoters within them [490]. At least two studies have shown some level of gene expression of transgenes under the CMV promoter in *E. coli* provided a bacterial ribosome-binding site occurs somewhere downstream within the gene [490, 491]. This is not a concern for the majority of the CMV genome but must be accounted for if transgenes are inserted into the BAC. Similarly, BACs containing features such as *loxP* cannot be stored in Cre expressing *E. coli*.

A major limiting factor to the use of BACs is the time and effort required to generate one. Multiple sequential steps of homologous recombination are required to integrate BAC DNA into the CMV genome [401, 481, 488, 492]. After each edit several laborious rounds of plaque purification are required to select for successfully edited virus [401, 481, 488, 492]. It can take months to years of effort to successfully clone a CMV as a BAC. Each round of plaque purification introduces a genetic bottleneck where mutations can also be selected.

Multiple BACs have been prepared from a variety of CMV viruses, including HCMV, MCMV, and GPCMV [481, 488, 489]. At the initiation of this project there existed only one macaque CMV that had been cloned as a BAC, RhCMV 68-1 [401]. No CyCMV BAC had been created.

1.4.4.2 RhCMV 68-1 Does Not Infect Cynomolgus Macaques

Only recently has it been accepted that each species of macaque constitutes a discrete model for biomedical research (See section 1.4.2). Significant differences exist between Indian rhesus macaques, Chinese rhesus macaques and cynomolgus macaques in their interaction to both CMV and SIV [379, 426, 428-430, 493, 494].

RhCMV 68-1 was first used in the rhesus macaque model to vaccinate against SIV with promising but mixed results; at this time, the correlates of the protection seen in a subset of animals have yet to be determined [187, 200, 244, 288] (See Section 1.1.3). However, the number of precise modifications to the CMV genome needed to prepare it as a viral vaccine vector for similar studies calls for the use of a suitable CMV cosmid or BAC (See Section 1.4.4.1).Cynomolgus macaques had been used for SIV research in Canada for more than twenty years due to the availability of animals from a breeding colony first developed for Polio vaccine quality control and research in 1983 [495]. The colony was maintained by Health Canada in Ottawa until the early 2000s, and cynomolgus macaques were readily available to Canadian researchers through the Public Health Agency of Canada. As RhCMV 68-1 was the only macaque CMV BAC available at the initiation of this project our lab attempted to utilize it in the available cynomolgus macaque model as the base of a new SIV vaccine [291, 401]. An infectivity study was undertaken to assess the viability of using RhCMV 68-1 to superinfect twelve CMV seropositive cynomolgus macaques from Mauritius [291].

Despite two high dose subcutaneous inoculations with enhanced green fluorescent protein (eGFP) expressing RhCMV 68-1, and a subsequent high dose intravenous inoculation in a subset of the cynomolgus macaques, no signs of infection were observed [291]. RhCMV 68-1 challenged macaques developed antibody and PBMC cytokine responses to RhCMV proteins and eGFP equivalent to controls challenged with inactivated virus [291]. Qualitative PCR confirmed endogenous CyCMV infection, but did not detect RhCMV gene translation [291]. This author successfully cultured CMV from biological secretions, but eGFP could not be detected visually or by western blot in *ex-vivo* cultures [291].

To confirm the addition of eGFP did not affect the fitness of the RhCMV, this author sequenced the complete virus by next-generation sequencing, and assembled the reads to the expected sequence using a Burrows-wheeler alignment algorithm [291]. Though RhCMV 68-1 differs from primary RhCMV isolates on a number of genes, no changes could be identified between RhCMV 68-1 and RhCMV 68-1 eGFP that are known to affect virulence [291, 379].

It has since been found that the species specificity of RhCMV 68-1 may be due not to intrinsic cross-species specificity but to attenuation now inherent in the RhCMV 68-1 genome [379]. RhCMV

68-1 cannot productively infect cynomolgus macaque cells as the virus lacks an equivalent of the anti-apoptotic *UL36*, and cell entry factors *UL128-UL131* [379]. Repair of these features in the RhCMV 68-1 BAC allows productive infection of, and viral shedding from cynomolgus macaques [379]. It is unknown how natural this infection is. To date no comparison of CMV viral load or in-host growth kinetics have been undertaken to compare infection with RhCMV in cynomolgus macaques to that in rhesus macaques. An important lesson of the HIV vaccine STEP trial (see section 1.1.3.1) is that quality of an immune response against a vector that does not naturally infect the animal model may not predict the response in humans [46]. We have concluded that working in a cynomolgus macaque model with a cynomolgus macaque strain of CMV will most accurately model natural CMV host dynamics. Further, the least attenuated strain of CMV, with the most natural host-CMV dynamics, will best predict the natural interaction between HCMV and its human hosts.

1.5 Thesis Objectives

1.5.1 Thesis Objectives

Overview: The overall aim of this thesis is to prepare a Mauritius strain of CyCMV as a SIV vaccine for use in cynomolgus macaques in order to evaluate the use of HCMV vectored vaccines against HIV. To this end a strain of CyCMV has been isolated from cynomolgus macaques from the island of Mauritius. This CyCMV strain was sequenced in full, and compared alongside other sequenced CMV strains to confirm infection with CyCMV Mauritius will parallel natural infection with HCMV in humans. The virus has been cloned as a BAC to permit rapid, site-specific modifications of the viral genome. The CyCMV Mauritius BAC was then modified to express two SIV vaccine transgenes throughout viral infection that permit the use of CyCMV Mauritius as a putative vaccine against SIV in the cynomolgus macaque model.

Main Objective: To prepare a Mauritius strain of CyCMV as a SIV vaccine for evaluation in cynomolgus macaques.

Specific Objectives:

1. Sequencing and genomic characterization of a Mauritian Cynomolgus Macaque strain of CMV (Chapter 2)
2. Comparative phylogenetic analysis of CyCMV Mauritius (Chapter 2)
3. Cloning of CyCMV Mauritius as an excisable BAC (Chapter 3)
4. Utilization of CyCMV Mauritius to produce an SIV vaccine candidate (Chapter 3)

Chapter Two: A Novel Strain of Cynomolgus Macaque Cytomegalovirus: Implications for Host-Virus Co-Evolution

Justen N. Hoffman Russell[1], Angie K. Marsh[1], David O. Willer[2, 3], Aruna P. N. Ambagala[2, 3,§], Misko Dzamba[4], Jacqueline K. Chan[2, 3], Richard Pilon[5], Jocelyn Fournier[6], Michael Brudno[4], Joseph M. Antony[2,3], Paul Sandstrom[5], Ben J. Evans[7] and Kelly S. MacDonald[1, 2,3,8*]

[1]Department of Immunology, University of Toronto, ON, Canada, M5S 1A8; [2]Department of Medicine, University of Toronto, ON, Canada, M5S 1A8 [3]Department of Microbiology, Mount Sinai Hospital, ON, Canada, M5G 1X5 [4]Department of Computer Science, University of Toronto, ON, Canada, M5S 1A8 [5]National HIV & Retrovirology Laboratories, Public Health Agency of Canada, ON, Canada, K1A 0K9 [6]Scientific Services Division, Health Products & Food Branch, Health Canada, ON, Canada K1A 0L2 [7]Biology Department, McMaster University, Life Sciences Building, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4K1.[8] Section of Infectious Diseases, Department of Internal Medicine, University of Manitoba, 745 Bannatyne Ave, Winnipeg, Manitoba, Canada, R3E 0J9

[§]Present Address: Canadian Science Centre for Human and Animal Health. National Centre for Foreign Animal Disease, 1015 Arlington Street, Winnipeg, Manitoba, R3E 3M4.

Published in BMC Genomics, 2016, 17:277,

Edited in conformity with the requirements of a thesis for the degree of Doctor of Philosophy.

2.1 Overview

A wildtype strain of CMV isolated from a cynomolgus macaque of Mauritian origin and minimally passaged was sequenced in full. ORFs were identified in the CyCMV sequence and used alongside known sequences of primate CMVs to confirm the fitness of the virus. The sequences were further utilized to understand the evolution of primate CMVs alongside their hosts and ascertain the relationship between the new CyCMV strain and other models of HCMV.

Main Objective: To prepare a Mauritius strain of CyCMV as a SIV vaccine for evaluation in cynomolgus macaques.

Specific Objectives:

1. **Sequencing and genomic characterization of a Mauritian Cynomolgus Macaque strain of CMV (Chapter 2)**
2. **Comparative phylogenetic analysis of CyCMV Mauritius (Chapter 2)**
3. Cloning of CyCMV Mauritius as an excisable BAC (Chapter 3)
4. Utilization of CyCMV Mauritius to produce an SIV vaccine candidate (Chapter 3)

2.2 Abstract

Background

CMV belongs to a large, ancient, genus of DNA viruses comprised of a wide array of species-specific strains that occur in diverse array of hosts.

Methods

In this study we sequenced the ~217 Kb genome of a CMV isolated from a Mauritius cynomolgus macaque, CyCMV Mauritius, and compared it to previously sequenced CMV from a cynomolgus

macaque of Filipino origin (CyCMV Ottawa) and two from Indian rhesus macaques (RhCMV 180.92 and RhCMV 68–1).

Results

Though more closely related to CyCMV Ottawa, CyCMV Mauritius is less genetically distant from both RhCMV strains than is CyCMV Ottawa. Several individual genes, including homologues of CMV genes *RL11B*, *UL123*, *UL83b*, *UL84* and a homologue of mammalian COX-2, show a closer relationship between homologues of CyCMV Mauritius and the RhCMVs than between homologues of CyCMV Mauritius and CyCMV Ottawa. A broader phylogenetic analysis of 12 CMV strains from eight species recovers evolutionary relationships among viral strains that mirror those amongst the host species, further demonstrating co-evolution of host and virus.

Conclusions

Phylogenetic analyses of rhesus and cynomolgus macaque CMV genome sequences demonstrate co-speciation of the virus and host.

2.3 Background

Macaque monkeys are an important animal model in biomedical research, particularly in infectious diseases. They are used in vaccine and infectivity studies of viruses that either do not infect or do not produce realistic pathogenic features in small animal models – viruses such as HIV, where SIV or SHIV can be studied as a surrogate in macaques.

Recently, the limited availability of Indian rhesus macaques in North America and elsewhere has led to the development of resources and standards for the use of cynomolgus macaques and Chinese rhesus macaques [428], the former in particular being an excellent model of HIV-1/SIV infection [494]. Though relatively closely related, divergence between populations or species of macaques and their corresponding CMV strains is substantial enough as to preclude direct replacement in any study. In fact, with increased usage in research, it has been questioned whether cynomolgus macaques may be in need of taxonomic re-classification [430]. Captive populations of cynomolgus macaques are

frequently interbred with little concern as to their origins. However, the genetic divergence between geographically distinct populations of cynomolgus macaques rivals that found between Indian and Chinese rhesus macaques, a factor that has the potential to differently influence experimental results [430]. For example, disease susceptibility is highly variable between isolated-tightly knit groups of macaques found across the Sunda Shelf of Southeast Asia. Macaque genome sequencing has enabled a better understanding of their disease susceptibility from an evolutionary and conservation standpoint, and provided insight into population structure and patterns of migration [408, 416, 430, 496].

Interest in non-human primate herpesviruses, and CMVs in particular, has increased in recent years in recognition of the unique immune response that they evoke. Specifically, CMV evokes a type of effector memory T cell response, which intermittently is boosted by reactivated virus throughout the life of its immunocompetent host. This type of immune memory qualitatively is different from conventional T cell memory in that it does not require priming and can respond immediately to antigen. Thus CMVs are also of interest to vaccinologists who are examining ways to expand the duration of rapidly inducible T cell responses to block primary viral infection. The evaluation of CMVs in rhesus and more recently, in cynomolgus macaques, has become a priority in order to facilitate the use of these animal models for vaccine research and development. CMV infects many primate and non-primate hosts including humans, baboons, green monkeys, chimpanzees, squirrel monkeys, macaques, oysters and rodents [290]. Since previous studies have demonstrated only limited mixing and horizontal transmission between mammalian populations [294], it is thought that CMV, like other herpesviruses, diversified via co-evolution with their host species [294, 497].

During speciation, pathogens may co-speciation with their host [498]. Herpesviruses, for instance, co-evolve with their mammalian hosts [499]. Herpesviruses were present in primates 70 mya [257], and have been used as surrogates to track mammalian – including human – evolution and migration [257, 500]. Herpesviruses, including CMVs, were also present in the most recent common ancestor of rhesus and cynomolgus macaques, and are presumed to have undergone geographic differentiation

similar to their host species [501]. However, macaques can become infected with multiple strains of CMV and these CMV strains may experience recombination. Hosts co-infected with two CMV strains occasionally release infectious viral particles (shed virus) from multiple strains (dual shedding), but generally have a dominant strain that makes up the majority of shed virus [502]. Interaction among populations of cynomolgus macaques can facilitate the exchange of CMV strains [503] but cross infection between species is generally rare [291].

In the present study, we sequenced a new strain of CyCMV from a cynomolgus macaque from Mauritius and compared it to three other strains – one from a cynomolgus macaque that originated in the Philippines: CyCMV Ottawa [289], and two from Indian rhesus macaques: RhCMV 68-1 [288] and RhCMV 180.92 [480]. We also characterized variation in gene content among these macaque CMVs. Our results indicate that evolutionary relationships over the complete genomes of cynomolgus CMV strains matches those of their hosts, thereby supporting their co-evolution and also the further use of CMVs in the study of mammalian biogeography and phylogeny. We also found phylogenetic correspondence between CMVs and their hosts in a broader phylogenetic analysis that included several other CMV strains. However, phylogenetic relationships among some genomic regions of the macaque CMVs deviated from the expected relationship, and this can be explained either by intra-strain recombination or alternatively could reflect phylogenetic error.

2.4 Materials and Methods

2.4.1 Isolating Virus

All animal procedures were performed as part of a series of experiments approved by the Health Canada National Nonhuman Primate Animal Care Committee. Urine samples were collected from a 15 year-old cynomolgus macaque imported from Mauritius using bladder catheterization. After filtration through a 0.45 μm filter, the urine samples were centrifuged at 900 x *g* for 30 mins at 4°C and supernatants were collected and mixed 1:1 with 2× Minimum Essential Media (MEM) supplemented with 2× antibiotic-

antimycotic (Gibco) and 20 mg/ml gentamycin (Invitrogen). The cell pellets were resuspended with 1 ml of Dulbecco's Modified Eagle's Medium (DMEM; Sigma) supplemented with 1× antibiotic-antimycotic, and 10 mg/ml gentamycin. Urine supernatants were ultracentrifuged at 20 000 x g for 30 mins at 4°C and the pellets resuspended in 500 µl of DMEM supplemented with 1× antibiotic-antimycotic and 10 mg/ml gentamycin.

CyCMV Mauritius virus was grown on human foetal lung fibroblast (MRC-5) cells [504] from the American Type Culture Collection (ATCC) as described previously [289]. MRC-5 cells were seeded 1:2 in a 12-well tissue culture plate and grown at 37˚C in a 5% CO_2 incubator for 2 days prior to inoculation. Resuspended viral pellets were each plated in triplicate in wells of MRC-5 cells, spinoculated at 2,000 x *g* for 30 mins at 4°C, and incubated for 2-3 hours at 37°C in a 5% CO_2 incubator. The inoculum was aspirated and replaced with 2 ml of DMEM [supplemented with 10% foetal bovine serum (FBS) (Wisent Bioproducts), 100 U/ml penicillin, 100 µg/ml streptomycin (Sigma)] and amphotericin B. The medium was changed the following day and every week thereafter. Cultures were monitored daily for CMV cytopathic effect (CPE). Cells showed 100% CPE after approximately 1 month. The virus was further propagated in MRC-5 cells at a 1:2 split ratio each time CPE reached 90-100% to achieve higher viral titers.

2.4.2 DNA Extraction

Viral DNA was isolated from cell free virus following centrifugation, digestion with pronase, and phenol-chloroform extraction. Fifteen 75-cm^2 tissue culture flasks containing CyCMV Mauritius-infected MRC-5 cells were used for viral DNA isolation using a Hirt extraction protocol as described previously [289]. To confirm the purity of the viral DNA isolation, 1 µg of isolated extracellular viral DNA was digested with 20 U of BamHI or HindIII restriction enzymes (New England Biolabs) and fractionated by gel electrophoresis on a 0.8% agarose gel.

2.4.3 Next Generation Sequencing

A paired-end library of CyCMV Mauritius DNA was prepared with 500 base pair (bp) insert size and high-throughput Illumina Genome Analyzer II paired end sequencing with 72 bp read length was performed as described previously [289]. The complete genome was sequenced at the Centre of Applied Genomics, Toronto, Canada.

2.4.4 Genome Assembly

De novo assembly of the CyCMV Mauritius genome was achieved from 2 268 822 reads on the Illumina Genome Analyzer II platform. Paired ends were filtered to match the barcode (1 134 411 paired reads) and assembled using Velvet (version 0.7.55) [505]. Results were best obtained using a kmer length of 39, an insert length of 500, and an expected coverage of 280X to generate multiple partial contigs. Three seed sequences, A: 5'-AAACCAGCGCCGTTGTTTTCCGTTCTACGTTCGGG-3',
B: 5'-ACTATCGAGGACAACGATGTTTTTTCCAACATAAA-3' and
C: 5'-TATCGGTATCTATTCCAAGCAGACCAAGTACGATT-3' were used to create contigs.

2.4.5 Gap Closing

Gap closing in the *de novo* genome build was accomplished using PCR. Primers were designed from the sequenced contigs, and amplified using purified viral DNA. The resulting product sizes were used to confirm the *de novo* DNA assembly.

2.4.6 Error Correction

Regions of low resolution or ambiguity were confirmed using Sanger sequencing. First, a 3 555 bp region (3 567-7 122 bp) was confirmed using the primers 5'-TCGGCAAAGTCAGGAGCGGC-3' and 5'-TGCACAATTTGCGATGCCTATCGTT-3'. Next, Sanger sequencing to amplify the 1 207 bp region spanning the origin of lytic replication (OriLyt) [82 561 bp–83 768 bp] was attempted using the primers 5'-TGGCGATCTGAAACCACACCCC-3' and 5'-CGCCCAAGAGAGAGCGCACC-3' but proved challenging due to the presence of inverted and repeated sequence motifs [506] as our previous experience

has shown [289]. Amplification of a 1 688 bp segment spanning 154 169 bp -155 856 bp using multiple primers also proved unsuccessful. However, a 560 bp region spanning 173 964 bp - 174 520 bp was amplified using the primers 5'-ACTTCGCTTCTGTTCTAGCGTTTAGG-3' and 5'-CCGCTGTGGCTTGCTGGCTC-3' and successfully confirmed by sequencing. The CyCMV Mauritius sequence was finally confirmed for errors by aligning with that of RhCMV 68-1 (accession: AY186194), RhCMV 180.92 (accession: DQ120516) and CyCMV Ottawa (accession: JN227533).

2.4.7 ORF Assignment

1408 putative ORFs, greater than 30 amino acids in length, and not contained within and in frame of another identified potential ORFs were identified *in silico* using Geneious Pro 6.1.4 (Biomatters Ltd., Auckland, New Zealand). BLAST-P (National center for biotechnology information; NCBI) with a BLOSUM62 matrix, gap opening cost of 11 and gap extension cost of 1 was used to screen the ORFs. Homologous protein lists were generated using a maximum E value of 10^{-1}. ORFs with irrelevant or no homology were not included.

2.4.8 Nomenclature

Putative genes were named with respect to homologues, as determined above, when applicable. Homologues of human genes are given the prefix 'Cy' followed by the region name: TRL, TRS, UL, or US and a number previously attributed to the homologue. Homologues of old world monkey CMV genes were given the prefix 'Cy' followed by 'O' and a gene number attributed to the homologue. Some of these were given alternative names, indicated in brackets, due to naming of CyCMV Ottawa homologues prior to this nomenclature [289]. Putative genes with previously named CyCMV Ottawa homologues were given the prefix 'Cy' followed by the previously assigned number. Putative genes lacking a CMV homologue were given the prefix 'Cy' followed by a unique number not used in naming within CyCMV Ottawa and corresponding to gene order within the genome. Capital letter suffixes (for example, *CyUL48***A**) indicate unique ORFs with a shared gene number as previously established [289]. Lower case single letter suffixes

(for example CyUS28**b**) indicate closely related or repeated genes that share homologous partners. Putative genes were assigned to one of 18 families, or designated Non-Core family, based on previously designated family of gene homologues.

2.4.9 Bit-Score Plots

Bit-scores were calculated using Geneious Pro 6.1.4 (Biomatters Ltd., Auckland, New Zealand) to compare individual CyCMV Mauritius or RhCMV 180.92 ORFs to homologous ORFs of CyCMV Ottawa (accession: JN227533), RhCMV 180.92 (accession: DQ120516), and RhCMV 68-1 (accession: AY186194) found using BLAST-P (NCBI) using a BLOSUM62 matrix, gap opening cost of 11, gap extension cost of 1 and a maximum E value of 10^{-1}. Calculated scores for individual ORF were plotted as seen previously [289] as a scatterplot comparing the bit-score for CyCMV Mauritius-CyCMV Ottawa, CyCMV Mauritius - RhCMV 180.92 and CyCMV Mauritius -RhCMV 68-1 to each other.

2.4.10 Gross Genome Comparison

CyCMV Mauritius has been deposited in the GenBank database (accession: KP796148). Other Genomes and annotations were obtained from GenBank using available strains as follows: CyCMV Ottawa (accession: JN227533), RhCMV 180.92 (accession: DQ120516), and RhCMV 68-1 (accession: AY186194). Comparisons were also made with human CMV strains [HCMV Towne (accession: AY315197), HCMV AD169 (accession: X17403) and HCMV HAN1 (accession: JX512199)], CCMV Heberling (accession: AF480884), Baboon CMV (BaCMV) OCOM4-37 (accession: AC090446), Cercopithecine herpesvirus 5 (GMCMV) Colburn (accession: FJ483969), Cercopithecine herpesvirus 5 (GMCMV) strain 2715 (accession: FJ483968), Aotine herpesvirus 1 strain S34E (accession: FJ483970), and Saimiriine (Squirrel monkey) betaherpesvirus 4 (Squirrel monkey CMV; SMCMV) strain SqSHV (accession: FJ483967). Genome identity between CMV strains was determined using global alignment with free end gaps set at a cost matrix of 65% similarity (5.0/-4.0), a gap open penalty of 12 and a gap extension penalty of 3 with automatic sequence direction determination on Geneious 6.1.4 (Biomatters Ltd., Auckland, New Zealand).

Whole genome alignments were generated using progressive MAUVE multiple genome alignment algorithm [507] available in Geneious 6.1.4 (Biomatters Ltd., Auckland, New Zealand).

2.4.11 Phylogenetic Analysis

Phylogenetic analysis was carried out with Geneious Pro 6.1.4 (Biomatters Ltd., Auckland, New Zealand). Alignments were created using MAFFT version 7.017 [508] with a scoring matrix of 200PAM/k=2 a Gap open penalty of 1.53 and offset value of 0.123. An appropriate phylogenetic model was selected for each tree using JModel test 2.1.7 [509]. Most trees were generated with the MrBayes plugin for Geneious Pro 6.1.4 [510] after 1 100 000 iterations and a burn-in of 110 000. The individual ORFs or whole genome sequences of CyCMV Mauritius, CyCMV Ottawa (accession: JN227533), RhCMV 68-1 (accession: AY186194), and RhCMV 180.92 (accession: DQ120516), were compared with the outgroup HCMV AD169 (accession: FJ527563) **(Figure 2-2**), Homo Sapiens PTGS2 (accession: BAA05698) (**Figure 2-16a**), Homo Sapiens CXCL5 (accession: CR457428) (**Figure 2-16b**), or Homo Sapiens TNFRSF14X1 (trimmed CDS from accession: XM_011542383) (**Figure 2-16c**). The 12 CMV full genome tree **(Figure 2-17)** was generated using MrBayes 3.2.4 run in MPI (Message Passaging Interface) on the SciNet supercomputer at the University of Toronto [511] using a generalized time reversible (GTR) tree with proportion invariant sites and gamma distribution. Genomes utilized were selected from fully sequenced CMV genomes at the time of publication due to similar organization. CMVs included were free from large rearrangements or inverted regions with respect to CyCMV Mauritius as determined from alignments using progressive MAUVE multiple genome alignment algorithm [507] in Geneious 6.1.4 (Biomatters Ltd., Auckland, New Zealand). RCMV Maastricht (accession: AF232689) was used as the out-group.

2.5 Results and Discussion

2.5.1 CyCMV Mauritius is a Unique CMV from Mauritian Cynomolgus Macaques

In the present study, we characterized a novel CMV strain, isolated from a Mauritian cynomolgus macaque, and compared its sequence to that of three other CMV strains, including that of CyCMV Ottawa [289], and two strains from rhesus macaques: RhCMV 68-1 [288] and RhCMV 180.92 [480]. Our analyses indicate that the sequence of CyCMV Mauritius is highly similar to the sequence of CyCMV Ottawa, and also that both are derived from a more recent common ancestor than either is with one of the rhesus macaque strains.

Our assembly of Illumina reads from CyCMV Mauritius produced three contigs that were then connected by Sanger sequencing. The assembly had an average 280X fold coverage that was 217 200 bp in length, and had 49.5% GC content. In comparison, CyCMV Ottawa is 841 bp longer and has a similar GC content (**Table 2-1**). RhCMV 68-1 and RhCMV180.92 are both similar in length and GC content, with RhCMV 180.92 being the shortest and RhCMV 68-1 the longest of the four viruses. Compared to other mammalian CMVs, all macaque CMVs examined are shorter than CCMV Heberling strain and shorter on average than HCMV strains AD169 and HAN1. Of the CMVs that have been sequenced thus far, these four macaque CMVs are most similar in length to Aotine (Owl Monkey) CMV (OMCMV) and GMCMV.

Table 2-1: Select Sequenced primate CMV strains

Virus	Strain	GenBank Accession	Host Species	Length (bp)	%GC
CyCMV	Mauritius	KP796148	Cynomolgus Macaque	217 200	49.5%
CyCMV	Ottawa	JN227533	Cynomolgus Macaque	218 041	49.5%
RhCMV	180.92	DQ120516	Rhesus Macaque	215 678	49.1%
RhCMV	68-1	AY186194	Rhesus Macaque	221 454	49.1%
OMCMV	S34E	FJ483970	Owl Monkey	219 474	56.3%
GMCMV	Colburn	FJ483969	African Green Monkey	219 526	51.2%
GMCMV	2715	FJ483968	African Green Monkey	226 205	50.8%
CCMV	Heberling	AF480884	Chimpanzee	241 087	61.7%
HCMV	AD169	FJ527563	Human	229 354	57.2%
HCMV	HAN1	JX512199	Human	235 006	62.4%

CyCMV Mauritius and CyCMV Ottawa share 95.3% identity by linear full genome alignment while RhCMV 68-1 and RhCMV 180.92 share 95.6% identity. CyCMV Mauritius and RhCMV 180.92 have 87.9% identity and CyCMV Mauritius and RhCMV 68-1 have 89.7% identity. CyCMV Ottawa has 89.8% identity with RhCMV 68-1 and 88.2% identity with RhCMV 180.92.

Due to the possibility of rearrangements in these viral genomes we evaluated sequence synteny using the progressive Mauve multiple genome alignment algorithm [507]. This analysis reveals a local co-linear region around 160-170 kbp where the RhCMV 68-1 sequence is reversed in comparison to CyCMV Mauritius, CyCMV Ottawa and RhCMV 180.92 (**Figure 2-1**). This region is immediately adjacent to the *UL128* to *UL130* deletion found in RhCMV 68-1. There are two locally collinear blocks whose order are reversed between rhesus (RhCMV 68-1 and RhCMV 180.92) and cynomolgus (CyCMV Mauritius and CyCMV Ottawa) macaque CMV genomes. Both collinear blocks are short and located immediately upstream of the reversed region in RhCMV 68-1.

2.5.2 Phylogenetic Relationships Among CMV Strains Match Those of Their Hosts.

Evolutionary relationships were estimated for CyCMV Mauritius, CyCMV Ottawa, RhCMV 180.92 and RhCMV 68-1 using HCMV AD169 as an outgroup (**Figure 2-2**). Evolutionary relationships among these complete macaque CMV genome sequences match those of their hosts (**Figure 2-2a)**. Within the cynomolgus macaque, the viral strains CyCMV Mauritius and CyCMV Ottawa are diverged by 0.036 substitutions per site (sps), whereas within the rhesus macaque, the viral strains RhCMV 180.92 and RhCMV 68-1 are diverged by only 0.017 sps. Thus, divergence between CMV strains of cynomolgus macaques is greater than between the CMV strains of rhesus macaques, and this may also be true of the respective host populations.

All four-macaque CMVs share similar genome architecture. Macaque CMV and the CMV of closely related primates, like GMCMVs Colburn and 2715, have terminal repeats, as well as unique long and unique short regions in genome structure, but lack the internal repeat regions of other primate CMVs [290].

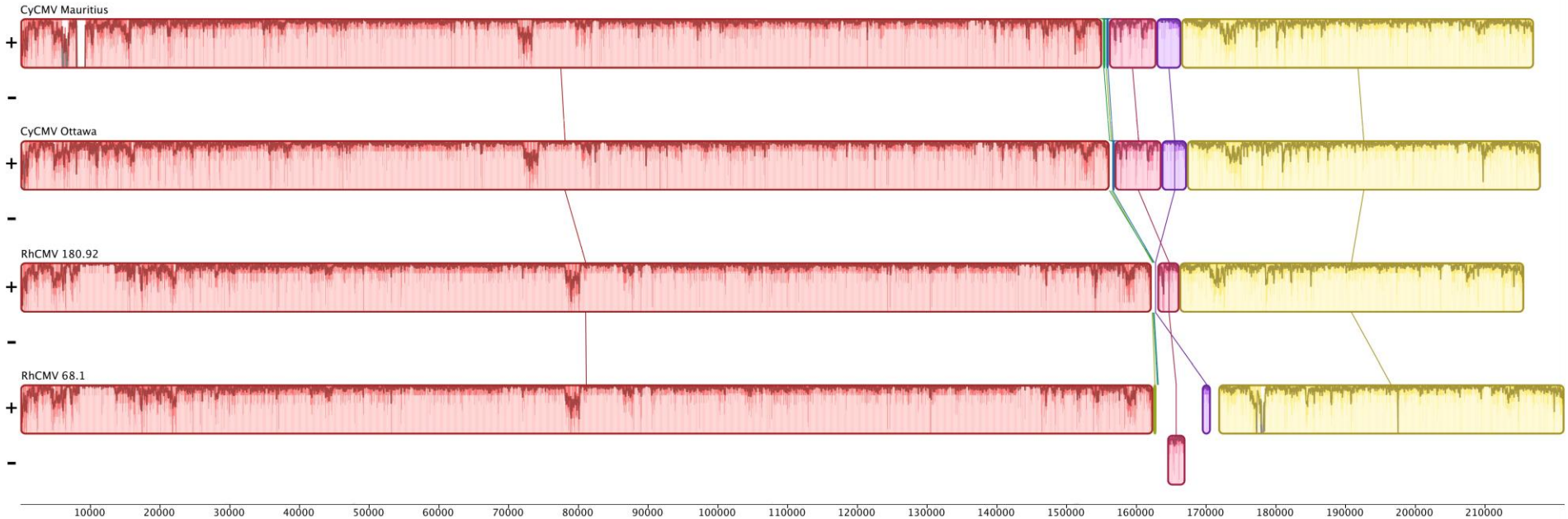


Figure 2-1: Multiple genome alignment of CyCMV Mauritius, CyCMV Ottawa, RhCMV 68-1, and RhCMV 180.92. Genome homology and rearrangement are presented using the MAUVE multiple genome alignment algorithm. Locally collinear blocks are indicated differentiated by colour and connected by linking line. Within collinear blocks column height within collinear blocks indicates average conservation of base pairs locally between homologous collinear blocks in other strains, full columns indicate regions fully conserved between species while empty columns indicate unique regions to the particular genome or absences. Collinear blocks shown above the x-axis are in the same orientation, with those shown below reversed in comparison to the same co-linear blocks in other species.

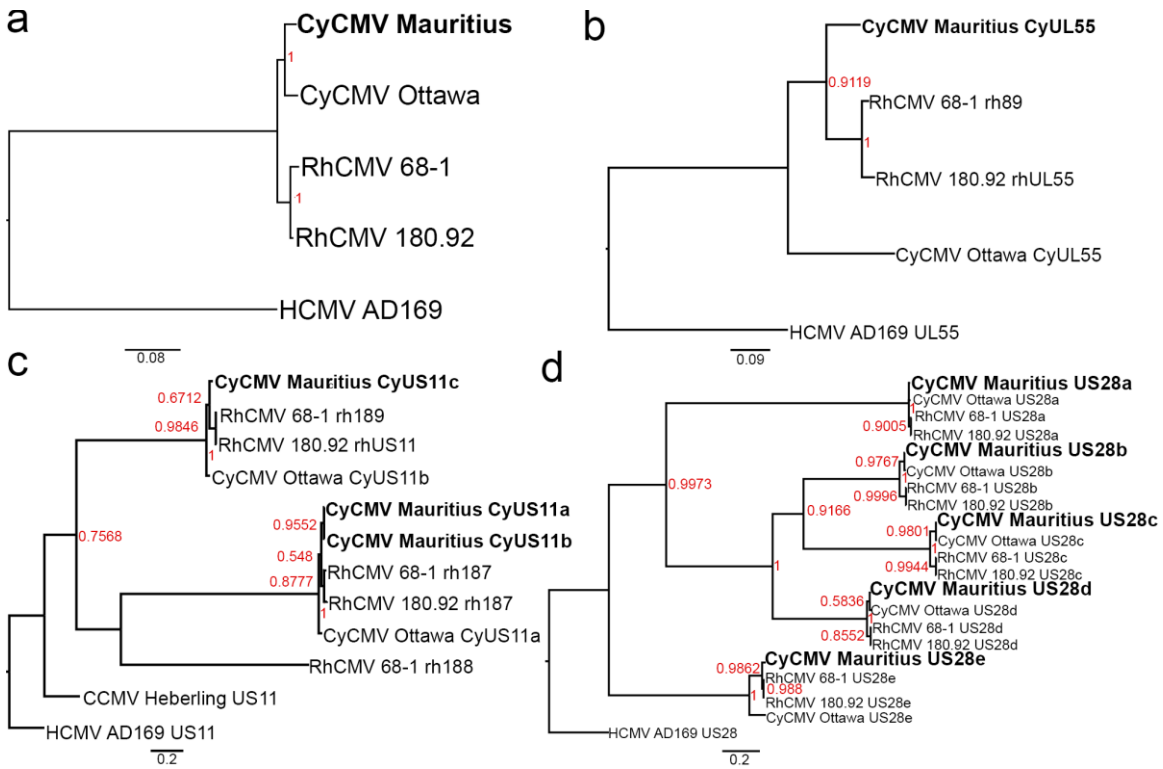


Figure 2-2: Phylogenetic tree comparison of Macaque CMV genome and select genes. Trees were generated comparing CyCMV Mauritius, CyCMV Ottawa, RhCMV 68-1 and RhCMV 180.92 using HCMV AD169 as an out-group. Trees were generated using Mr. Bayes, following MAFFT alignment, using a model of evolution selected by JModel test a) from a full genome global alignment; b) from an alignment of putative gB (UL55) ORF, a typical marker of phylogeny; c) from an alignment of US11, an MHC-I downregulatory gene necessary for superinfection of CMV, Chimpanzee CMV is included; d) from an alignment of US28 genes, for which 5 copies are encoded in each macaque CMV (4 copies of the gene and one copy of a non-functional pseudogene for RhCMV 68-1). CyCMV genes CyUS28a-CyUS28e and RhCMV genes rh214-rh220 are relabeled as US28a – US28e, inclusive of the RhCMV 68-1 pseudogene homologous to other macaque CMV US28c, based on order in genome. Numbers at loci indicate posterior probability with color scaled according to probability; genetic distances measured in sps are given by scale below.

2.5.3 CyCMV genomes show diversity in gene conservation

Viral genes that are important for host specificity may tend to be more diverged among strains in different hosts. To explore these possibilities, we compared divergence and gene content of the four completely sequenced macaque CMV strains in order to better understand how these factors varied among strains. Annotation of the sequenced CyCMV Mauritius genome identified 290 putative ORFs (**Figure 2-3;**

Table 2-2). By contrast, CyCMV Ottawa has 262 putative ORFs [289], RhCMV 68-1 has 230 ORFs [288], and RhCMV 180.92 has 258 ORFs [480]. To better understand the conservation of individual ORFs in these CMV genomes we used bit-score, a log-scaled measure indicating the size of a random search string required to find an equivalently or more similar sequence than the observed match. Bit-scores were generated by comparing ORFs in each of the six pairwise comparisons between the four macaque CMV genomes. Bit-scores comparing homologous ORFs of CyCMV Mauritius and CyCMV Ottawa, RhCMV 68-1 or RhCMV 180.92 were examined grouped by gene family (**Figure 2-4).** The average bit-score for genes in each family, but not the bit-score of each individual gene, is higher for the comparison of genes between the two CyCMVs than between CyCMV Mauritius and RhCMV 180.92 or RhCMV 68-1. The only exception to this is the COX-2 family, which is absent from CyCMV Ottawa. Two-dimensional bit-score plots were used to further visualize variation in individual gene conservation in the pairwise comparisons between the four macaque CMV strains (**Figure 2-5**, **Figure 2-6**, **Figure 2-7**, **Figure 2-8** and **Figure 2-9),** as in [289]. When plotted this way, ORFs found along the *x=y* diagonal have equivalent sequence conservation in both pairwise comparisons. Genes that are equally well conserved between the two RhCMVs and between the two CyCMVs (Bit Score >1000 in one comparison) cluster along the diagonal. However, this analysis of pairwise comparisons highlights a non-uniform level of sequence conservation in some genes. The cyclooxygenase-2 gene (*CyCOX2*), for example, is absent in CyCMV Ottawa. Similarly *CyTRL1,* an epithelial cell tropism factor [512], is more highly conserved between the RhCMVs than between the two CyCMVs. There was a larger discordance between the bit-scores generated comparing CyCMV genes and those comparing RhCMV for genes with a lower maximum bit-score. Of those genes with known functions, those involved in immune modulation or as temperance factors, and membrane proteins are more dispersed than other groups with many individual genes being more highly conserved in viruses from one species of macaque than the other. This could suggest that these viral genes face differing evolutionary pressures in the two macaque species.

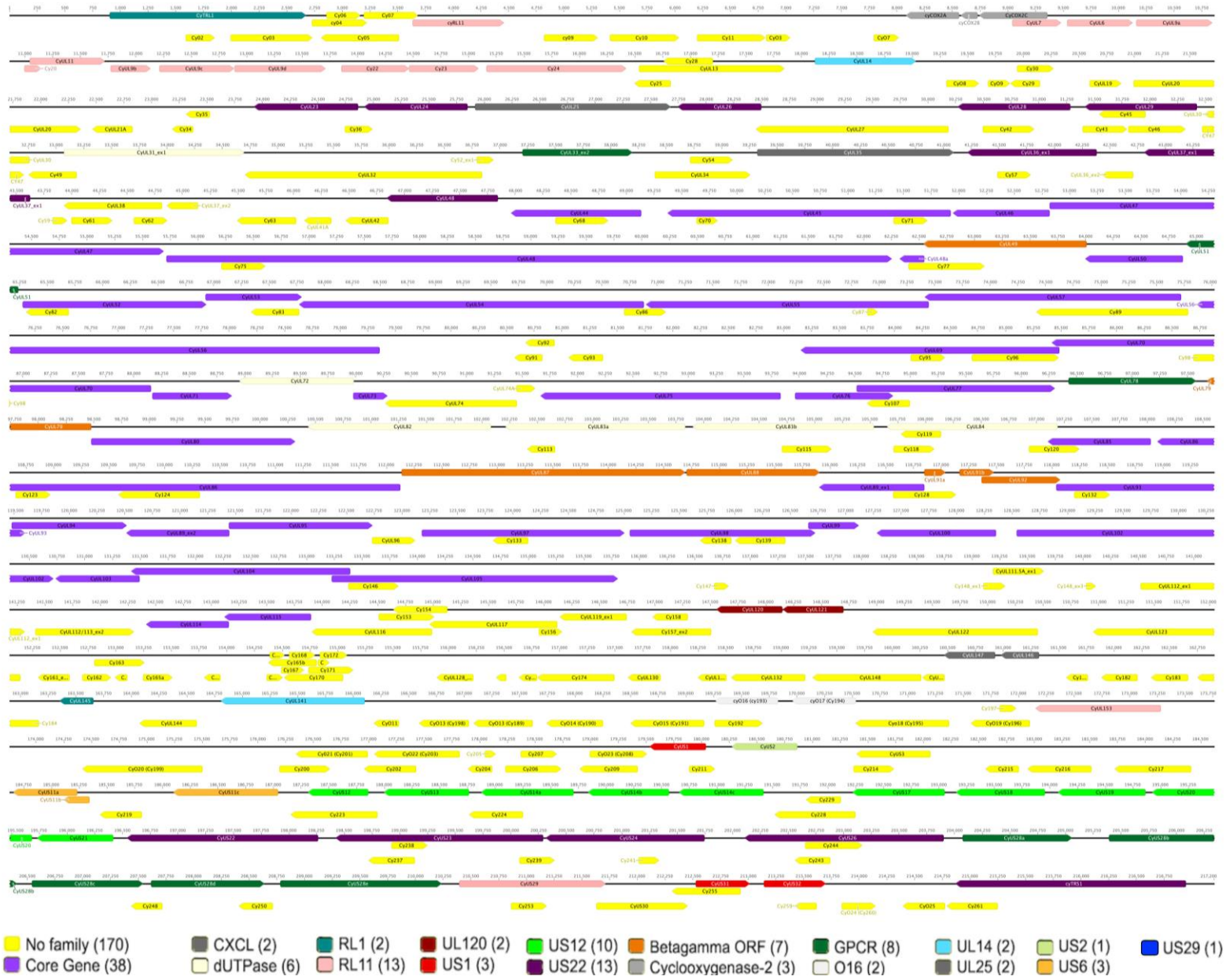


Figure 2-3: **Map of ORFs in CyCMV Mauritius genome.** CyCMV Mauritius encodes 290 putative ORFs that are annotated by gene name and colour coded based on gene families. Of the CyCMV Mauritius ORFs, 268 (92%) share homologues with CyCMV Ottawa, 239 share homologues with RhCMV 68-1 (82%), and 158 (54%) share homologs with HCMV strains. CyCMV Mauritius like RhCMV but unlike CyCMV Ottawa or HCMV contains ORFs with homology to COX-2. CyCMV ORFs with an HCMV homologue are annotated by "Cy" followed by the HCMV name. Arrowheads indicate the directions of the ORFs. Core genes are herpes virus core genes.

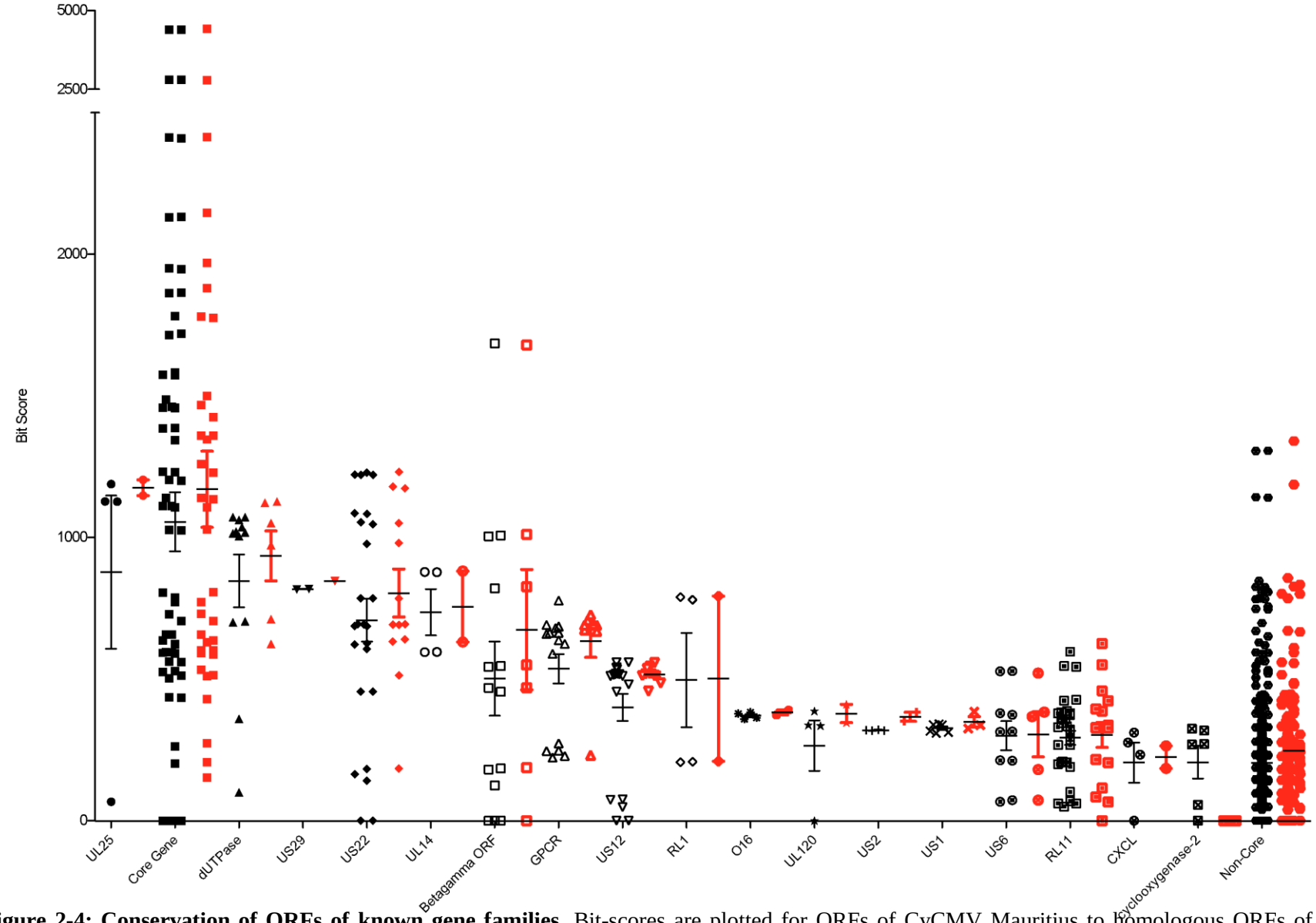


Figure 2-4: Conservation of ORFs of known gene families. Bit-scores are plotted for ORFs of CyCMV Mauritius to homologous ORFs of RhCMV 68-1 and RhCMV 180.92 (Black) and to homologous ORFs of CyCMV Ottawa (Red) grouped by gene gamily. Each point represents a single bit-score comparison for a single ORF. Horizontal line indicates the mean ± standard deviation.

Table 2-2: ORFs of CyCMV Mauritius

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyTRL1	RL1	-	903	2675	+	590	65.808	RL01 [24]	CyTRL1 (90.4%)	Rh01 (85.9%)	RhRL1 (86.7%)
Cy02	-	-	1589	1849	+	86	9.794	-	Cy02 (85.3%)	Rh02 (76.8%)	Rh2 (78%)
Cy03	-	-	1993	2730	+	245	26.341	-	Cy03 (69.4%)	-	Rh2.1 (71%)
Cy04	-	-	2727	3224	+	165	18.209	-	Cy04 (95.2%)	Rh03 (91.3%)	Rh3 (90.7%)
Cy05	-	-	3510	2812	-	232	27.064	-	Cy05 (97%)	Rh04 (90.9%)	Rh4 (90.5%)
CyO1 (Cy06)	-	-	2860	3162	+	100	11.511	-	Cy06 (98%)	-	Rh3.1 (94%)
Cy07	-	-	3197	3682	+	161	18.051	-	Cy07 (96.9%)	-	Rh3.2 (92%)
CyRL11A	RL11	IgG Fc binding glycoprotein modulator of antibody activity	3637	4458	+	273	30.151	-	CyRL11 (97.4%)	Rh05 (96%)	Rh5 (95.6%)
CyRL11B (Cy09)	-	-	4821	5303	+	160	18.479	-	Cy09 (44.2%)	Rh06 (86.4%)	Rh6 (86.2%)
CyRL11C (Cy10)	-	-	5415	6038	+	208	23.627	-	Cy10 (58.6%)	Rh07 (59.5%)	Rh7 (60%)
CyRL11D (Cy11)	-	-	6201	6812	+	203	23.135	-	Cy11 (31.2%)	Rh08 (33.8%)	Rh8 (29.1%)
CyO3	-	-	6821	7039	+	72	8.144	-	-	-	-
CyO7	-	-	7795	8019	+	74	8.419	-	-	-	-
CyCOX2A	COX-2	Putative COX-2 Homologue[6]	8563	8081	-	160	16.88	-	-	Rh10 (94.3%)	Rh10 (94.3%)
CyCOX2B	COX-2	Putative COX-2 Homologue[6]	8728	8579	-	49	6.126	-	-	Rh10 (92.6%)	-
CyCOX2C	COX-2	Putative COX-2 Homologue[6]	9357	8743	-	204	23.458	-	-	Rh10 (98%)	Rh10 (98.1%)
CyUL7	RL11	Putative membrane glycoprotein	9040	9480	+	146	16.945	-	CyUL7 (90%)	Rh11 (89.9%)	Rh11 (89.9%)
CyUL6	RL11	Putative membrane glycoprotein	9538	10128	+	192	21.782	UL6 [34]	CyUL6 (82.2%)	Rh20 (88.8%)	RhUL6 (88.9%)
CyUL9a	RL11	Temperance factor & Putative membrane glycoprotein	10160	10843	+	227	26.516	UL9 [10]	CyUL9 (97.4%)	Rh21 (61.8%)	Rh21 (66.7%)
Cy20	RL11	Putative membrane glycoprotein[6]	10997	11146	+	49	6.069	-	-	Rh22 (58.7%)	Rh22 (66.7%)
CyUL11	RL11	Membrane glycoprotein	11044	11721	+	225	25.308	UL9 [23]	CyUL11 (85.4%)	Rh23 (79.7%)	RhUL1 (80.6%)
CyUL9b	RL11	Temperance factor & Putative membrane glycoprotein	11772	12134	+	120	13.631	UL9 [21]	CyUL9 (95%)	Rh24 (88.3%)	Rh24 (88.3%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL9c	RL11	Temperance factor & Putative membrane glycoprotein	12213	12887	+	224	24.673	UL9 [26]	CyUL9 (92%)	Rh25 (92%)	Rh26 (29.5%)
CyUL9d	RL11	Temperance factor & Putative membrane glycoprotein	12892	13710	+	272	31.067	UL9 [21]	CyUL9 (89%)	Rh26 (81.4%)	Rh26 (81.8%)
CyRL11Q (Cy22)	RL11	-	13855	14463	+	202	22.99	-	Cy22 (93.6%)	Rh27 (93.6%)	Rh27 (90.1%)
CyRL11R (Cy23)	RL11	-	14465	15088	+	207	23.948	-	Cy23 (75.4%)	Rh28 (86%)	Rh28 (88.4%)
CyRL11S (Cy24)	RL11	Putative membrane glycoprotein[3]	15163	16422	+	419	45.885	-	Cy24 (70.5%)	Rh29 (69.1%)	Rh29 (70.7%)
Cy25	-	-	16821	16498	-	107	11.613	-	Cy25 (97.2%)	Rh30 (97.2%)	Rh30 (96.1%)
CyUL13	-	Putative secreted protein	16538	17848	+	436	50.654	-	CyUL13 (94.7%)	Rh31 (94%)	Rh31 (94.3%)
Cy28	-	-	17203	16760	-	147	16.953	-	Cy28 (95.8%)	Rh32 (89.1%)	Rh32 (89.1%)
CyUL14	UL14	Putative membrane glycoprotein	18123	19034	+	303	35.334	UL14 [31]	CyUL14 (99.7%)	Rh33 (98%)	RhUL1 (97%)
CyO8	-	-	19309	19599	+	96	11.399	-	-	-	-
CyO9	-	-	19683	19880	+	65	7.73	-	-	-	-
Cy29	-	-	20149	19886	-	87	9.952	-	-	Rh34 (95.4%)	Rh34 (95.4%)
CyUL17 (Cy30)	-	Putative 7-transmembrane protein	19946	20272	+	108	12.773	-	Cy30 (95.4%)	Rh35 (96.3%)	-
CyUL19	-	-	20598	20885	+	95	19.748	UL19 [22]	CyUL19 (96.8%)	Rh35.1 (95.8%)	RhUL1 (94.7%)
CyUL20	-	T cell receptor gamma chain homologue	20994	22364	+	456	51.387	UL20 [13]	CyUL20 (85.3%)	Rh36 (83.3%)	RhUL2 (83.8%)
CyUL21A	-	Fibroblast temperance factor & CC chemokine-binding protein	22831	22468	-	120	13.841	UL21A [9]	CyUL21A (99.2%)	Rh37 (98.3%)	RhUL2 (99.2%)
Cy34	-	-	23375	23187	-	62	6.33	-	Cy34 (68%)	-	-
Cy35	-	-	23529	23314	-	71	8.232	-	-	Rh39 (88.9%)	Rh39 (71.4%)
CyUL23	US22	Temperance factor & Tegument protein	24864	23926	-	312	35.906	UL23 [34]	CyUL23 (98.1%)	Rh40 (95.8%)	-
Cy36	-	-	24746	24994	+	82	9.49	-	Cy36 (97.6%)	Rh41 (92.7%)	Rh41 (91.5%)
CyUL24	US22	Temperance factor & Tegument protein	25850	24921	-	309	35.228	UL24 [34]	CyUL24 (100%)	Rh42 (98.1%)	UL24 (95.8%)
CyUL25	UL25	Temperance factor & Tegument phosphoprotein	25917	27683	+	588	67.28	UL25 [31]	CyUL25 (98.3%)	Rh43 (95.1%)	RhUL25 (95.6%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL26	US22	Transcriptional activator of major immediate early promoter & regulator of tegument protein phosphorylation	28499	27744	-	251	28.188	UL26 [35]	CyUL26 (98.8%)	Rh44 (97.5%)	RhUL26 (97.5%)
CyUL27	-	Maribavir resistance	30186	28453	-	577	65.694	pUL27 [43]	CyUL27 (99.1%)	Rh46 (97.6%)	RhUL27 (97.6%)
CyUL28	US22	-	31284	30271	-	337	38.679	UL28 [8]	CyUL28 (98.8%)	Putative Rh50/47 (98.2%)	RhUL28 (98.5%)
Cy42	-	-	30496	30957	+	153	17.046	-	Cy42 (91.1%)	Rh48 (87.6%)	-
Cy43	-	-	31396	31794	+	132	15.247	-	Cy43 (96.3%)	Rh49 (91.7%)	Rh49 (92.4%)
CyUL29	US22	Temperance factor	32424	31414	-	336	38.843	UL29 [8]	CyUL29 (99.4%)	Rh50 (98.5%)	UL29 (98.2%)
Cy45	-	-	31961	31545	-	138	15.642	-	Cy45 (97%)	Rh51 (93.3%)	Rh51 (94%)
Cy46	-	-	31810	32322	+	170	19.225	-	Cy46 (96.5%)	Rh52 (92.1%)	-
Cy47	-	-	32476	32709	+	77	9.002	-	Cy47 (100%)	Rh52 (100%)	-
CyUL30	-	-	32764	32498	-	88	10.32	UL30 [33]	CyUL30 (100%)	Rh50.1 (96.4%)	-
Cy49	-	-	33188	32754	-	144	17.408	UL30A [8]	Cy49 (97.7%)	Rh53 (94.8%)	Unknown (94.5%)
CyUL31	dUTPase	Immediate early protein	33071	34696	+	541	61.074	UL31 [35]	CyUL31 (99.3%)	Rh54 (97.8%)	UL31 (94.7%)
CyUL32	-	Major Tegument phosphoprotein	36839	34707	-	710	79.61	pp150 (UL32) [35]	CyUL32 (97.3%)	Rh55 (92.4%)	RhUL32 (92.5%)
Cy52_ex1	-	-	36793	36945	+	50	6.024	-	Cy52 exon 1 (100%)	-	G protein coupled UL33 like (94%)
CyUL33_ex2	GPCR	Constitutive signalling	37207	38196	+	329	37.331	-	CyUL33 (98.5%)	G protein coupled UL33 like (93%)	RhUL33 exon 2 (95.7%)
CyUL34	-	Repressor of US3 transcription	38401	39258	+	285	32.858	UL34 [34]	CyUL34 (99.6%)	Rh57 (96.6%)	RhUL34 (99.3%)
Cy54	-	-	38715	39101	+	128	13.906	-	Cy54 (93%)	Rh58 (91.4%)	Rh58 (93.8%)
CyUL35	UL25	Tegument phosphoprotein, Virion transactivation and assembly regulation & interacts with UL82	39321	41093	+	590	66.924	UL35 [9]	CyUL35 (99.3%)	Rh59 (98.1%)	RhUL25 (22%)
CyUL36_ex1	US22	Immediate early tegument protein / Inhibitor of caspase-8-induced apoptosis	42380	41217	-	387	44.731	UL36 [44]	CyUL36 (97.4%)	Rh60 (97.4%)	RhUL36 (97.7%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
Cy57	-	-	41487	41786	+	99	11.289	-	Cy57 (93.9%)	-	Rh59.1 (89.9%)
CyUL36_ex2	US22	Immediate early tegument protein / Inhibitor of caspase-8-induced apoptosis	42712	42434	-	92	10.341	UL36 [34]	CyUL36 (97.8%)	Rh61 (94.6%)	RhUL36 (97.4%)
CyUL37_ex1	-	Immediate early protein / Mitochondrial inhibitor of apoptosis	43628	42810	-	272	31.148	UL37 [40]	CyUL37 (97.1%)	Rh62 (98.2%)	RhUL37 (98.2%)
Cy59	-	-	43830	43961	+	43	5.049	-	Cy59 (82.5%)	Rh63 (77.5%)	-
CyUL38	-	Virion Envelope Glycoprotein	44816	43932	-	294	33.344	UL38 [35]	CyUL38 (99.3%)	Rh64 (96.6%)	RhUL38 (96.9%)
Cy61	-	-	44001	44366	+	121	13.733	-	Cy61 (95.9%)	Rh65 (94.2%)	Rh65 (93.4%)
Cy62	-	-	44562	44861	+	99	11.442	-	Cy62 (91.9%)	-	Rh65.1 (91.9%)
CyUL37_ex2	-	Immediate early protein / Mitochondrial inhibitor of apoptosis	45142	44858	-	94	10.703	UL37 [45]	CyUL37 (94.8%)	Rh66 (93.6%)	RhUL37 (93.4%)
CyUL40 (Cy63)	-	-	46023	45493	-	176	18.952	-	Cy63 (100%)	Rh67 (93.2%)	-
CyUL41A	-	Virion Envelope Protein	46340	46101	-	79	9.408	UL41A [31]	CyUL41A (100%)	Rh67.1 (97.5%)	RhUL41 (98.7%)
CyUL42	-	Putative membrane protein	46858	46472	-	128	14.102	UL42 [20]	CyUL42 (91.4%)	Rh68 (92.3%)	RhUL42 (91.4%)
CyUL43	US22	Tegument protein / Putative multiple transmembrane protein	47843	46842	-	333	38.49	UL43 [31]	CyUL43 (99.7%)	Rh69 (98.2%)	-
CyUL44	Core Gene	DNA polymerase accessory subunit / Increases DNA polymerase product length	49134	47962	-	390	44.011	UL44 [8]	CyUL44 (99%)	Rh70 (98.5%)	-
Cy68	-	-	48364	48840	+	158	17.891	-	Cy68 (96.2%)	Rh71 (93.7%)	-
CyUL45	Core Gene	Enzymatically inactive large subunit ribonucleotide reductase homologue tegument protein	51927	49375	-	850	96.916	UL44 [33]	CyUL45 (99.6%)	Rh72 (98.1%)	Rh72 (97.9%)
Cy70	-	-	49636	49827	+	63	6.807	-	Cy70 (98.4%)	Rh73 (92.1%)	-
Cy71	-	-	51413	51715	+	100	11.284	-	Cy (0%)	Rh74 (99%)	-
CyUL46	Core Gene	Component of Capsid triplexes (minor capsid binding protein)	52818	51946	-	290	33.128	UL46 [31]	CyUL46 (99.7%)	Rh75 (98.3%)	-
CyUL47	Core Gene	UL48 binding tegument protein involved in intracellular capsid transport	52817	55693	+	958	110.632	UL37 [28]	CyUL47 (99.1%)	Rh76 (97.7%)	RhUL47 (97.6%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL48	Core Gene	Largest Capsid protein involved in intracellular capsid transport transport	55714	62253	+	2179	246.998	UL48 [24]	CyUL48 (99.2%)	Rh78 (98.5%)	RhUL48 (98.6%)
Cy75	-	-	56213	56605	+	130	15.039	-	Cy75 (92.3%)	-	Rh78.1 (89.2%)
CyUL48A	Core Gene	Smallest Capsid protein (located on tip of hexons) putatively involved in capsid transport	62547	62326	-	73	8.439	UL48A [31]	CyUL48a (100%)	-	-
Cy77	-	-	62405	63091	+	228	24.686	-	Cy77 (99.5%)	Rh79 (99.1%)	-
CyUL49	Betagamma ORF	Membrane protein	64009	62540	-	489	56.125	UL49 [24]	CyUL49 (99.6%)	Rh80 (99.2%)	RhUL49 (99.4%)
CyUL50	Core Gene	Inner nuclear membrane protein involved in nuclear egress of capsids	64880	63999	-	293	32.346	UL50 [45]	CyUL50 (100%)	Rh81 (99.3%)	RhUL50 (99.3%)
CyUL51	GPCR	Terminase component (TER3) / DNA packaging	65241	64906	-	111	12.356	UL33 [8]	CyUL51 (100%)	Rh82 (99.1%)	RhUL51 (97.3%)
CyUL52	Core Gene	Putatively involved in capsid transport in nucleus	65279	66937	+	552	62.559	UL52 [33]	CyUL52 (98.7%)	Rh83 (98.7%)	RhUL52 (98.9%)
Cy82	-	-	65697	65311	-	128	13.632	-	Cy82 (98.8%)	Rh84 (98.8%)	Rh84 (98.8%)
CyUL53	Core Gene	Tegument protein / Nuclear matrix protein /Capsid and nuclear egress lamina protein	66930	67796	+	288	33.024	UL53 [35]	CyUL53 (99.3%)	Rh85 (98.6%)	-
Cy83	-	-	67772	67341	-	143	16.017	-	-	Rh86 (97.9%)	-
CyUL54	Core Gene	DNA polymerase catalytic subunit (POL)	70881	67774	-	1035	116.571	UL54 [32]	CyUL54 (99.9%)	Rh87 (98.9%)	RhUL54 (98.8%)
Cy86	-	-	70702	71079	+	125	14.097	-	Cy86 (96%)	Rh88 (96.8%)	Rh88 (96%)
CyUL55	Core Gene	Virion Envelope Glycoprotein B (gB) / Forms homomultimers / Heparin binding / Role in entry and signalling	73449	70900	-	849	97.466	UL55 [33]	CyGlycoprotein B (cyUL55) (77.7%)	Rh89 (88.7%)	RhUL55 (88%)
Cy87	-	-	72895	72987	+	30	3.516	-	-	Rh90 (72.4%)	Rh90 (72.4%)
CyUL57	Core Gene	ssDNA binding protein	75721	73415	-	768	88.185	UL57 [32]	CyUL57 (99.8%)	Rh92 (99.7%)	RhUL57 (99.6%)
Cy89	-	-	75785	74418	-	455	48.398	-	Cy89 (98.9%)	-	Rh91.1 (97.4%)
Cy90	-	Bat Herpesvirus Homology[6]	75666	75842	+	58	6.832	-	-	-	-

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL56	Core Gene	Terminase component (TER2) / binds DNA packaging motif & nuclease activity	79359	75868	-	1163	129.202	UL56 [26]	CyUL56 (99.1%)	Rh91 (99%)	RhUL56 (99.1%)
Cy91	-	-	80825	80577	-	82	9.184	-	Cy91 (87.8%)	Rh93 (86.6%)	Rh93 (86.6%)
Cy92	-	-	80934	80674	-	86	8.889	-	Cy92 (71.4%)	-	-
Cy93	-	-	81372	81064	-	102	10.76	-	Cy93 (98.9%)	Rh95 (96.7%)	Rh95 (97.8%)
CyUL69	Core Gene	Multiple regulatory protein tegument protein / Contributes to cell cycle block and nucleocytoplasmic export of unspliced mRNA	85484	83157	-	775	87.204	UL69 [8]	CyUL69 (98.1%)	Rh97 (96.7%)	RhUL69 (96.9%)
Cy95	-	-	84148	84456	+	102	11.793	-	Cy95 (100%)	Rh98 (99%)	-
Cy96	-	-	84699	85481	+	260	28.354	-	Cy96 (96.9%)	Rh99 (93.8%)	Rh99 (94.2%)
Cy98	-	-	86693	86884	+	63	7.075	-	Cy98 (88.9%)		Rh99.1 (88.9%)
CyUL70	Core Gene	DNA Helicase primase subunit involved in DNA unwinding	88156	85418	-	912	105.521	Ul70 [8]	CyUL70 (99.6%)	Rh100 (98.5%)	RhUL70 (98.5%)
CyUL71	Core Gene	Tegument protein putatively involved in Cytoplasmic egress	88169	88885	+	238	26.247	UL51 [20]	CyUL71 (98.3%)	Rh100.1 (95%)	RhUL71 (95.4%)
CyUL72	dUTPase	Enzymatically inactive dUTPase homologue virion protein	89984	88953	-	343	39.208	UL72 [37]	CyUL72 (99.7%)	Rh101 (98.3%)	RhUL72 (98%)
CyUL73	Core Gene	Virion Envelope Glycoprotein N (gN) / Complexes with gM involved in entry	89979	90290	+	103	11.713	UL73	CyUL73 (97.1%)	Rh102 (96.2%)	-
CyUL74	Non-Core	Virion Envelope Glycoprotein O (gO) / Complexes with gH:gL involved in entry	91452	90271	-	393	46.01	UL74 [11]	CyUL74 (98.2%)	Rh103 (95.4%)	RhUL74 (95.2%)
CyUL74A	Non-Core	Virion Envelope Glycoprotein 24	91451	91621	+	56	6.415	UL74A [33]	CyUL74A (100%)	-	-
CyUL75	Core Gene	Virion Envelope Glycoprotein H (gH) / Associates with gO & complexes with GL or UL128-131 involved in entry	93831	91669	-	720	81.442	UL75 [45]	CyUL75 (98.5%)	Rh104 (97.8%)	RhUL75 (97.6%)
CyUL76	Core Gene	Viral associated regulatory protein	93964	94848	+	294	32.85	UL24 [8]	CyUL76 (98.6%)	Rh105 (100%)	RhUL76 (99.7%)
CyUL77	Core	DNA packaging & Portal	94517	96304	+	595	67.331	UL25 [31]	CyUL77 (99.5%)	Rh106 (99.7%)	RhUL77 (99.5%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
	Gene	Capping protein									
Cy107	-	-	94996	94613	-	127	14.328	-	Cy107 (98.4%)	-	Rh106.1 (97.6%)
CyUL78	GPCR	Putative chemokine receptor protein	96431	97576	+	381	42.149	UL78 [35]	CyUL78 (92.9%)	Rh107 (91.3%)	RhUL78 (90.8%)
CyUL79	Betagamma ORF	-	98478	97678	-	266	30.455	UL79 [24]	CyUL79 (100%)	Rh108 (98.9%)	RhUL79 (99.2%)
CyUL80	Core Gene	Capsid maturation protease and capsid assembly scaffold protein precursor	98477	100315	+	612	66.39	Ul80 [37]	CyUL80 (98.7%)	Rh109 (98%)	RhUL80 (98%)
CyUL82	dUTPase / UL82	Tegument phosphoprotein (pp65) (lower matrix protein) / Virion transactivator	102080	100431	-	549	61.672	Ul82 [36]	CyUL82 (98.9%)	Rh110 (94.8%)	RhUL83 (24.8%)
CyUL83a	dUTPase / UL82	Major Tegument phosphoprotein (pp65) (upper matrix protein) / Interferon response suppression	103832	102210	-	540	62.106	UL83 [39]	CyUL83 (98.5%)	Rh111 (97%)	RhUL83 (96.9%)
Cy113	-	-	102658	102410	-	82	9.098	-	Cy113 (95.2%)	-	-
CyUL83b	dUTPase / UL82	Major Tegument phosphoprotein (pp65) / Interferon response suppression	105532	103901	-	543	61.741	UL83 [39]	CyUL83 (95.8%)	Rh112 (96.9%)	RhUL83 (97.2%)
Cy115	-	-	104705	105151	+	148	16.367	-	Cy115 (90.3%)	Rh113 (91.9%)	Rh113 (91.2%)
CyUL84	dUTPase/UL82	Nucleocytoplasmic shuttling and DNA replication organization/Binds IE2	107189	105651	-	512	57.359	UL84 [13]	CyUL84 (97.9%)	Rh114 (99.4%)	RhUL84 (99.2%)
Cy118	-	-	105712	106080	+	122	13.366	-	Cy118 (98%)	Rh115 (97.5%)	Rh115 (96.7%)
Cy119	-	-	106140	105781	-	119	13.13	-	Cy119 (91.6%)	-	Rh115.1 (92.4%)
Cy120	-	-	106931	107389	+	152	16.124	-	Cy120 (98%)	Rh116 (99.3%)	-
CyUL85	Core Gene	Capsid triplex subunit[2]	108030	107104	-	308	34.656	UL85 [31]	CyUL85 (99.4%)	Rh117 (100%)	-
CyUL86	Core Gene	Major capsid protein	112122	108091	-	1343	151.345	UL86 [28]	CyUL86 (99%)	Rh118 (99.4%)	RhUL86 (99.3%)
Cy123	-	-	108656	108970	+	104	11.804	-	Cy123 (99%)	Rh119 (97.1%)	-
Cy124	-	-	110317	109583	-	244	28.846	-	Cy124 (96.3%)	Rh120 (97.5%)	-
CyUL87	Betagamma	-	112137	114692	+	851	96.581	UL87 [35]	CyUL87 (99%)	Rh122 (98.7%)	-

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL88	ORF Betagamma	Terminase Component putatively involved in Cytoplasmic egress	114704	115906	+	400	45.605	UL88 [13]	CyUL88 (99.3%)	Rh123 (98.3%)	-
CyUL89_ex1	ORF Core Gene	Terminase Component (TER1) / ATPase subunit	116850	115903	-	315	35.829	Ul89 [27]	CyUL89 (100%)	UL89 DNA packaging Protein (100%)	RhUL89 (100%)
Cy128	-	-	116566	117135	+	189	20.879	-	Cy128 (96.3%)	Rh125 (93.2%)	Rh125 (92.7%)
CyUL91a	Betagamma	Essential for transcription of true late (γ2) genes[5]	116849	117040	+	63	7.289	UL91 [14]	-	-	RhORF6 (100%)
CyUL91b	ORF Betagamma	-	117167	117475	+	102	10.84	UL91 [35]	CyUL91 (98.1%)	Rh126 (93.2%)	RhUL91 (96.1%)
CyUL92	ORF Betagamma	-	117363	118073	+	236	26.494	UL92 [35]	CyUL92 (98.3%)	Rh127 (97.5%)	RhUL92 (99.2%)
CyUL93	ORF Core Gene	Tegument protein putatively involved in Capsid transport	118039	119604	+	521	59.752	UL17 [29]	CyUL93 (97.1%)	Rh128 (97.7%)	RhUL93 (97.7%)
Cy132	-	-	118200	118520	+	106	11.189	-	Cy132 (96.2%)	-	Rh128.1 (92.5%)
CyUL94	Core Gene	ssDNA binding cytoplasmic egress tegument protein	119480	120520	+	346	37.682	UL16 [9]	CyUL94 (98%)	-	RhUL94 (97.4%)
CyUL89_ex2	Core Gene	Terminase Component / ATPase subunit	121440	120517	-	307	35.795	Ul89 [35]	CyUL89 (99.3%)	UL89 (99.7%)	RhUL89 (100%)
CyUL95	Core Gene	Putative encapsulation chaperone protein	121439	122734	+	431	47.029	UL95 [8]	CyUL95 (99.2%)	Rh130 (99.2%)	RhUL95 (99%)
CyUL96	-	Tegument protein	122731	123120	+	129	14.796	UL14 [31]	CyUL96 (100%)	Rh131 (97.7%)	RhUL96 (96.9%)
CyUL97	Core Gene	Viral serine-threonine protein kinase tegument protein	123177	125003	+	608	67.997	UL97 [30]	CyUL97 (99.2%)	Rh132 (95.9%)	RhUL97 (96.4%)
Cy133	-	-	124136	123822	-	104	11.853	-	-	Rh133 (94.2%)	-
Cy134	-	Murid herpesvirus homology[6]	124487	124248	-	79	8.681	-	-	-	-
CyUL98	Core Gene	DNAse	125054	126724	+	556	63.381	UL98 [33]	CyUL98 (99.6%)	Rh134 (99.8%)	-
Cy138	-	-	125967	125689	-	92	10.188	-	Cy138 (100%)	Rh135 (98.9%)	-

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
Cy139	-	-	126454	126005	-	149	16.874	-	Cy139 (99.3%)	Rh136 (98.7%)	Rh136 (98%)
CyUL99	Core Gene	Myristoylated tegument phosphoprotein (pp28) involved in cytoplasmic egress	126661	127116	+	151	16.654	UL99 [33]	CyUL99 (92.7%)	Rh137 (92.8%)	-
CyUL100	Core Gene	Virion Envelope Glycoprotein M (gM) / Complexes with gN involved in entry	128353	127283	-	356	41.054	UL100 [37]	CyUL100 (99.2%)	Rh138 (98.6%)	-
CyUL102	Core Gene	DNA Helicase primase subunit	128542	130716	+	724	80.418	UL102 [36]	CyUL102 (99%)	Rh139 (98.2%)	RhUL10 (98.1%)
CyUL103	Core Genes	Tegument protein putatively involved in nuclear egress /envelope fusion protein	131493	130738	-	251	28.842	UL7 [31]	CyUL103 (98.4%)	Rh140 (97.2%)	-
CyUL104	Core Gene	Capsid Portal Protein[4]/DNA encapsulation	133390	131420	-	656	75.513	Ul104 [24]	CyUL104 (99.8%)	Rh141 (99.2%)	RhUL104 (99.1%)
CyUL105	Core Gene	DNA Helicase primase subunit	133227	135806	+	859	97.481	UL105 [34]	CyUL105 (99.4%)	Rh142 (99.5%)	-
Cy146	-	-	133376	133828	+	150	16.912	-	Cy146 (98%)	-	Rh142.1 (95%)
Cy147	-	-	136674	136805	+	43	5.151	-	Cy147 (100%)	-	Rh142.3 (97.6%)
Cy148_ex1	-	Latency-associated viral Interleukin 10[4]	139102	139299	+	65	7.05	-	Cy148_exon 1 (96.9%)	Interleukin-10-like protein precursor (93.8%)	RhUL111 (96.8%)
CyUL111.5A_ex2	-	Latency-associated viral Interleukin 10[4]	139187	139648	+	153	17.642	-	CyUL111.5A (96.1%)	Rh143 (93.5%)	RhUL111a (91%)
Cy148_ex3	-	Interleukin 10 like protein precursor[4]	140026	140115	+	29	3.312	-	Cy148_ex3 (100%)	Interleukin-10-like protein precursor (91.3%)	RhUL111a (91.3%)
CyUL112_ex1	-	Early Phosphoprotein (p50)	140520	141317	+	265	28.286	UL112 [24]	CyUL112 (94.3%)	Rh144/Rh145 (94.9%)	RhUL112 (90.6%)
CyUL112/113_ex2	-	Early phosphoprotein (p84)[4]/Transcriptional activation and DNA replication	141416	142300	+	294	30.422	UL112 [9]	CyUL112/113 (98.3%)	Rh144/Rh145 (94.9%)	RhUL112 (94.9%)
CyUL114	Core Gene	Uracil-DNA glycosylase	143158	142415	-	247	28.253	UL114 [34]	CyUL114 (99.6%)	Rh146 (99.2%)	-

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL115	Core Gene	Virion Envelope Glycoprotein L (gL)/Associated with GH and complexes with gO or US128-130-131 for entry	143897	143121	-	258	29.176	UL115 [42]	CyUL115 (99.6%)	Rh147 (98.4%)	RhUL115 (98.1%)
CyUL116	-	Putative membrane glycoprotein	144987	143908	-	359	38.048	UL116 [33]	CyUL116 (96.7%)	Rh148 (92%)	RhUL116 (83.6%)
Cy153	-	-	144512	145009	+	165	16.703	-	Cy153 (99%)	Rh147.1 (89.2%)	-
Cy154	-	-	145126	144641	-	161	20.015	-	Cy154 (97.5%)	Rh149 (93.5%)	Rh149 (91.9%)
CyUL117	-	-	146117	144969	-	382	42.552	UL117 [20]	CyUL117 (100%)	Rh150 (99.7%)	RhUL117 (99.5%)
Cy156	-	-	145953	146162	+	69	7.751	-	Cy156 (96.9%)	Rh149.1 (96.9%)	-
CyUL119_ex1	-	IgG Fc binding glycoprotein modulator of antibody activity/ virion glycoprotein	146743	146144	-	199	23.478	UL119 [20]	CyUL119 (93%)	Rh141 (92.5%)	RhUL119 (92.5%)
Cy157_ex2	-	-	147504	146788	-	238	24.188	-	Cy157_exon 2 (88.8%)	Rh152 (93.9%)	RhUL119 (78.7%)
Cy158	-	-	147296	146979	-	105	11.721	-	Cy158 (100%)	Rh153 (87.6%)	Rh153 (87.5%)
CyUL120	UL120	Putative membrane glycoprotein	148149	147553	-	198	22.577	UL120 [38]	CyUL120 (100%)	Rh154 (92.4%)	-
CyUL121	UL120	Putative membrane glycoprotein	148699	148151	-	182	21.046	UL121 [15]	CyUL121 (97.8%)	Rh155 (95.6%)	Rh155 (95.1%)
CyUL122	-	Immediate Early 2 Transactivator/DNA-binding repressor activity & cell cycle modulation	150452	148965	-	495	53.824	IL122 [46]	CyUL122 (92%)	Rh156 (94.3%)	RhUL122 (92.3%)
CyUL123	-	Major immediate early 1 co-transactivator/ enhancer of activation by IE2 & indirect effect on transcription machinery	152140	150953	-	395	44.25	IL123 [41]	CyUL123 (61.9%)	Rh156 (83.5%)	RhUL123 (83.8%)
Cy161_ex3	-	Immediate early protein[4]	152582	152298	-	94	10.526	-	Cy161_exon 3 (65.9%)	Rh156 (80.6%)	RhUL122 (80.6%)
Cy162	-	-	152697	152954	+	85	10.229	-	Cy162 (93.1%)	Rh156.1 (90.7%)	-
Cy163	-	-	152806	153255	+	149	15.47	-	Cy163 (91.4%)	Rh156.2 (94.7%)	Rh156.2 (95.3%)
Cy161_ex4	-	Immediate early protein[4]	153101	152991	-	36	4.201	-	Cy161_exon 4 (94.9%)	Rh156 (IE1) (87.2%)	RhUL123 (82.1%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
Cy165a	-	-	153246	153512	+	88	9.609	-	Cy165 (43.1%)	-	-
CyUL126	-	Putative Map Kinase interacting Kinase[6]	153942	153796	-	48	5.661	UL126 [8]	CyUL126 (97.9%)	-	-
Cy166	-	-	154358	154507	+	49	5.827	-	Cy166 (98%)	-	Rh157.1 (59.1%)
Cy165b	-	-	154811	154374	-	145	16.433	-	Cy165 (93.8%)	-	Rh157.1 (90.3%)
Cy165c	-	-	154386	154520	+	44	4.918	-	Cy165 (57.9%)	-	Rh157.1 (55.3%)
Cy167	-	-	154492	154695	+	67	7.168	-	Cy167 (94%)	-	Rh157.3 (55.1%)
Cy170	-	-	155050	154517	-	177	20.109	-	Cy170 (73.6%)	Rh157 (75%)	Rh157.3 (74.9%)
Cy168	-	-	154563	154796	+	77	8.476	-	Cy168 (90.9%)	-	Rh157.1 (56.3%)
Cy171	-	-	154735	155139	+	134	15.833	-	Cy171 (77.3%)	Rh157 (67.5%)	Rh157 (73.7%)
Cy165	-	-	154823	154927	+	34	3.748	-	Cy165 (66.7%)	-	
Cy172	-	-	154839	155084	+	81	9.239	-	Cy172 (75%)	-	Rh157.2 (69.3%)
CyUL128_ex3	-	Putative secreted CC chemokine/ Involved in endothelial and epithelial tropism/ complexes with gH & gL	156220	155894	-	108	12.075	-	CyUL128 (100%)	-	RhUL128 (97.2%)
CyUL128_ex2	-	Putative secreted CC chemokine/ Involved in endothelial and epithelial tropism/ complexes with gH & gL	156519	156430	-	29	3.372	-	-	-	RhUL128 (93.1%)
CyUL128_ex1	-	Putative secreted CC chemokine/ Involved in endothelial and epithelial tropism/ complexes with gH & gL	156798	156634	-	54	6.287	UL128 [17]	CyUL128 (94.4%)	-	RhUL128 (93.6%)
Cy174	-	-	157495	156800	-	231	25.044	-	Cy174 (86.4%)	-	Rh157.4 (77%)
CyUL130	-	Putative Secreted protein involved in endothelial and epithelial tropism/ Complexes with gH & gL	157909	157616	-	97	11.524	UL130 [49]	CyUL130 (95.9%)	-	RhUL130 (85.6%)
CyUL131A	-	Putative Secreted protein involved in endothelial and epithelial tropism/ Complexes with gH & gL	158506	158249	-	85	9.717	UL131A [18]	CyUL131A (100%)	Rh157.6 (96.5%)	RhUL131a (96.5%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyUL132	-	Envelope Glycoprotein /post-entry temperance factor in fibroblasts	159215	158550	-	221	24.275	UL132 [38]	CyUL132 (96.6%)	Rh160 (97.4%)	RhUL132 (96.6%)
CyUL148	-	Putative membrane glycoprotein	160261	159281	-	326	36.884	UL148 [49]	CyUL148 (96.9%)	Rh159 (90.2%)	RhUL148 (87.5%)
CyUL147A	-	Putative membrane protein	160469	160272	-	65	6.957	-	-	-	-
CyUL147	CXCL	CXCL5 chemokine like secreted glycoprotein[6]	160927	160466	-	153	17.559	UL147 [47]	CyUL147 (89.6%)	UL147 (96.7%)	UL147 (98.7%)
CyUL146	CXCL	CXCL5 chemokine like secreted glycoprotein[6]	161322	160978	-	114	12.922	-	CyUL146 (77.9%)	-	RhUL146 (98.2%)
Cy181	-	Putative CXC chemokine like protein[4]	161762	161568	-	64	7.234	-	Cy181 (90.6%)	-	RhUL146a (98.4%)
Cy182	-	Putative CXC chemokine like protein[4]	162211	161885	-	108	12.033	-	Cy182 (93.5%)	-	RhUL146b (95.4%)
Cy183	-	Putative CXC chemokine like protein[4]	162666	162334	-	110	12.498	-	Cy183 (99.1%)	Rh161 (35.2%)	Rh161.1 (93.6%)
Cy184	-	Putative CXC chemokine like protein[4]	163165	162749	-	138	15.767	-	Cy184 (97.8%)	Rh161 (97.3%)	Rh161.2 (97.1%)
CyUL145	RL1	-	163656	163351	-	101	11.259	UL145 [12]	CyUL145 (100%)	Rh162 (99%)	RhUL145 (98%)
CyUL144	-	TNF receptor homologue membrane protein	164590	164075	-	171	18.686	-	CyUL144 (98.8%)	Rh163 (98.8%)	RhUL144 (98.8%)
CyUL141	UL14	NK cell inhibiting membrane glycoprotein	166101	164806	-	431	48.822	UL141 [34]	CyUL141 (97.2%)	Rh164 (97%)	RhUL141 (96.8%)
CyO11	-	-	166412	166185	-	75	8.539	-	-	Rh164.1 (100%)	-
Cy188	-	-	167041	166595	-	148	17.239	-	Cy188 (98%)	Rh165 (94.6%)	Rh165 (94.6%)
CyO13 (Cy189)	-	-	167615	167088	-	175	19.479	-	Cy189 (93.1%)	Rh166 (94.9%)	Rh166 (94.3%)
CyO14 (Cy190)	-	-	168249	167746	-	167	18.177	-	Cy190 (97.6%)	Rh167 (96.4%)	RhUL148 (83.7%)
CyO15 (Cy191)	-	-	169162	168503	-	219	24651	-	Cy191 (97.3%)	Rh168 (95.9%)	Rh168 (95.4%)
Cy192	-	-	169259	169690	+	143	16.256	-	Cy192 (86.7%)	-	Rh168.1 (83.9%)
CyO16 (Cy193)	O16	-	169828	169265	-	187	20.667	-	Cy193 (94.7%)	Rh169 (92.5%)	Rh169 (91.4%)
CyO17 (Cy194)	O16	-	170528	169962	-	188	21.24	-	Cy194 (99.5%)	Rh170 (96.3%)	Rh170 (96.8%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
CyO18 (Cy195)	-	-	171370	170531	-	279	30.633	-	Cy195 (96.1%)	Rh171 (92.5%)	Rh171 (92.1%)
CyO19 (Cy196)	-	-	172101	171571	-	176	19.975	-	Cy196 (98.3%)	Rh172 (95.5%)	Rh172 (96.6%)
Cy197	-	-	171826	171978	+	50	5.352	-	Cy197 (96%)	-	Rh171.1 (92%)
CyUL153	RL11	Putative membrane protein[4]	173280	172153	-	375	40.855	RL13 [19]	CyUL153 (96.5%)	Rh173 (57.9%)	Rh173 (90.9%)
CyO20 (Cy199)	-	-	175499	174417	-	360	40.025	-	Cy199 (92.5%)	Rh174 (93.6%)	Rh174 (93.6%)
Cy200	-	-	176196	176651	+	151	16.309	-	Cy200 (96.7%)	Rh175 (94%)	Rh175 (94.7%)
CyO21 (Cy201)	-	-	176992	176345	-	215	23.507	-	Cy201 (93.1%)	Rh176 (93%)	Rh176 (94.9%)
Cy202	-	-	177425	176958	-	155	17.491	-	Cy202 (93%)	Rh177 (79.9%)	Rh177 (79.2%)
CyO22 (Cy203)	-	-	177819	177052	-	255	28.114	-	Cy203 (91.8%)	Rh178 (87.5%)	Rh178 (86.4%)
Cy204	-	-	178111	177896	-	71	7.586	-	Cy204 (87.7%)	-	Rh178.2 (88.9%)
Cy205	-	-	178044	178142	+	32	3.415	-	Cy205 (87.9%)	-	Rh178.1 (76.7%)
Cy206	-	-	178231	178734	+	167	17.956	-	Cy206 (91.6%)	-	Rh178.3 (70.8%)
Cy207	-	-	178371	178694	+	107	11.841	-	Cy207 (87.9%)	-	Rh178.3 (84.6%)
Cy209	-	-	179424	178903	-	173	17.196	-	Cy209 (92.6%)	Rh180 (91.3%)	Rh180 (94.2%)
O23 (Cy208)	-	-	178994	179509	+	171	18.631	-	Cy208 (91.8%)	Rh179 (93%)	Rh179 (94.2%)
CyUS1	US1	-	180039	179533	-	168	19.267	US1 [37]	CyUS1 (100%)	Rh181 (97.6%)	RhUS1 (98.8%)
Cy211	-	-	179888	180121	+	77	8.31	-	Cy211 (97.4%)		Rh180.1 (96.1%)
CyUS2	US2	Membrane glycoprotein/MHC processing and transport inhibition	180864	180274	-	196	23.139	US2 [9]	CyUS2 (92.9%)	Rh182 (77.9%)	Rh182 (76.9%)
CyUS5 (Cy214)	-	-	181370	181735	+	121	13.682	-	Cy214 (89.7%)	Rh183 (86%)	Rh183 (86.8%)
CyUS3	US2	Immediate early membrane glycoprotein/MHC processing and transport inhibition	182062	181397	-	221	25.397	-	CyUS3 (94%)	Rh184 (92.6%)	Rh184 (92.6%)
Cy215	-	-	182860	182561	-	99	11.146	-	Cy215 (97%)	-	Rh184.1 (94.9%)
CyUS6 (Cy216)	US6	Putatively inhibits TAP peptide translocation	183514	182942	-	190	21.417	-	Cy216 (96.8%)	Rh185 (95.9%)	Rh185 (95.3%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
Cy217	-	-	184413	183724	-	229	27.241	-	Cy217 (83.1%)	Rh186 (83.2%)	Rh186 (84.1%)
CyUS11a	US6	MHC I degrading membrane glycoprotein	185230	184658	-	190	21.79	-	CyUS11 (92.6%)	Rh187 (94.7%)	Rh187 (94.7%)
CyUS11b	US6	MHC I degrading membrane glycoprotein	185343	185113	-	76	8.399	-	CyUS11 (86.8%)	Rh187 (86.8%)	Rh187 (78.9%)
Cy219	-	-	185813	185439	-	124	14.579	-	Cy219 (92.7%)	Rh188 (98.4%)	-
CyUS11c	US6	MHC I degrading membrane glycoprotein	187041	186100	-	313	36.658	US11 [34]	CyUS11 (91.8%)	Rh189 (90.8%)	RhUS11 (90.8%)
Cy223	-	-	187938	187156	-	260	30.051	-	Cy223 (93.6%)	Rh191 (85.5%)	-
CyUS12	US12	Putative multiple transmembrane protein	187855	187313	-	180	19.783	US12 [34]	CyUS12 (97.3%)	Rh190 (98.1%)	-
CyUS13	US12	Putative multiple transmembrane protein	188760	187996	-	254	29.726	US13 [36]	CyUS13 (99.6%)	Rh192 (100%)	RhUS13 (99.6%)
Cy224	-	-	189251	188766	-	161	18.461	-	-	Rh193 (93.2%)	Rh193 (92.5%)
CyUS14a	US12	Putative multiple transmembrane protein	189706	188873	-	277	31.395	US14 [24]	CyUS14 (98.2%)	Rh194 (98.6%)	Rh194 (98.9%)
CyUS14b	US12	Putative multiple transmembrane protein	190565	189837	-	242	27.324	US14 [31]	CyUS14 (99.2%)	Rh195 (97.9%)	RhUS14 (26.8%)
CyUS14c	US12	Putative multiple transmembrane protein	191416	190658	-	252	29.503	US14 [9]	CyUS14 (99.6%)	Rh196 (98.8%)	RhUS14 (99.2%)
Cy228	-	-	192247	191522	-	241	27.85	-	Cy228 (96.3%)	Rh197 (95.9%)	Rh197 (95.9%)
Cy229	-	-	192116	191802	-	104	12.443	-	Cy229 (92.1%)	-	Rh196.1 (91.3%)
CyUS17	US12	Putative multiple transmembrane protein	193049	192225	-	274	30.415	US17 [45]	CyUS17 (98.5%)	Rh198 (97.8%)	RhUS17 (97.8%)
CyUS18	US12	Putative multiple transmembrane protein	193955	193155	-	266	30.015	US18 [31]	CyUS18 (98.9%)	Rh192 (23.9%)	RhUS18 (98.5%)
CyUS19	US12	Temperance factor and putative multiple transmembrane protein	194860	194075	-	261	30.0734	US19 [24]	CyUS19 (96.2%)	Rh200 (96.2%)	Rh200 (95.8%)
CyUS20	US12	Putative multiple transmembrane protein	195682	194921	-	253	28.578	US20 [34]	CyUS20 (99.2%)	Rh201 (99.2%)	RhUS13 (27%)
CyUS21	US12	Membrane protein[6]	196416	195730	-	228	26.086	US21 [28]	CyUS21 (100%)	Rh202 (98.2%)	-
CyUS22	US22	Tegument protein	198262	196538	-	574	65.879	US22 [9]	CyUS22 (98.4%)	Rh203 (97.6%)	RhUS22 (97.4%)
CyUS23	US22	Tegument protein	200292	198421	-	623	72.542	US23 [34]	CyUS23 (97.6%)	Rh204 (95.4%)	RhUS23 (95.7%)
Cy237	-	-	199133	198720	-	137	16.402	-	Cy237 (81.2%)	Rh206 (89%)	Rh206 (89.8%)

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
Cy238	-	-	198925	199248	+	107	12.684	-	Cy238 (94.2%)	Rh205 (88.4%)	Rh205 (88.4%)
Cy239	-	-	200077	200394	+	105	11.668	-	Cy239 (95.2%)	Rh208 (90.5%)	Rh208 (91.4%)
CyUS24	US22	Tegument protein	201746	200316	-	476	56.539	US24 [32]	CyUS24 (99.6%)	Rh209 (99.2%)	RhUS23 (27.1%)
Cy241	-	-	201153	201338	+	61	6.852	-	Cy241 (91.7%)	Rh210 (85.2%)	-
US26	US22	-	203902	202109	-	597	67.842	US26 [9]	CyUS26 (98.2%)	Rh211 (99.2%)	RhUS26 (99%)
Cy243	-	-	202878	202564	-	104	11.597	-	Cy243 (96.2%)	Rh212 (96.2%)	-
Cy244	-	-	202655	203170	+	171	19.018	-	Cy244 (95.3%)	Rh213 (94.7%)	-
CyUS28a	GPCR	CC and CXC chemokine receptor like virion envelope glycoprotein	204074	205060	+	328	37.497	US28 [42]	CyUS28 (99.4%)	Rh214 (98.5%)	Rh214 (98.8%)
CyUS28b	GPCR	CC and CXC chemokine receptor like virion envelope glycoprotein	205395	206408	+	337	38.662	US28 [16]	CyUS28 (99.1%)	Rh215 (93.8%)	Rh218 (38.1%)
CyUS28c	GPCR	CC and CXC chemokine receptor like virion envelope glycoprotein	206543	207544	+	333	38.093	US28 [36]	CyUS28 (99.4%)	Rh218 (37.2%)	Rh218 (37.5%)
Cy248	-	-	207720	207439	-	93	10.528	-	Cy248 (96.7%)	Rh217 (89.1%)	Rh217 (91.3%)
US28d	GPCR	Membrane protein	207617	208639	+	340	39.229	US28 [48]	CyUS28 (98.2%)	Rh218 (96.5%)	Rh218 (97.4%)
Cy250	-	-	208714	208409	-	101	11.545	-	Cy250 (96%)	Rh219 (94.1%)	Rh219 (93.1%)
CyUS28e	GPCR	Membrane protein	208784	210235	+	483	53.429	US28 [9]	CyUS28 (85.6%)	Rh220 (98.6%)	RhUS28 (98.7%)
CyUS29	US29	Putative membrane glycoprotein	210396	211715	+	439	49.253	US29 [9]	CyUS29 (97.3%)	Rh221 (94.8%)	RhUS29 (94.6%)
Cy253	-	-	210862	211185	+	107	12.663	-	Cy253 (97.2%)	Rh222 (95.3%)	-
CyUS30	-	Putative membrane glycoprotein	211633	212454	+	273	30.767	US30 [9]	CyUS30 (98.2%)	Rh223 (96%)	Rh223 (96%)
Cy255	-	-	212933	212316	-	205	22.469	-	Cy255 (96.7%)	Rh224 (94.8%)	Rh224 (94.3%)
CyUS31	US1	-	212530	213015	+	161	18.547	US31 [8]	CyUS31 (97.5%)	Rh225 (93.2%)	RhUS31 (91.9%)
CyUS32	US1	-	213142	213702	+	186	22.205	US32 [28]	CyUS32 (100%)	Rh226 (96.8%)	RhUS32 (96.2%)
Cy259	-	-	213616	213422	-	64	7.203	-	Cy259 (100%)	Rh227 (93.7%)	-
CyO24 (Cy260)	-	-	213845	214150	+	101	10.809	-	Cy260 (98%)	Rh228 (93.1%)	Rh228 (88.3%)
CyO25	-	Putative secreted protein	214774	214394	-	126	14.167	-	-	Rh228.1 (89.7%)	-
Cy261	-	Putative membrane protein	215251	214793	-	152	16.754	-	Cy261 (92.1%)	Rh229 (85.5%)	Rh229 (84.9%)
CyTRS1	US22	Immediate early tegument	216947	214872	-	691	77.289	TRS1 [35]	CyTRS1 (96.7%)	Rh230 (92.5%)	RhTRS1

ORF	Gene Family	Putative Function[1]	Translation		Strand	Size (aa)	Putative Molecular Mass (kDa)	Nearest HCMV homologue[2]	Non-Human Primate Homologues		
			Start	Stop					CyCMV Ottawa	RhCMV 68-1	RhCMV 180.92
		protein									(92.8%)

Footnotes:

1 Functions annotated based on studies of HCMV [274] unless otherwise indicated.
2 Nearest homologous HCMV gene based on Bit score using BLASTP search. Strain of HCMV gene indicated with footnote.
3 Based on function of RhCMV68-1 homologue [288].
4 Based on function of CyCMV Ottawa homologue [289]
5 Based on studies of HCMV Towne-BAC [513]
6 Inferred through function of homologous proteins [274]
8 HCMV strain AD169
9 HCMV strain AF1
10 HCMV strain ASM72
11 HCMV strain C154
12 HCMV strain C425,
13 HCMV strain CINCY+Towne
14 HCMV strain Coz
15 HCMV strain David
16 HCMV strain FRCMV-14L
17 HCMV strain GSV6
18 HCMV strain GSV9
19 HCMV strain HAN1
20 HCMV strain HAN13
21 HCMV strain HAN16
22 HCMV strain HAN20
23 HCMV strain HAN3
24 HCMV strain HAN38
25 HCMV strain HKS40
26 HCMV strain I-10
27 HCMV strain IS17
28 HCMV strain JHC
29 HCMV strain JP
30 HCMV strain L2
31 HCMV strain Merlin
32 HCMV strain RK (Human Herpesvirus 7)
33 HCMV strain TB40/E
34 HCMV strain Toledo
35 HCMV strain Towne
36 HCMV strain TR
37 HCMV strain U11
38 HCMV strain U8
39 HCMV strain VR1814
40 HCMV strain 35
41 HCMV strain 66
42 HCMV strain 452
43 HCMV strain 553
44 HCMV strain 3157
45 HCMV strain 3301
46 HCMV strain 5234
47 HCMV strain 401058
48 HCMV strain 26M
49 HCMV strain 51C

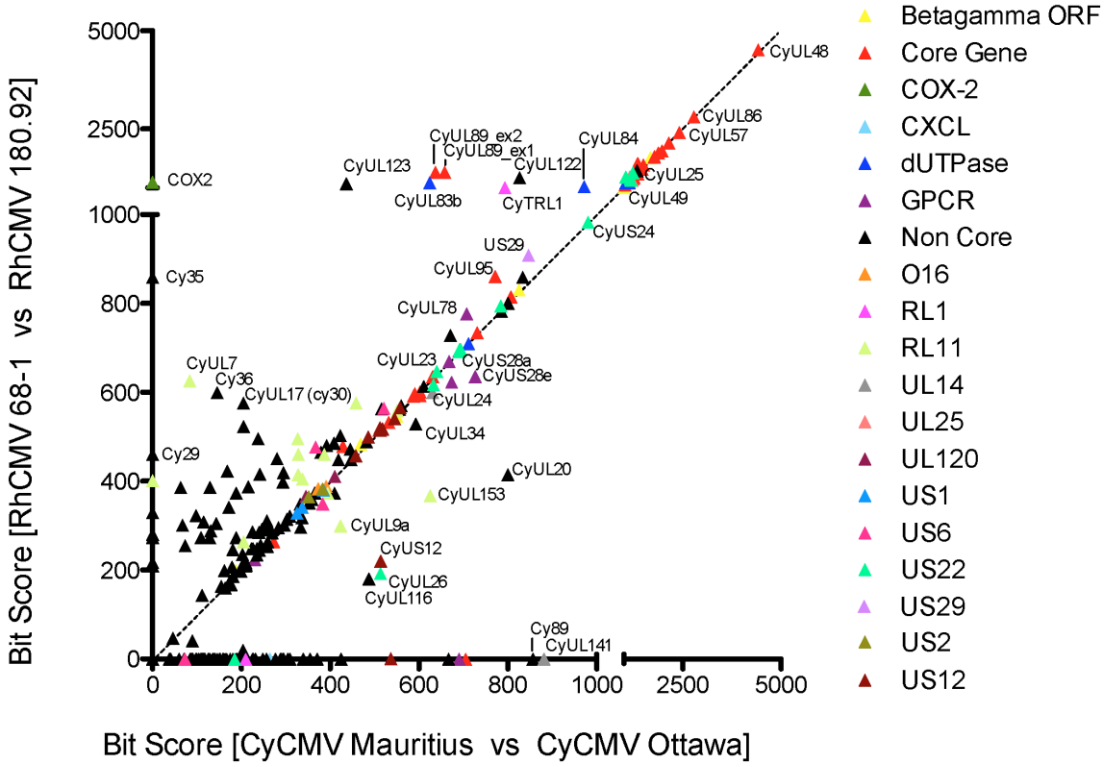


Figure 2-5: Two-dimensional bit-score plots between CMV strains. Graph represents comparisons of gene homologue bit-scores across CyCMV Mauritius versus CyCMV Ottawa and RhCMV 68-1 versus RhCMV 180.92. Bit-score among RhCMV homologues of the indicated gene are tracked on the y-axis while CyCMV homologues are tracked on the x-axis. A high bit-score indicated a high level of similarity between genes, while a bit-score of zero indicates one or both species in the comparison lacked a copy of the gene. Genes that fall on the diagonal are equally conserved in all comparisons, for example *CyUL48* (*rh78*) and *CyUL74* (*rh74*); genes that fall to one side of the diagonal are better conserved between viruses of one species than between viruses of the other species, for example COX-2 (*rh10*) which is highly conserved between RhCMV 180.92 and RhCMV 68-1, but absent from CyCMV Ottawa, and *CyUL94* (*rh129*). Genes are coloured according to known gene family. Select genes are annotated according to their CyCMV Mauritius names. Where genes belong to multiple families only one is indicated, for breakdown of chart by gene family see **Figure 2-6**, **Figure 2-7**, **Figure 2-8** and **Figure 2-9** for alternative pairwise comparisons.

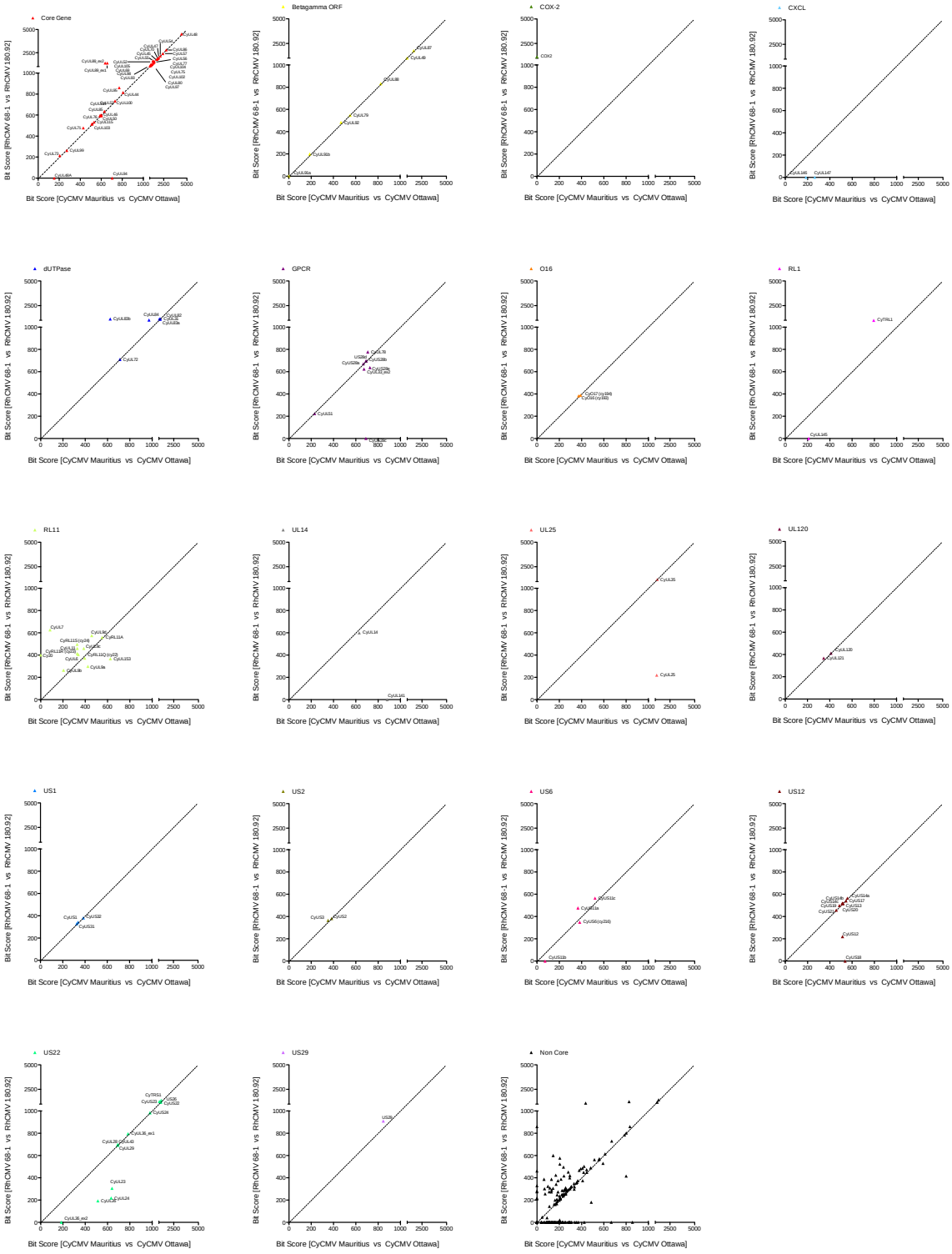


Figure 2-6: Break down of two-dimensional bit-score plots between CMV Mauritius-CyCMV Ottawa versus RhCMV 68-1-RhCMV 180.92 by ORF family. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to ORF family. CyCMV Mauritius names for ORF labels.

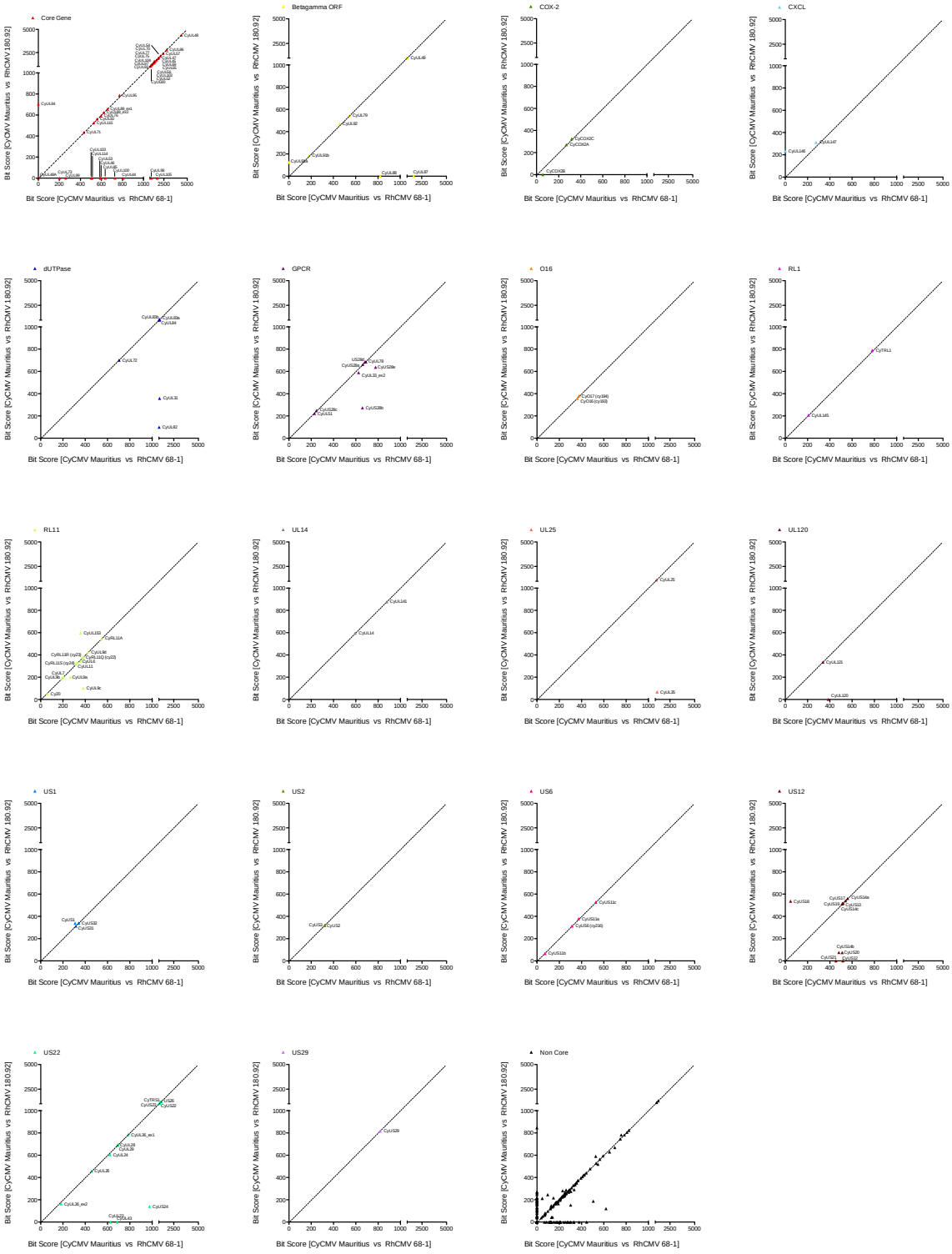


Figure 2-7: Break down of two-dimensional bit-score plots between CyCMV Mauritius-RhCMV 68–1 versus CyCMV Mauritius-RhCMV 180.92 by ORF family. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to ORF family. CyCMV Mauritius names for ORF labels.

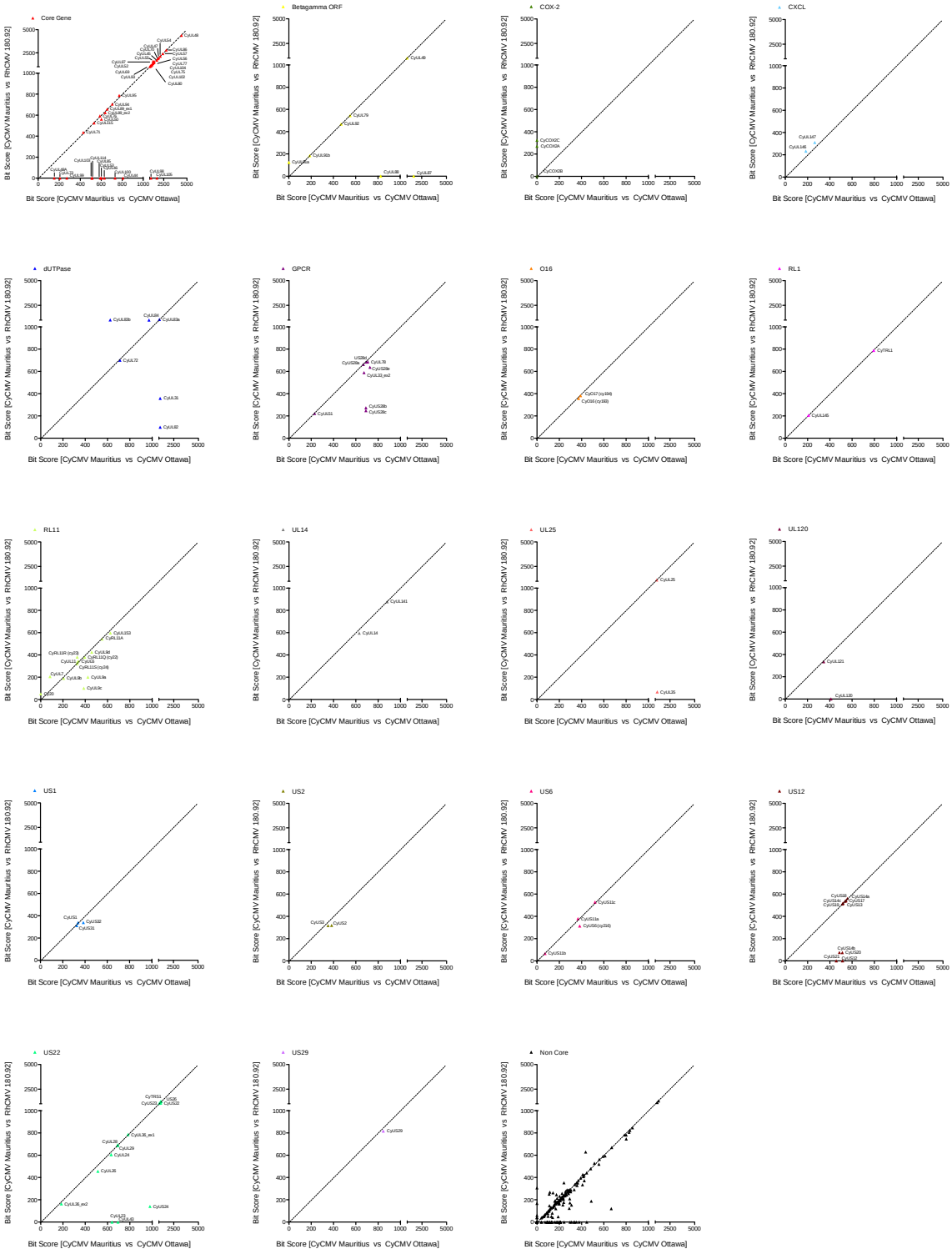


Figure 2-8: Break down of two-dimensional bit-score plots between CMV Mauritius-CyCMV Ottawa versus CyCMV Mauritius-RhCMV 180.92 by ORF family. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to ORF family. CyCMV Mauritius names used for ORF labels.

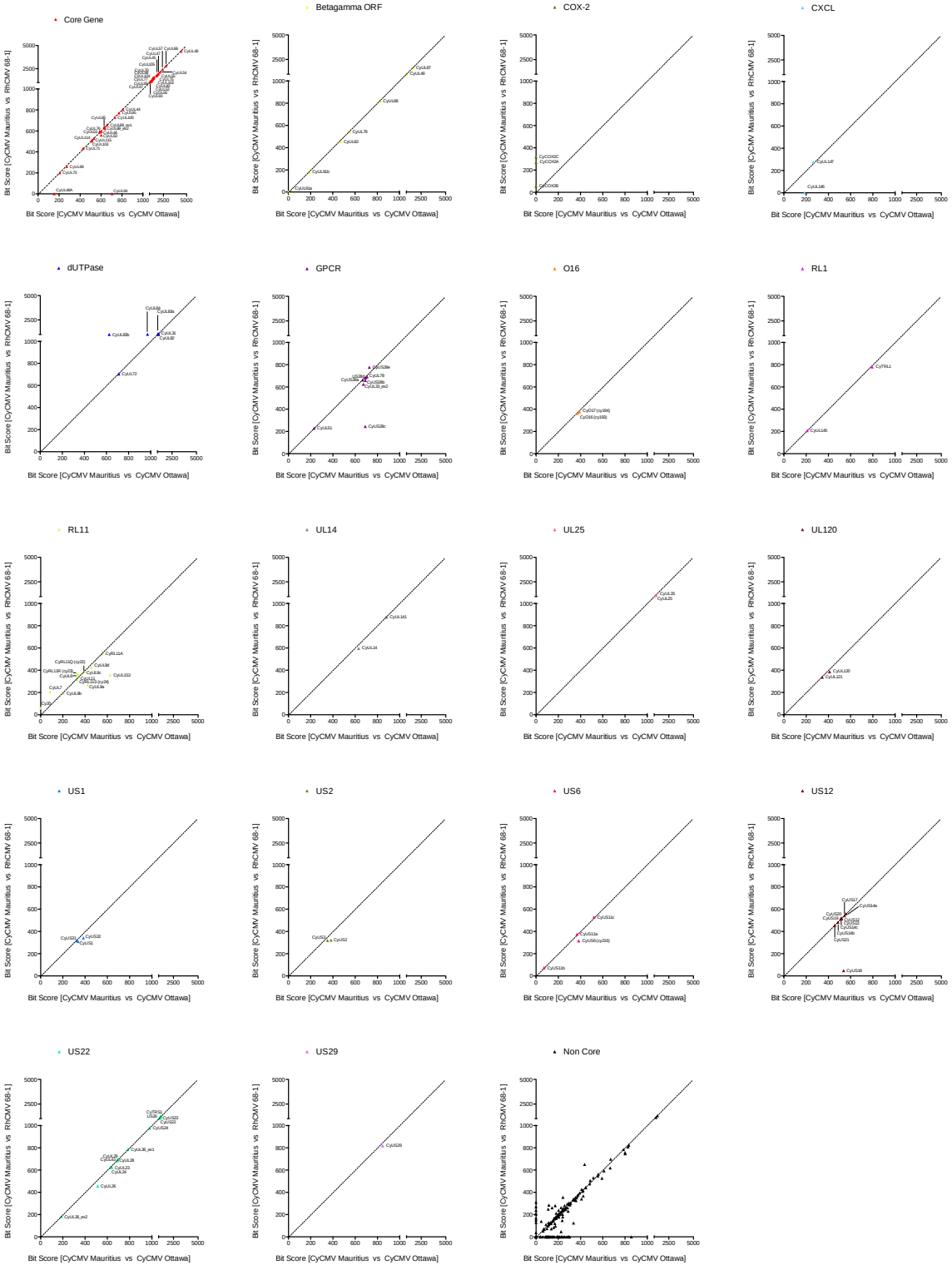


Figure 2-9: Break down of two-dimensional bit-score plots between CMV Mauritius-RhCMV 68–1 versus CyCMV Mauritius-CyCMV Ottawa by ORF family. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to ORF family. CyCMV Mauritius names used for ORF labels.

When grouped by gene function, capsid, DNA binding/nuclear, and tegument groupings contain highly conserved genes and less deviation in bit-score from the diagonal is observed (Fig**ure 2-10**; **Figure 2-11**, **Figure 2-12**, **Figure 2-13** and **Figure 2-14**). In each of these functional categories, genes deviate towards a higher bit-score in CyCMV Mauritius versus CyCMV Ottawa and CyCMV Mauritius versus RhCMV 68-1, and towards a lower bit-score in CyCMV Mauritius versus RhCMV 180.92. A similar trend for higher bit-scores in CyCMV Mauritius versus CyCMV Ottawa and CyCMV Mauritius versus RhCMV 68-1, and a lower bit-score in CyCMV Mauritius versus RhCMV 180.92 was observed with ORFs involved in viral temperance, though the average bit-score was lower, and was observed to a lesser absolute extent with immune regulatory, entry, and membrane associating ORF groupings.

Though closely related to CyCMV Ottawa, CyCMV Mauritius is less genetically diverged from both RhCMVs than CyCMV Ottawa is at several individual genes, including homologues of CMV genes *RL11B, UL123, UL83b, UL84* and a homologue of mammalian COX-2 (**Table 2-2**). This suggests a non-uniform rate of evolution of some genes (**Figure 2-2a**) and could reflect differences in selective pressure on the CMV strains in different hosts. The high average bit-score of capsid, DNA binding/nuclear, and tegument ORFs indicate a slow rate of evolution of these genes. Immune regulatory, viral entry, and membrane associated ORFs, in contrast, exhibit signs of more rapid evolution. This may be because the proteins these genes encode interact directly with the host immune system and may consequentially face greater directional or diversifying selection.

Macaque CMV genomes lack an internal repeat region between the UL and US regions. We explored the effect of this feature of genome architecture on the genomic location of genes using gene bit-scores. Our analyses indicate that there is more divergence between genes from different macaque CMV strains that are closer to the terminal regions of both the UL and US genome regions than between genes in different strains that are far from these genomic features (**Figure 2-15**). This decrease in gene conservation near the UL/US border was observed even in the absence of an internal terminal repeat region between the UL and US regions. Additionally, these terminal areas contain a higher proportion of genes that are strain specific. The center of the UL region contains a region with a high density of highly conserved genes. This is expected since in general, the rate of evolution of genes in a virus is slower in key genes, faster in less important genes, and fastest in non-coding and non-regulatory regions [514, 515], and because the UL region includes a cluster of core genes that are conserved across all herpesviruses.

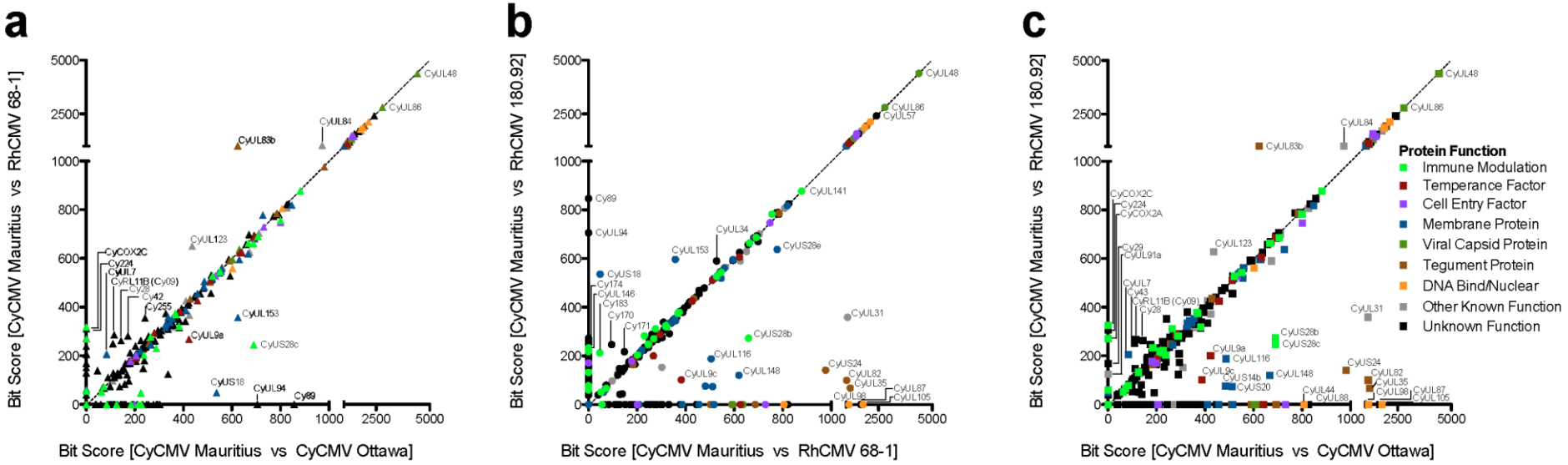


Figure 2-10: Two-dimensional bit-score plots between CMV strains. Graphs represent comparisons of gene homologue bit-scores across three strains simultaneously. Comparison of CyCMV Mauritius versus CyCMV Ottawa and CyCMV Mauritius versus RhCMV 68-1 (a), CyCMV Mauritius versus RhCMV 180.92 and CyCMV Mauritius versus RhCMV 68-1 (b), and CyCMV Mauritius versus RhCMV 180.92 and CyCMV Mauritius versus CyCMV Ottawa (c). Genes are coloured according to known function. Immune ORFs are involved in host immune regulation or evasion, entry ORFs are known cell entry factors, temperance ORFs are involved in temperance of viral growth, membrane ORFs are presented in cell membrane on infected cells or in virions, DNA/nuclear are known to interact with DNA or shuttle to the cell nucleus, capsid ORFs form the virus capsid, tegument proteins are found in the viral tegument, the function of ORFs with other known function vary, and the function is unknown at the time of writing for ORFs plotted as unknown. Individual ORF may have several functions; see **Figure 2-11**, **Figure 2-12**, **Figure 2-13** and **Figure 2-14** for visualization without overlap. Select genes are annotated according to their CyCMV Mauritius names.

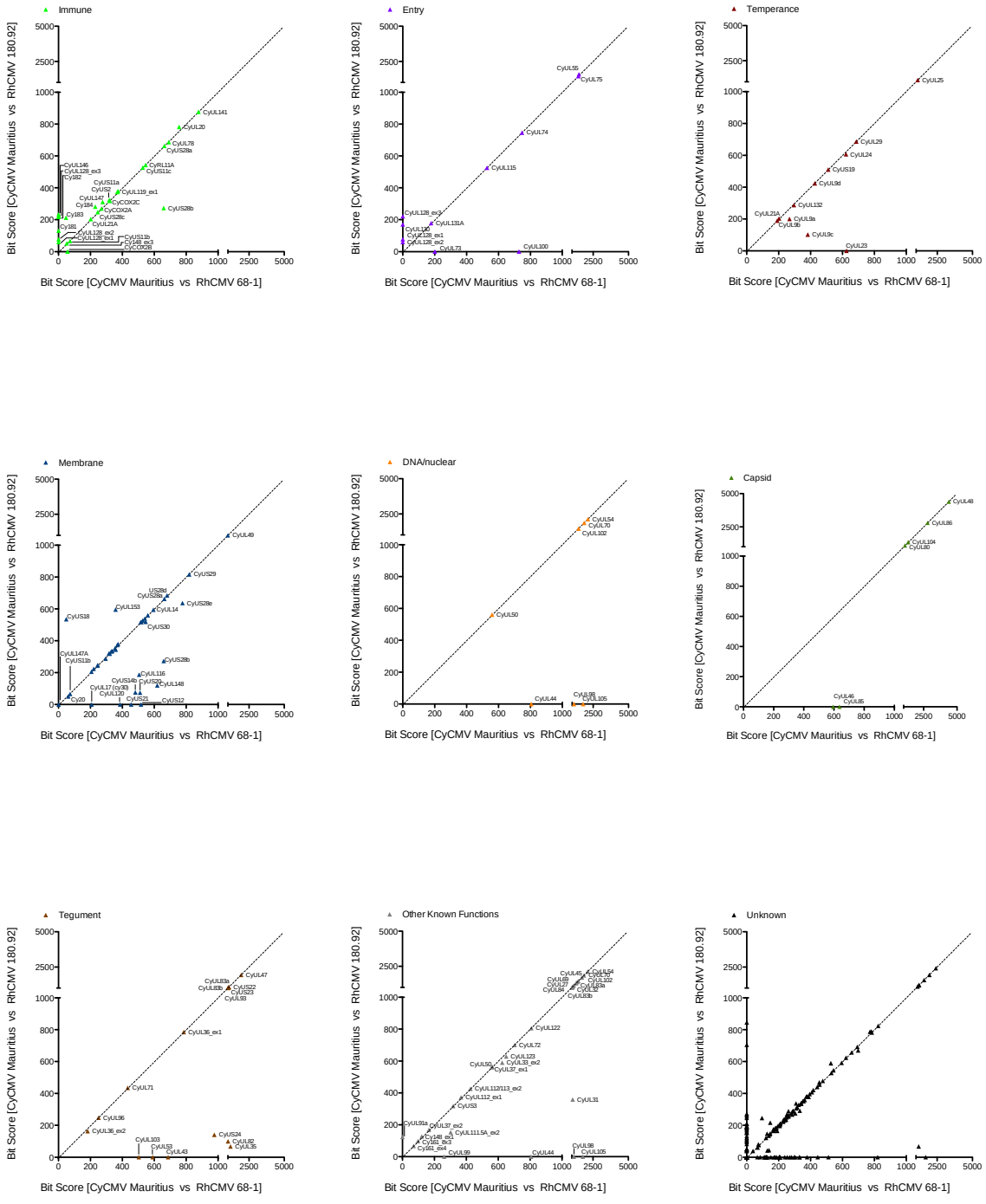


Figure 2-11: Break down of two-dimensional bit-score plots between CMV Mauritius-RhCMV 68-1 versus CyCMV Mauritius-RhCMV 180.92 by ORF function. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to known function. Individual ORF may have several functions and are present in multiple plots. Immune ORFs are involved in host immune regulation or evasion, entry ORFs are known cell entry factors, temperance ORFs are involved in temperance of viral growth, membrane ORFs are presented in cell membrane on infected cells or in virions, DNA/nuclear are known to interact with DNA or shuttle to the cell nucleus, capsid ORFs form the virus capsid, tegument proteins are found in the viral tegument, the function of ORFs with other known function vary, and the function is unknown at the time of writing for ORFs plotted as unknown. ORFs are annotated according to CyCMV Mauritius names except for ORFs of unknown function or membrane ORF where some ORF are left unlabelled.

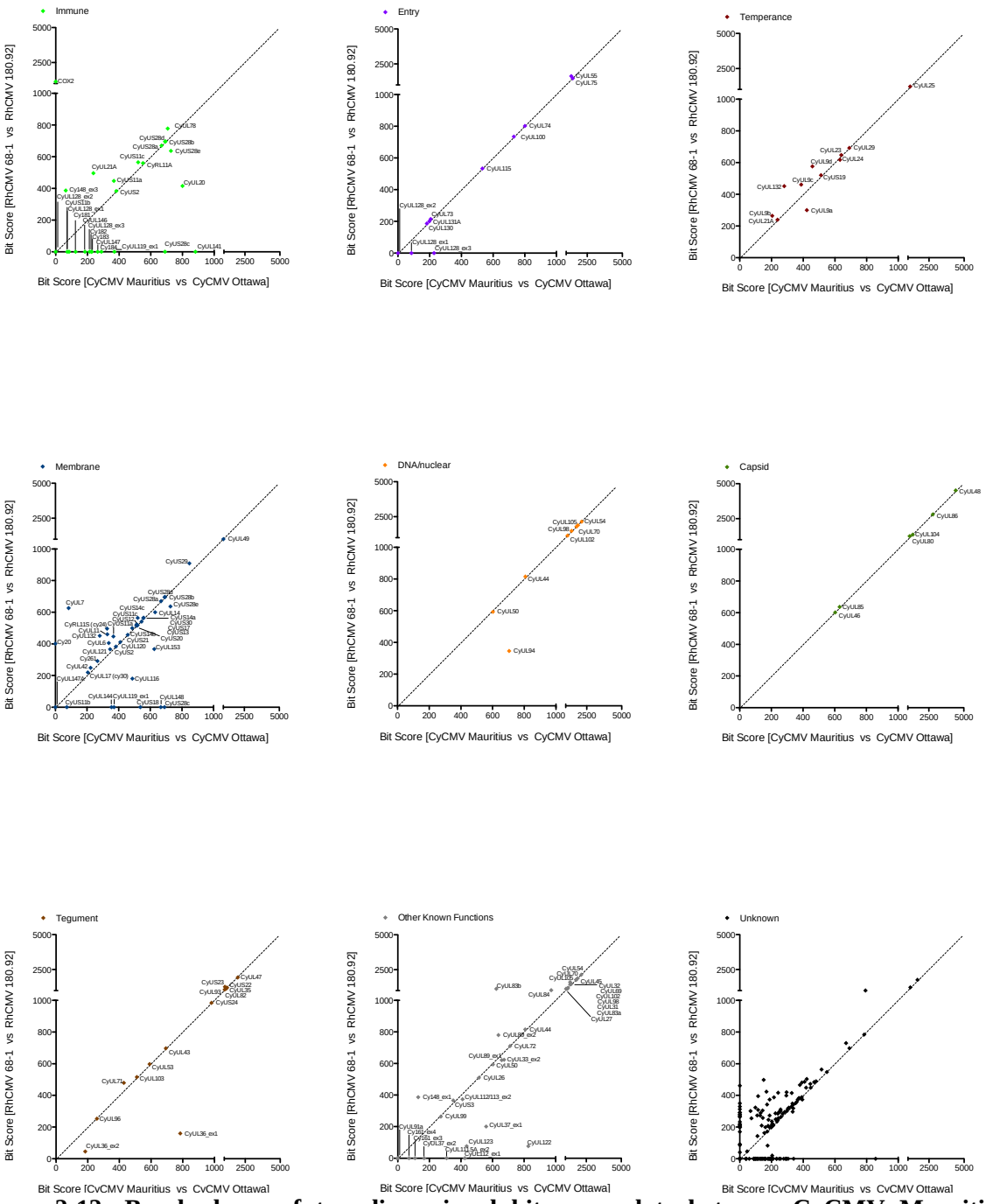


Figure 2-12: Break down of two-dimensional bit-score plots between CyCMV Mauritius-CyCMV Ottawa versus RhCMV 68-1-RhCMV 180.92 by ORF function. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to known function. Individual ORF may have several functions and are present in multiple plots. Immune ORFs are involved in host immune regulation or evasion, entry ORFs are known cell entry factors, temperance ORFs are involved in temperance of viral growth, membrane ORFs are presented in cell membrane on infected cells or in virions, DNA/nuclear are known to interact with DNA or shuttle to the cell nucleus, capsid ORFs form the virus capsid, tegument proteins are found in the viral tegument, the function of ORFs with other known function vary, and the function is unknown at the time of writing for ORFs plotted as unknown. ORFs are annotated according to CyCMV Mauritius names except for ORFs of unknown function or membrane ORF where some ORF are left unlabelled.

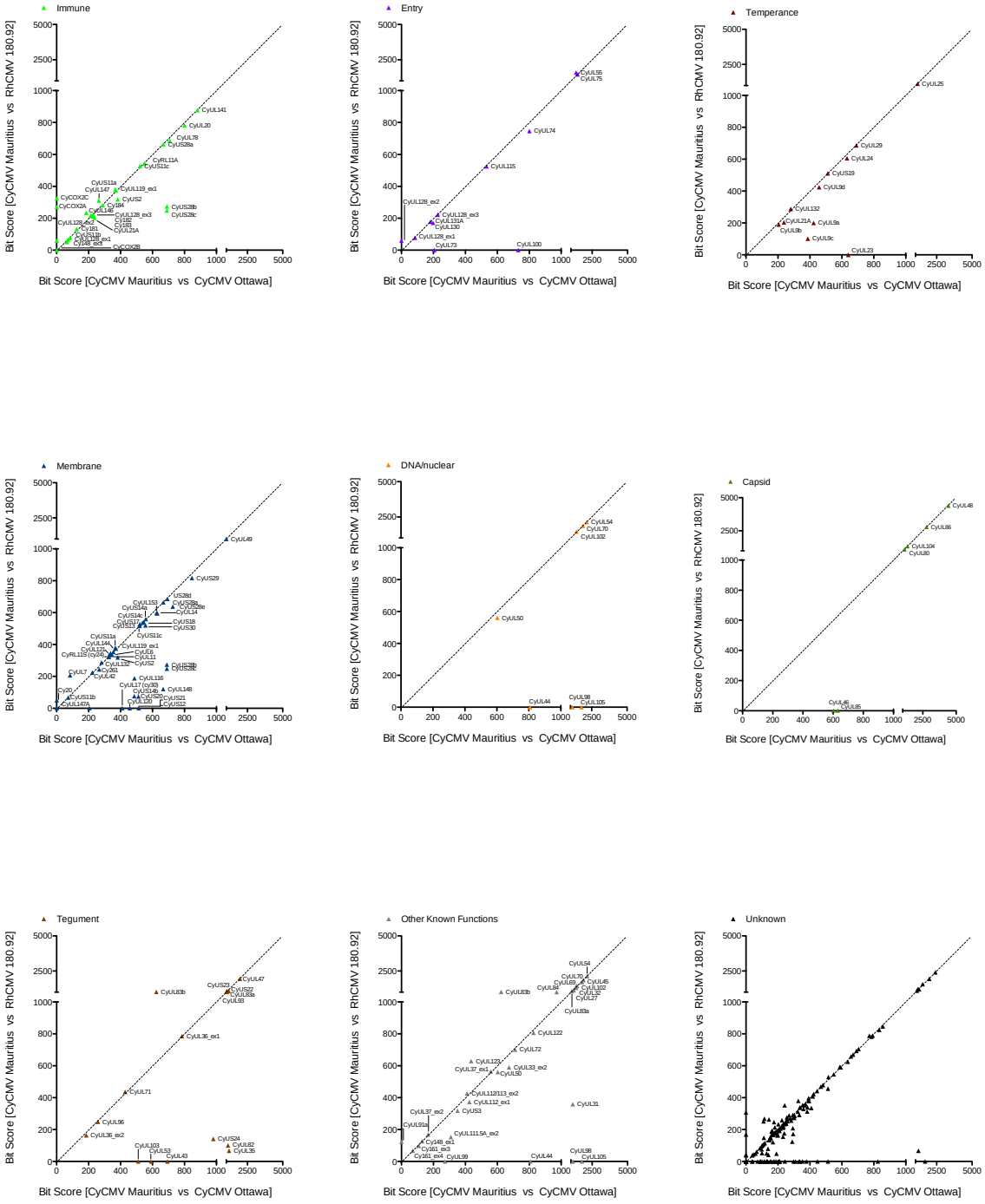


Figure 2-13: Break down of two-dimensional bit-score plots between CyCMV Mauritius-CyCMV Ottawa versus CyCMV Mauritius-RhCMV 180.92 by ORF function. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to known function. Individual ORF may have several functions and are present in multiple plots. Immune ORFs are involved in host immune regulation or evasion, entry ORFs are known cell entry factors, temperance ORFs are involved in temperance of viral growth, membrane ORFs are presented in cell membrane on infected cells or in virions, DNA/nuclear are known to interact with DNA or shuttle to the cell nucleus, capsid ORFs form the virus capsid, tegument proteins are found in the viral tegument, the function of ORFs with other known function vary, and the function is unknown at the time of writing for ORFs plotted as unknown. ORFs are annotated according to CyCMV Mauritius names except for ORFs of unknown function or membrane ORF where some ORF are left unlabelled.

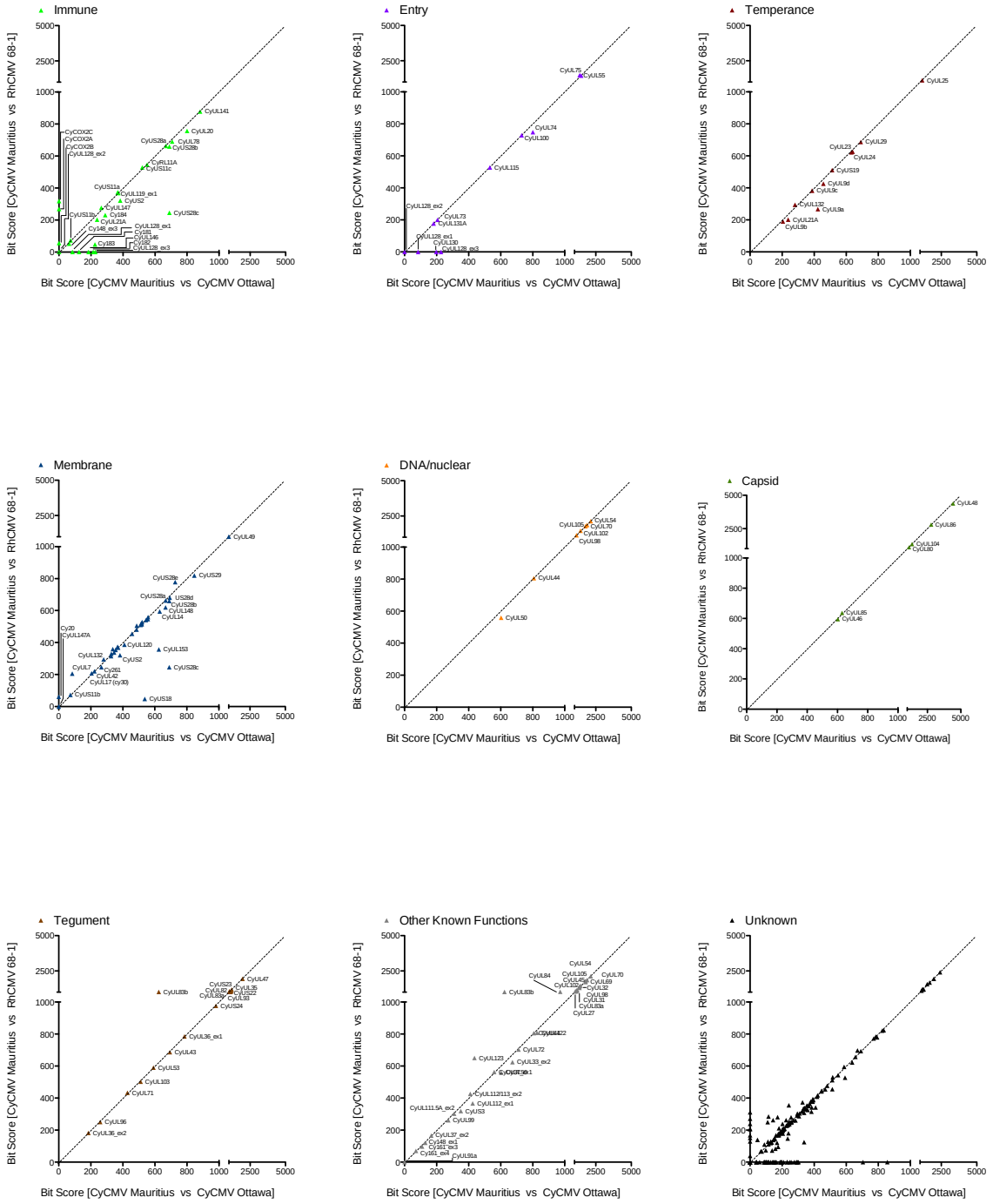
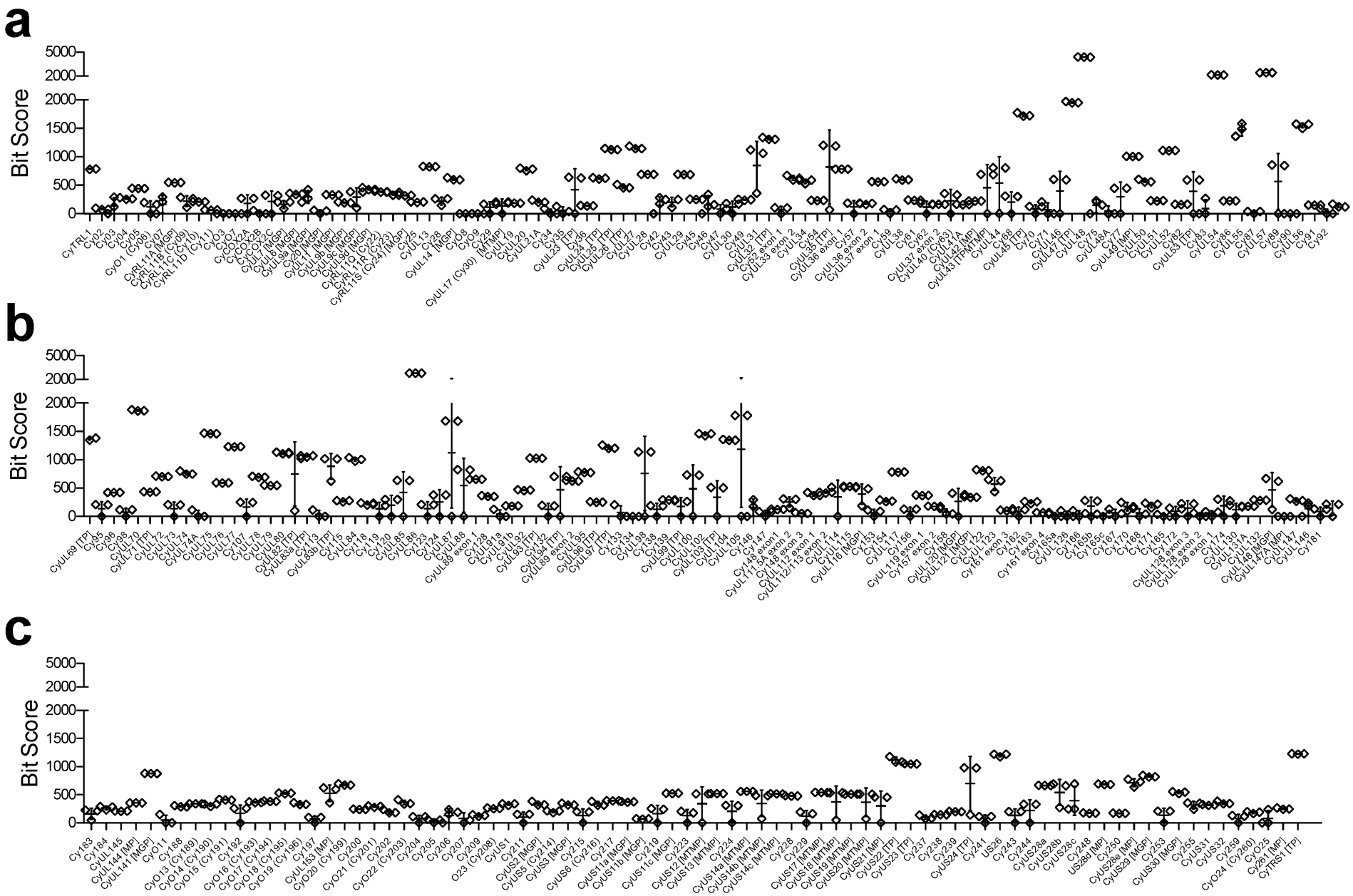


Figure 2-14: Break down of two-dimensional bit-score plots between CyCMV Mauritius-CyCMV Ottawa versus CyCMV Mauritius-RhCMV 68–1 by ORF function. Graphs represent comparisons of ORF homologue bit-scores across three strains simultaneously. ORFs are coloured according to known function. Individual ORF may have several functions and are present in multiple plots. Immune ORFs are involved in host immune regulation or evasion, entry ORFs are known cell entry factors, temperance ORFs are involved in temperance of viral growth, membrane ORFs are presented in cell membrane on infected cells or in virions, DNA/nuclear are known to interact with DNA or shuttle to the cell nucleus, capsid ORFs form the virus capsid, tegument proteins are found in the viral tegument, the function of ORFs with other known function vary, and the function is unknown at the time of writing for ORFs plotted as unknown. ORFs are annotated according to CyCMV Mauritius names except for ORFs of unknown function or membrane ORF where some ORF are left unlabelled.



Supplementary Figure 2-15: **Plot of bit-scores of ORFs across the CyCMV Mauritius genome.** ORFs of CyCMV Mauritius are ordered as seen in CyCMV Mauritius genome. Bit-score is plotted for each ORF between CyCMV Mauritius and CyCMV Ottawa, RhCMV 68-1, and RhCMV 180.92. Absence of the indicated gene from one or more compared genomes results in a bit-score of zero. Values are indicated as the means ± standard deviation for each gene.

During CMV replication, the genome circularizes and large genome segments may be reordered, reversed or lost [287]. This mechanism could explain why the *UL128-UL131* region is absent in RhCMV 68-1, but present in RhCMV 180.92 [516], and why *rh12-rh16* is absent from CyCMV Mauritius and CyCMV Ottawa but present in both RhCMVs. Loss of the *UL128-UL131* region is typical of attenuation (that is, the a decrease in virulence after passaging) in fibroblast cell lines, and demonstrates an inability to infect endothelial and epithelial cell lines. Thus our findings do not necessarily prove an absence of this region in the original wildtype RhCMV 68-1 [288, 290, 517]. How reversal and reordering of CMV genome segments [518] affect the evolution of CMV in wild macaque populations is yet to be fully understood.

2.5.4 Gene Content of CyCMV Mauritius is Distinguished from Other Strains

Several CMV genes found in CyCMV Ottawa, RhCMV 68-1 and RhCMV 180.92 are absent in CyCMV Mauritius. In particular, nine CyCMV Ottawa ORFs lack homologues in CyCMV Mauritius (**Table 2-3)** whereas 20 CyCMV Mauritius ORFs lack homologues in CyCMV Ottawa, and of the latter, 9 have homologues that are present in both RhCMV strains (**Table 2-2**). In all, 20 genes of RhCMV 68-1 and RhCMV 180.92 are not identified in CyCMV Mauritius (**Table 2-3).** Similarly, comparison to human CMV data, identified 84 ORFs present in strains of HCMV but absent in CyCMV Mauritius (**Table 2-4**). The majority of genes absent in all macaque CMV strains but present in HCMV strains have unknown functions though some, such as UL65 and UL108, are known to have effects on CMV growth kinetics (**Table 2-4**).

Table 2-3: RhCMV Genes absent from CyCMV Mauritius

Gene	Alternative Gene Name	Human Homologue	Necessary for growth in Towne[1]	Rh-Human Similarity (%)[2]	Function[3]
Rh9	-	-	-	-	-
Rh12	RL11F	-	-	-	-
Rh13	RL11G	-	-	-	-
Rh13.1	-	-	-	-	-
Rh14	-	-	-	-	Membrane Protein
Rh15	-	-	-	-	-
Rh16	-	-	-	-	-

Rh17	RL11H	UL11	-	-	Early Glycoprotein
Rh18	-	-	-	-	-
Rh19	RL11I	UL07	-	34%	Membrane Protein
Rh38	-	-	-	-	-
Rh45	-	-	-	-	-
Rh77	-	-	-	-	-
Rh94	-	-	-	-	-
Rh96	-	-	-	-	-
Rh121	-	-	-	-	-
Rh129	RhUL94	UL94	Essential	64%	Virion Protein
Rh142.2	-	-	-	-	-
Rh151.1	-	-	-	-	-
Rh158	RhUL147	UL147	Dispensable	-	Viral CXC Chemokine Homologue

Footnotes:
1 Based on study of gene mutation in HCMV Towne [519]
2 Similarity annotation based on comparisons of RhCMV 68-1 and HCMV in [288]
3 Function annotated based on studies of HCMV [274].

Table 2-4: Human Genes absent from CyCMV Mauritius

Gene	CyCMV Ottawa /RhCMV 68-1 Homologue	Function[1]	Effect of deletion on viral growth kinetics[2]	Family
RL2	-/-	-	no effect	-
RL3	-/-	-	no data	-
RL4	-/-	-	no effect	-
RL5	-/-	-	no data	RL11
RL6	-/-	-	no effect	RL11
RL7	-/-	-	no data	-
RL8	-/-	-	no data	-
RL9	-/-	-	no effect	-
RL10	-/-	-	no effect	RL11
RL11	cyRL11/rh05	IgG Fc-binding glycoprotein	no effect	RL11
RL12	-/-	Putative membrane glycoprotein	no effect	RL11
RL13	-/-	Putative membrane glycoprotein	no effect	-
RL14	-/-	-	no data	-
UL1	-/-	Putative membrane protein	no data	RL11
UL2	-/-	-	modest effect	RL11
UL3	-/-	-	no effect	-
UL4	-/-	Virion envelope glycoprotein	no effect	RL11
UL5	-/-	-	no effect	RL11

Gene	CyCMV Ottawa /RhCMV 68-1 Homologue	Function[1]	Effect of deletion on viral growth kinetics[2]	Family
UL8	-/-	Putative membrane glycoprotein	no effect	RL11
UL10	-/-	Temperance factor in retinal tissue[2]	no effect	RL11
UL12	-/-	-	modest effect	-
UL15	-/-	-	no effect	-
UL16	-/-	Membrane glycoprotein involved in inhibiting Natural Killer cell cytotoxicity	no effect	-
UL17	-/rh35	7-transmembrane glycoprotein	no effect	-
UL18	-/-	MHCI homologue putative membrane protein	no effect	UL18
UL22	-/-	-	no data	-
UL39	-/-	-	no effect	-
UL40	-/rh67	Membrane glycoprotein	no data	-
UL58	-/-	-	no data	-
UL59	-/-	-	no effect	-
UL60	-/-	-	required for replication	-
UL61	-/-	-	no data	-
UL62	-/-	-	no effect	-
UL63	-/-	-	no data	-
UL64	-/-	-	no effect	-
UL65	-/-	-	modest effect	-
UL66	-/-	-	no data	-
UL67	-/-	-	no effect	-
UL68	-/-	-	no data	-
UL80.5	-/rh109.1	Capsid assembly protein precursor	no data	Core
UL81	-/-	-	no data	-
UL90	-/-	-	required for replication	-
UL101	-/-	-	no data	-
UL106	-/-	-	no data	-
UL107	-/-	-	no data	-
UL108	-/-	-	modest effect	-
UL109	-/-	-	no effect	-
UL110	-/-	-	no effect	-
UL118	-/rh151	-	no data	-
UL124	-/rh156.2	Membrane glycoprotein / latent protein	variable critical effect	-
UL125	-/-	-	no data	-
UL127	-/-	-	no effect	-
UL129	-/-	-	modest effect	-
UL143	-/-	-	no data	-

Gene	CyCMV Ottawa /RhCMV 68-1 Homologue	Function[1]	Effect of deletion on viral growth kinetics[2]	Family
UL142	-/-	Involved in Inhibiting Natural Killer cell cytotoxicity	no data	-
UL140	-/-	Putative membrane protein	no data	-
UL139	-/-	Putative membrane glycoprotein	no data	-
UL138	-/-	Putative membrane protein	no data	-
UL137	-/-	-	no data	-
UL136	-/-	Putative membrane protein	no data	-
UL135	-/-	Putative secreted protein	no data	-
UL134	-/-	-	no data	-
UL133	-/-	Putative membrane protein	no data	-
UL148A	-/-	Putative membrane protein	no data	-
UL148B	-/-	Putative membrane protein	no data	-
UL148C	-/-	Putative membrane protein	no data	-
UL148D	-/-	Putative membrane protein	no data	-
UL149	-/-		no data	-
UL150	-/-	Putative Secreted Protein	no data	-
IRS1	-/-	Immediate early tegument protein & transcriptional activator / Inhibitor of protein kinase R-mediated repression of translation	no effect	US22
US4	-/-	-	no data	-
US5	-/rh183	-	no data	-
US6	-/rh185	Putative membrane glycoprotein / Inhibits TAP mediated endoplasmic reticulum peptide transport	no effect	US6
US7	-/-	Membrane glycoprotein	no effect	US6
US8	/-rh187	MHCI binding Membrane glycoprotein	no effect	US6
US9	-/-	Membrane glycoprotein involved in cell to cell transmission	no effect	US6
US10	-/-	Membrane glycoprotein / Delays MHCI trafficking	no effect	US6
US15	-/-	Putative multiple transmembrane protein	no effect	US12
US16	-/-	Temperance factor / Putative multiple transmembrane protein	no effect	US12
US25	-/-	-	no effect	-
US27	-/-	Virion envelope glycoprotein/Potentiates CXCR4 receptor[3]	no effect	GPCR
US33	-/-	-	no effect	-
US34	-/-	Putative secreted protein	no effect	-
US34A	-/-	Putative membrane protein	no data	-

Footnotes:

1 Function annotated based on studies of HCMV [274] unless otherwise indicated.

2 Based on study of gene mutation in HCMV Towne [519]

3 Based on study of US27 in HCMV AD169 [520]

2.5.5 Anomalous Gene Trees Reveal Patterns of Gene Family Evolution

Analysis of individual genes reveals several interesting features of these viral genomes. The Bayesian phylogeny for the important viral surface *gB* gene (*UL55*), for example, indicates an atypically high level of divergence among these strains as compared to the whole genomes (**Figure 2-2b**). This is of interest since it is a target of immune responses. In the analysis of the *UL55* gene, CyCMV Mauritius and CyCMV Ottawa do not cluster together, and CyCMV Mauritius instead is inferred to be more closely related to the RhCMVs than to CyCMV Ottawa. CyCMV Mauritius *UL55* is diverged from CyCMV Ottawa by 0.408 sps but diverged from RhCMV 68-1 and RhCMV 180.92 by only 0.175 and 0.188 sps respectively. **Figure 2-2c** shows a phylogenic estimate for the *US11* gene, which encodes an MHC down regulatory protein essential for superinfection [374]. Included in the analysis are multiple gene homologs of US11 generated by gene duplication in various CMVs. Inferred evolutionary relationships within US11 paralogs again indicate a closer evolutionary relationship between CyCMV Mauritius and the RhCMVs than to CyCMV Ottawa, although with weak statistical support (87.7% posterior probability). Interestingly, rh188 is not present in RhCMV 180.92.

Figure 2-2d shows results of a phylogenetic analysis of the multiple copies of *US28*, a macaque CMV capsid protein that induces COX-2 in target cells upon entry. CyCMV Mauritius genes were named following the synteny-based system established for CyCMV Ottawa [289]. The five copies of CyCMV Ottawa, CyCMV Mauritius and RhCMV 180.92, as well as the four copies and single pseudogene of RhCMV 68-1, cluster first by synteny and then by species. This likely indicates that divergence of all five copies occurred before speciation and that loss of one copy in RhCMV 68-1 occurred after speciation. Unlike the other *US28* genes, however, phylogenetic sequence analysis of *US28e* (*rhUS28*) in RhCMV180.92, rh220 in RhCMV 68-1, *CyUS28e* (in both CyCMVs) suggests that CyCMV Mauritius and CyCMV Ottawa are each more closely related to RhCMV *US28e* than to CyCMV *US28e*. CyCMV Mauritius is diverged from CyCMV Ottawa by 0.175 sps and diverged from RhCMV 180.92 and RhCMV

68-1 by 0.031 sps and 0.031 sps, respectively. CyCMV Ottawa is diverged from RhCMV 180.92 and RhCMV 68-1 by 0.168 sps and 0.169 sps, respectively.

Phylogenetic relationships were further estimated for several CMV genes known to have homology to mammalian host genes (**Figure 2-16**). Strain specific absence of several human and rhesus homologues supports independent evolution of each strain. This is evidenced, for example, by the absence of the *COX-2* gene in CyCMV Ottawa. Similarly, a COX-2 homologue (*rh10, CyCOX2*, this study), which promotes the formation of arachidonic acid in infected cells [298], is present in CyCMV Mauritius, RhCMV 180.92 and RhCMV 68-1, but absent in CyCMV Ottawa [298]. Viral COX-2 (rh10) is necessary for the infection of endothelial cells [298] and it has been shown that viral COX-2, but not cellular COX-2, protein is expressed when RhCMV 68-1 infects cells [298]. Since CMV has been implicated in vascular inflammation [521-525], the study of *rh10*, which confers cellular tropism and its selection in macaques, is of particular interest. US28 is a viral capsid protein capable of inducing host COX-2 expression in target cells [387]. CyCMV Mauritius, CyCMV Ottawa and RhCMV 180.92 have five copies each while RhCMV 68-1 has only four copies of *US28* genes and a *US28* pseudogene. The pseudogene in RhCMV 68-1 shows that it, similar to other macaque CMVs, had five copies of *US28*, and that the fifth copy was likely pseudogenized after speciation. The number of copies of *US28* is variable across mammalian CMVs. Baboon CMV, similar to macaque CMV, has four distinct copies, HCMV has a single copy and both RCMV and MCMV lack *US28* homologues [526, 527]. These multiple copies are more divergent from each other within a genome than they are from corresponding homologues between macaque species (**Figure 2-2d)** and can be assumed to have emerged prior to the speciation of these macaques.

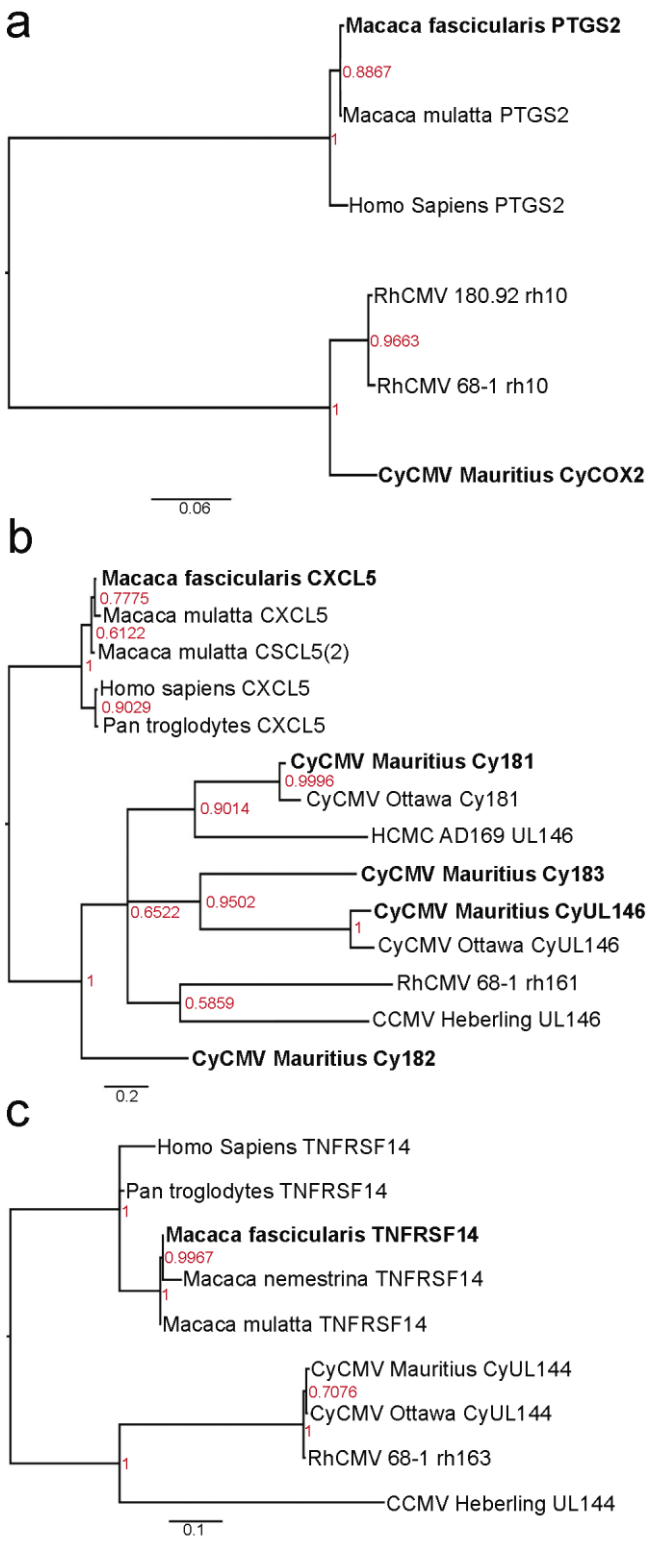


Figure 2-16: Phylogenetic tree comparison of CMV genes with mammalian homologues. Trees were prepared comparing CyCMV Mauritius, CyCMV Ottawa, RhCMV 68-1, RhCMV 180.92, CCMV Heberling and HCMV AD169 with mammalian homologue genes using the relevant *Homo sapiens* homologue as the outgroup. Trees were generated by Mr. Bayes following MAFFT alignment using a model of evolution selected by JModel test comparing a) CyCOX-2/*rh10* putative protein product to its mammalian homologue *PTGS2*; b) *UL146* and *UL146*-like CMV genes with mammalian homologue CXC chemokine ligand 5 (*CXCL5*); c) *UL144* with mammalian homologue tumor necrosis factor receptor superfamily 14 (*TNFRSF14*). Numbers at loci indicate posterior probability with color scaled according to probability; genetic distances measured in substitutions per site (sps) are given by scale below.

COX-2 homologues are absent from CyCMV Ottawa and HCMV but present in both RhCMV 68-1 and RhCMV 180.92 [288, 289, 298, 387] Interestingly, however, the CyCMV Mauritius genome has three ORFs with homology to COX-2: *CyCOX-2A*, *CyCOX-2B*, and *CyCOX-2C*. It is perhaps not coincidental that RhCMV 68-1, which has a COX-2 homologue the least diverged from mammalian COX-2, is also the strain with fewest functioning *US28* genes. Meanwhile, a high degree of divergence between COX-2 homologues in the three strains that retain it suggests a more recent integration or only a low fitness advantage. The absence of the COX-2 homologue in CyCMV Ottawa and the fact that RhCMV 68-1 and CyCMV Mauritius are genetically less distant than RhCMV 68-1 and CyCMV Ottawa may again suggest a faster rate of divergence of CyCMV Ottawa. Comparison of *COX-2* homologues showed a clustering of CMV homologues that are substantially diverged from their macaque homologs, *PTGS2* in *Macaca fascicularis* and *Macaca mulatta* (**Figure 2-16a**). There is no HCMV equivalent of *PTGS2*. It can be inferred that the *COX-2* gene was copied from a mammalian host once to an ancestral CMV prior to CyCMV and RhCMV speciation. A mammalian origin of the viral *COX-2* genes is further evidenced by the presence of introns in these genes. Perhaps the unique MHC haplotypes of Mauritian macaques favor the retention of an immune-modulatory gene, such as COX-2 homologue in CyCMV Mauritius.

Similarly, mammalian CXC chemokine ligand 5 (*CXCL5*) genes substantially diverged from CMV encoded *UL146*-like genes (**Figure 2-16b**). Interestingly, viral chemokine-encoding *UL146*-like genes such as CCMV *UL146* clustered with RhCMV rh161, whereas HCMV *UL146* clustered with CyCMV *Cy181* with weak support. Two interesting observations were noted. First, both CCMV *UL146* and HCMV *UL146* are more diverged from mammalian *CXCL5* genes (closest is 2.075 sps and 2.069 sps respectively) than they are from multiple macaque CMV genes (closest are 1.299 sps and 1.870 sps respectively). Second, the CyCMV Mauritius gene Cy182 is less diverged from all the mammalian *CXCL5* genes than it is from any CMV genes (the closest mammalian gene is *Macaca mulatta CXCL5* at 1.234 sps, and the closest macaque gene is CyCMV Mauritius Cy181 1.448 sps). This probably reflects variation in the rate of evolution among these genes and among CMV strains, with the CyCMV Mauritius being slower than the others (**Figure 2-16b**). The clustering of CCMV and HCMV with other CMVs may indicate the acquisition of a viral *CXCL5*

homologue in a common ancestral CMV strain followed by gene duplication and mutation to create the UL146-like genes. This phylogeny supports at least two independent gene duplication events before diversification of Old World Primates. Possibly after this, other gene duplication events generated additional gene copies in macaques. Alternatively, more than two gene duplication events occurred in the ancestor of Old World primates followed by more extensive gene loss in apes than macaques. It is interesting that no *UL146*-like genes were identified in RhCMV180.92. **Figure 2-16c** shows a well-supported close relationship between CMV UL144 genes of CyCMV Mauritius, CyCMV Ottawa, RhCMV 68-1 and CCMV Heberling to the exclusion of the homologous tumor necrosis factor receptor superfamily 14 (TNFRS14) genes in their macaque and human hosts, which is again consistent with a single copying event in the ancestor of these CMV strains.

2.5.6 CMVs and Their Hosts Share Similar Evolutionary Patterns

In order to test for co-evolution between host species and the CMVs that they harbor over a broader phylogeny, we performed a phylogenetic analysis of 12 available primate CMV viral genomes (**Figure 2-17**). The resulting phylogenetic tree of these CMVs was then compared to a previously estimated phylogenetic tree of mammalian genomes [528]. Phylogenetic relationships among multiple primate CMVs exhibit an identical pattern of CMV and host diversification (**Figure 2-17;** [528]). There is evidence of attenuation in some sequenced CMVs caused by growth of the isolated strains in tissue culture prior to sequencing. This can result in the mutation or loss of viral genes necessary for entry or growth in certain cell types, as is the case with the deletion of the *UL128-UL130* region in RhCMV 68-1 [200], and other laboratory adapted CMV strains [268, 529]. Despite this, there is no indication of cross species contact or attenuation of some strains during CMV diversification prior to laboratory isolation. DNA polymerase sequence comparisons have suggested a speciation for CyCMV Ottawa and RhCMV 68-1 of around 0.5 mya [290]. This places speciation and divergence from a common CMV strain at ~1.2 mya, in line with Y-DNA segregation and the suspected end to macaque migration between islands of the Sunda shelf [419, 420].

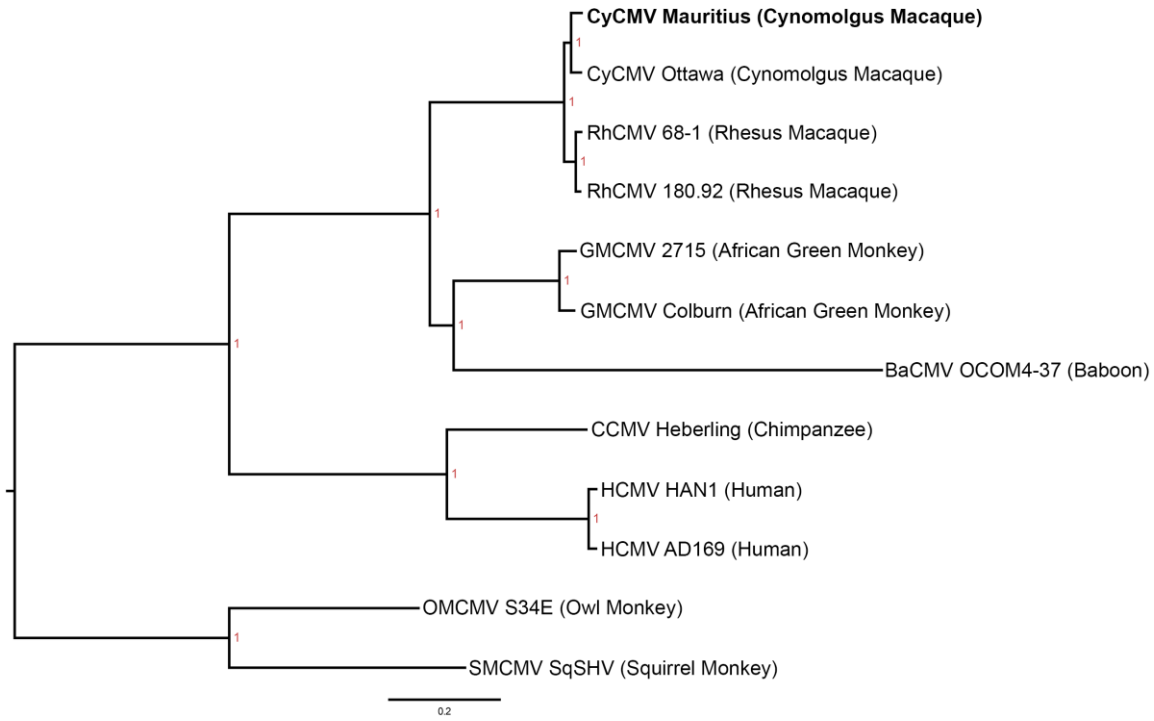


Figure 2-17: Bayesian phylogenetic tree of selected CMV strains. GTR tree with proportion invariant sites and gamma distribution of multiple sequenced primate CMV strains. Whole genome sequences were aligned using MAFFT with scoring matrix of 200PAM/k=2 a gap open penalty of 1.53 and offset value of 0.123. New world primates, OMCMV S34E and SMCMV SqSHV were used as the out-group to root the tree. RhCMV 68-1, RhCMV 180.92, CyCMV Ottawa, CyCMV Mauritius, CCMV Heberling, HCMV AD169, HCMV HAN1, GMCMV Colburn and 2715, and BaCMV OCOM4-37 were compared. Numbers at loci indicate posterior probability with color scaled according to probability; genetic distances measured in sps are given by scale below.

It is possible for multiple CMV strains to co-exist within a single macaque, or within a single macaque population, because CMV can superinfect seropositive individuals [530]. However, co-infection of a host may only be possible for closely related CMV strains – cynomolgus macaques, for example, are not susceptible to infection or co-infection by a RhCMV [291]. The correspondence between viral and host evolutionary relationships suggests further study of CMV would serve as a molecular tool to understand primate evolution.

2.6 Conclusions

In this study, we generated a novel genome sequence from a CMV strain isolated from a cynomolgus macaque from Mauritius. When analyzed with other CMV genomes from macaques and from other host species, inferred phylogenetic relationships among the viruses generally matched those among their hosts. Comparisons of this new genome with other macaque CMVs identify several functional categories of genes with atypically high levels of divergence, variation in gene content, and several genes with inferred phylogenetic relationships that differed from the genome-wide estimate. These results have implications for use of CMV as a vaccine vector and molecular tool, in CMV pathogenesis studies, as well as providing a tool to assist in tracing both viral and macaque migration and distribution.

2.7 Acknowledgments

This work was supported by the Canadian Institutes for Health Research (CIHR, grant number 107743). A CIHR Doctoral research award to A.K.M, an OHTN Junior Investigator Development Award to D.O.W and an OHTN Career Scientist Award to K.S.M. and postdoctoral fellowship to A.P.A are acknowledged. Computations were performed on the general purpose cluster supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto. The authors thank the veterinary and technical staff at the NHP colony of Health Canada for their help. The authors wish to thank the late Dr Allan J Baker of the Royal Ontario Museum for his thoughts on refining macaque evolution with co-evolved viruses and Catia Perciani for critical reading of the manuscript.

Chapter Three: Generation of a Novel SIV Vaccine Candidate Using Excisable, Full-length Cynomolgus Macaque Cytomegalovirus Bacterial Artificial Chromosome Technology.

Justen N. Hoffman Russell[1, 2, 4], Angie K. Marsh[2], Aruna P, Ambagala[3], Jacqueline K. Chan[2], Fatima C. Csordas[2], Meika Richmond [1, 4], Joseph M. Antony[2], David O. Willer[5] and Kelly S. MacDonald[1, 2, 4]

[1]Dept. of Internal Medicine, Section of Infectious Diseases, University of Manitoba, 745 Bannatyne Ave, Winnipeg, Manitoba, Canada, R3E 0J9 [2]Dept. of Immunology, University of Toronto, ON, Canada, M5S 1A8 [3]National Centre for Foreign Animal Diseases, Canadian Science Centre for Human and Animal Health, Winnipeg, MB, Canada. R3E 3M4 [4]JC Wilt Infectious Diseases Research Centre, Public Health Agency of Canada, Winnipeg, MB [5]GSK Vaccines, Rockville, MD, USA

3.1 Overview

This chapter described the construction of a CyCMV BAC and insertion of SIV transgenes to make possible testing of a CyCMV vector-based SIV vaccine in the Cynomolgus macaque vaccine model. The CyCMV Mauritius strain, isolated, sequenced and characterized, was cloned onto a BAC. To assess viral fitness the resulting CyCMV Mauritius BAC was sequenced in full and reconstituted as a virus in tissue culture. Abnormalities to the integrated DNA were successfully corrected by BAC recombination. The CyCMV Mauritius BAC was then designed as a vaccine against SIV. Two SIV transgenes introduced by bacterial recombination were expressed throughout viral infection

Main Objective: To prepare a Mauritius strain of CyCMV as a SIV vaccine for evaluation in cynomolgus macaques.

Specific Objectives:

1. Sequencing and genomic characterization of a Mauritian Cynomolgus Macaque strain of CMV (Chapter 2)
2. Comparative phylogenetic analysis of CyCMV Mauritius (Chapter 2)
3. **Cloning of CyCMV Mauritius as an excisable BAC (Chapter 3)**
4. **Utilization of CyCMV Mauritius to produce an SIV vaccine candidate (Chapter 3)**

3.2 Abstract

Cynomolgus macaques respond differently than rhesus macaques to a number of infectious diseases, including HIV/SIV [493, 494]. As interest in CMV as a vaccine vector has increased, the paucity of molecular reagents has hindered utilization of the cynomolgus macaque model.

To facilitate the use of this model, a cynomolgus macaque strain of CMV, CyCMV Mauritius, was cloned as a full length, excisable BAC. The BAC vector was integrated near the center of the linear CyCMV genome between *CyO19* and *CyO20*. A number of polymorphisms accumulated throughout the cloning process occurring predominantly on highly variable transmembrane RL11 family genes and Non-core family genes. Single amino acid changes were observed in two core genes, *CyUL47* and *CyUL102*, and in *CyUL87*, a Betagamma family gene. Partial attenuation of the *CyUL128*/*CyUL130*/*CyUL131* genes suggests that CyCMV Mauritius BAC is biased towards growth in fibroblasts. Despite this, reconstituted CyCMV Mauritius BAC growth kinetics were unaltered compared to wildtype virus grown in cynomolgus macaque, rhesus macaque, and human fibroblast cell lines. After confirming the integrity of the BAC, two SIV vaccine constructs composed of the structural and non-structural SIV proteins respectively were integrated into CyCMV Mauritius by BAC recombination and expressed in macaque cell lines. CyCMV Mauritius BAC is a powerful molecular tool expanding CMV vaccine vector studies to the cynomolgus macaque model.

3.2.1 Importance

Animal models are integral to the development of new vaccines. Of key importance is the concurrent use of complementary unique animal models, as no singular model can perfectly replicate vaccination in humans. The use of each animal model, however, requires the development of specialized biological tools and reagents. We created a tool based on the virus, CMV, which can be used in cynomolgus macaque monkeys. This work expands the possibility for CMV researchers to utilize this dynamic and unique primate model. In particular, cynomolgus macaque monkeys respond differently from other macaque monkey species to the simian analog of HIV known as SIV in important ways. To demonstrate the potential for this molecular tool, we have employed it to generate two SIV vaccine vectors that can be assessed in the cynomolgus macaque SIV challenge vaccine model.

3.3 Introduction

CMVs are large, ancient DNA viruses that have co-evolved with a wide array of hosts into highly species-specific variants. Unlike other chronic diseases, such as HIV or Hepatitis C, the immune response against natural CMV infection does not become exhausted [340, 341]. Throughout lifelong infection, robust CD8+ and CD4+ responses, including fast-acting CD8+ T effector memory cells, are maintained while viremia is limited to brief asymptomatic episodes unless there is immune suppression [185, 199, 309, 312, 342]. This continues to elicit the interest of immunologist, including vaccine researchers.

A number of CMVs have been cloned as BACs, most of them from lab-adapted strains of CMV, including multiple strains of human CMVs [481, 492, 531] and the RhCMV strain 68-1 [401]. When contained within a single low-copy BAC, the slow growing CMV viral genome can be faithfully preserved and replicated, or modified with relative ease by site-specific genetic manipulation [486, 492, 532]. A self-excising BAC integrates with the CMV genome without the loss of genes, and can be excised later by Cre/*loxP* recombination to reconstitute the unencumbered virus.

Recently our group isolated and characterized a novel strain of CyCMV [278] (See Chapter 2). CyCMV Mauritius, like its macaque host, was segregated from Sumatra in the 16th century common era (CE) as a small founder population [278, 430, 431, 433, 434, 436]. The virus is closely related to other isolated CyCMV strains, while genetically more distant from sequenced RhCMV strains [278, 379].

Due to their availability and controlled genetics, Mauritian cynomolgus macaques with their limited MHC variability have increasingly been the macaque of choice for a variety of research applications [426, 428, 437]. This is particularly poignant in HIV/SIV vaccine studies where, compared to rhesus macaques, cynomolgus macaques' show less abrupt loss of CD4+ cells, lower set-point viral loads, and slower disease progression [437, 463, 493, 494].

Here we describe the incorporation of CyCMV Mauritius into a BAC, creating a powerful molecular tool to allow more complete utilization of the cynomolgus macaque model. We further demonstrate the potential for the CyCMV Mauritius BAC by utilizing BAC recombination techniques to introduce two SIV vaccine transgenes within the CyCMV Mauritius viral genome. After reconstituting the CyCMV Mauritius BAC as a virus, both transgenes were retained and expressed faithfully within infected cells. The CyCMV Mauritius BAC offers an abundance of opportunity for CMV and CMV vaccine vector-based research.

3.4 Materials and Methods

3.4.1 Plasmids

Plasmids pWC132, encoding an eGFP cassette flanked by two *loxP* sites, pWC205, for Cre expression, and pWC155, containing the BAC vector were previously described [401]. Plasmid pGS284 [486] and pEF1/V5-His-A (Invitrogen) were used to generate allelic exchange shuttle vectors.

3.4.2 Viruses and Cells

CyCMV Mauritius (Accession: KP796148), isolated from a Mauritian cynomolgus macaque [278] (See Section 2.4.1), was primarily propagated in telomerase-immortalized rhesus macaque fibroblast cells (Telo-RF)[533]. Virally infected cells were maintained until cytopathic effect was observed in greater than 70% of cells. Infected cells were split at a 1:3 ratio and incubated for two to four hours over a healthy cell monolayer at 37˚C in an incubator with 5% CO_2 after which the inoculum was aspirated and replaced with fresh media. Telo-RF, Cynomolgus macaque-derived mellow skin fibroblast (MSF-T) [534], and human lung fibroblast-derived MRC-5 [504] cells were all maintained in DMEM (Sigma) supplemented with 10% FBS (Wisent Bioproducts), 100U/mL penicillin and 100µg/mL streptomycin (Sigma) at 37˚C in an incubator with 5% CO_2. Plasmids

pWC132, pWC2205, pWC155, pGS284, and Telo-RF cells[535] were generously provided by Drs. Peter Barry and William Chang (University of California, Davis).

3.4.3 Extraction of CyCMV Viral DNA

The cell culture media were collected and infected cells were washed with PBS then subjected to a HIRT extraction protocol with viral DNA separated from cellular DNA by sucrose gradient centrifugation, all as described previously [289].

Cellular debris was removed from collected cell media by centrifugation at 1835 × *g* and 4˚C for ten minutes followed by 0.45µm filtration. The filtrate was then centrifuged at 20 000 × *g* and 4˚C for one hour. The pellet was dissolved in 2.5 mL 10mM Tris-HCl pH 8.0 with 10mM EDTA. 30µg/mL RNAse A and 80 U/mL RNAse T1 (Thermo Scientific) were added to the pellet and incubated in a 37˚C water bath for four hours with Pronase added to a final concentration of 1mg/ml for the final two hours. Protein was removed with two phenol (pH 7.5) extractions followed by two phenol-chloroform: isoamyl alcohol (24:1) extractions. Sodium acetate was added to a 0.3M concentration. The DNA solution was mixed with two volumes of absolute ethanol and stored -20˚C for twelve hours. DNA was pelleted at 17 000 × *g* at 4˚C for 30 minutes, air dried and resuspended in 10mM Tris-HCl pH 8.0 with 10mM EDTA.

3.4.4 Restriction Enzyme Digests

Restriction enzyme digests were used to confirm the integrity of extracted DNA. DNA was digested for 18 hours with 1µL BamHI-HF, HindIII-HF or EcoRI (New England Biolabs) and visualized by gel electrophoresis.

3.4.5 Cloning CyCMV as a BAC

3.4.5.1 Generation of CyCMV Mauritius –eGFP

An amplification plasmid was prepared by cloning the CyUS1 to CyUS2 spanning region of CyCMV Ottawa (Accession: JN227533) [407] together between the *Not*I and *Pst*I sites. A *loxP* linker was ligated between cyUS1 and cyUS2 using *Asc*I. CyUS1 to *loxP* and CyUS2 to *loxP* regions were separately amplified and ligated into pWC-132 digested with *Nde*I/*Afl*II and *Pvu*II/*Sal*I. The resulting plasmid was transfected into Telo-RF cells in a six well plate with 3.2 µg of plasmid DNA using Lipofectamine 2000 (Invitrogen) reagent and grown for twenty-four hours prior to infection with CyCMV Mauritius at a multiplicity of infection (MOI) of 0.2, 2, and 20. Cells were cultured to 100% CPE; supernatant was harvested and frozen -80˚C with 0.2 M sucrose. CyCMV Mauritius-eGFP was selected through six rounds of plaque purification.

3.4.5.2 Plaque Purification

Viral supernatant was overlaid on Telo-RF cells in 10-fold serial dilutions for two hours. Medium containing inoculum was removed and the cells were overlaid with a 1:1 mixture of 2× MEM containing 10% FBS and 4mM L-glutamine (Quality Biologics Inc.) with 100U/mL penicillin and 100µg/mL streptomycin (Sigma), and 2% low-melting agarose (Type VII; Sigma). Individual plaques were selected based on eGFP fluorescence or absence and propagated on Telo-RF cells for subsequent rounds of plaque purification.

3.4.5.3 PCR Confirmation

The primers 5’-ACCGGACTTTTTGGCGCATGACT-3’ and 5’-GCCAGTCCCCGCACTGAC-3’ were designed to amplify a region of CyCMV Mauritius between *CyUS1* and *CyUS2* flanking the site of insertion for eGFP, *loxP*, and the BAC. Following viral DNA isolation, Sanger sequencing was performed to confirm the amplified sequence.

3.4.5.4 Generation of CyCMV Mauritius-loxP

Telo-RF cells were transfected in a six well plate with 3.2 µg/well pWC205 DNA using Lipofectamine 2000 reagent (Invitrogen) and grown twenty-four hours prior to infection with CyCMV Mauritius-eGFP in 10-fold serial dilutions. Supernatant was harvested at 100% CPE and frozen at -80˚C with 0.2M sucrose. CyCMV-loxP, which no longer expressed eGFP, was selected through four rounds of plaque purification.

3.4.5.5 Generation of CyCMV Mauritius BAC:

Purified CyCMV Mauritius-loxP propagated in Telo-RF cells was concentrated on a 20% sorbitol cushion with 50mM Tris-HCl (pH 7.2) and 1 mM $MgCl_2$ by centrifugation at 71 000 × *g* for 90 minutes at room temperature. Concentrated virus was titred on Telo-RF cells using a standard plaque assay (see below). Telo-RF cells were co-transfected in a six well plate with 3.2µg total DNA/well of pWC205 and pWC155 in a 3:2 ratio and grown twenty-four hours prior to infection with 0.3 or 0.03 MOI of concentrated CyCMV-loxP. Supernatant was harvested at forty-eight hours or at ninety-six hours and frozen at -80˚C with 0.2M sucrose. BAC containing virus was selected by 2-4 rounds plaque purification and then propagated on Telo-RF cells.

3.4.5.6 Electroporation into Bacterial Cells

Virus was concentrated by centrifugation on a 20% sorbitol cushion with 50mM Tris-HCl (pH7.2) and 1mM $MgCl_2$ by centrifugation at 71 000 × *g* for 90 minutes at room temperature. Concentrated virus was titred on Telo-RF cells using a plaque assay (See Below)

Telo-RF cells were grown to confluency and infected with 0.06 and 0.0006 MOI CyCMV Mauritius BAC and cultured for 120 hours. Circularized viral DNA was extracted using a modified HIRT method as described previously [492].

Next, 2-15 µg of extracted DNA was electroporated into 20 µL GS500 *Escherichia coli* (*E. coli*) cells (Invitrogen) using 1mm cuvette with 1.8 kV, 25 µF capacitance, 200Ω resistance, and

4.4 millisecond time constant. Cells were recovered in 1mL recovery media (Invitrogen) with shaking at 225 rotations per minute (rpm) 37˚C for one hour. Bacteria were pelleted at 20 000 × *g* for five minutes, resuspended in 80 µL recovery media and plated out on Miller's Luria-Bertani broth (LB) agar plates with 20 µg/mL chloramphenicol. Positive colonies were selected after growth for twenty-four hours at 37˚C.

3.4.5.7 Confirmation of Positive Colonies

Positive colonies selected above were confirmed first by colony PCR to contain DNA of both CyCMV Mauritius and the BAC. Primers 5'-CCCGCCCTGCCACTCATCG-3' and 5'-ACATTTTGAGGCATTTCAGTCAGTTGC-3' were used to amplify a segment of the chloramphenicol resistance gene of the BAC and primers 5'-CTGACTTCCGCCGCCGCTTC-3' 5'-GGGCAGCAACCCCGATCACC-3' were used to amplify the *parA/repE* gene of the BAC. Primers 5'-TTGTCAAGAACACCGTGCGCAAC-3' and 5'-AATGAGAGCTGCGGCCGTGACA-3' were used to amplify a region of *CyUL54* to confirm the presence of the CMV genome.

3.4.6 Next Generation Sequencing

2-4 µg of CyCMV Mauritius BAC DNA from two clones was isolated from bacterial culture and prepared for sequencing with an Illumina HighSeq MiSeq Sequencer. A DNA library was generated and run to create 1 575 392 and 1 031 166 paired end reads of 155pb for ~100× coverage. Reads were aligned using a Burrows-Wheeler aligner algorithm [536] run against the expected sequences generated *in silico* using Geneious pro 6.1.4 (Biomatters Ltd. Auckland, New Zealand) based off the sequenced CyCMV Mauritius genome (Accession: KP796148). This was confirmed against the three largest contigs generated from paired end reads in a *de novo* assembly to confirm the BAC insertion site. Homology of the construct was then compared to CyCMV Mauritius using a global pairwise alignment with a cost matrix of 70% similarity (5.0-4.5), a gap open penalty of twelve, gap extension penalty of three, and two refinement iterations. Homology of predicted protein sequences was

calculated using MAFFT Alignment [508, 537] with a BLOSUM62 scoring matrix, gap open penalty of 1.53 and offset value of 0.123.

3.4.7 Growth Curves

The growth kinetics of two CyCMV Mauritius BAC clones were compared to the corresponding wildtype CyCMV Mauritius strain by infecting Telo-RF cells. Cells were grown to a confluent monolayer in multi-well plates and then growth media replaced with 0.5 mL fresh media with virus at 0.25, 0.025 or 0.0025 MOI. Growth media was topped up to 3 mL after two hours then replaced with fresh media after twenty-four hours. After two, four, six, eight or ten days of growth monolayers of duplicate six well plates were disrupted using a cell scraper and homogenized with the media by vigorous pipetting. Cell/media homogenate from multiple wells were pooled and frozen at -80˚C with 0.2 M sucrose until titreing with a Plaque Assay.

3.4.8 Plaque Assay

The viral titer was determined in duplicate by first infecting Telo-RF monolayers. Then monolayers were incubated at 37˚C and 5% CO_2 in 0.5 mL growth media with serial dilutions of thawed viral homogenate. After three hours, media was replaced with 3 mL of a 1:1 mixture of 2% low melting temperature agarose (Lonza) and 2× MEM (Quality Biologic) supplemented with 10% FBS (Wisent Bioproducts), 100U/mL penicillin and 100µg/mL streptomycin (Sigma), 4mM L-alanyl-L-Glutamine (Gibco), and 100 µg/mL phenol red (Sigma). Media was solidified ten minutes at room temperature followed by ten minutes at 4˚C, then the plates were cultured for fourteen days at 37˚C with 5% CO_2. Solid media was carefully removed and the monolayer stained with 1% crystal violet. Plaques were averaged from all countable dilutions of each replicate.

3.4.9 *LoxP* Correction

1968 bp and 1845 bp sequences incorporating the upstream and downstream *loxP* flanking regions of sequenced CyCMV-BAC respectively were synthesized with the *loxP* site corrected to 5’-

ATAACTTCGTATAATGTATACTATACGAAGTTAT-3’ (See Section 3.8; **supplementary Materials 3-1** and **Supplementary Materials 3-2**). Synthesized DNA was ligated into pGS284 cut with NheI and BglII and grown in S17-1 λpir *E. coli*. Serial rounds of BAC recombination (see below) were utilized to correct each BAC *loxP* site, with successful exchanged screened for using the primers 5’-ACTAGCGATGTCATTCCGCC-3’ and 5’-CACACCTCCCCCTGAACC-3’ to amplify the upstream *loxP* site or 5’-GTATGCCTGCTGTGGATTGC-3’ and 5’-CTTTCTCGTTATGCCGCTGC-3’ to amplify the downstream *loxP* site. Corrected *loxP* sites were differentiated from uncorrected sites after colony PCR amplification by successful digestion with AccI.

3.4.10 BAC Recombination

S17-1 λpir *E. coli* containing pGS284-based shuttle plasmid and GS5000 *E. coli* containing CyCMV-BAC were streaked on LB Agar with 100µg/mL carbenicillin or 20µg/mL chloramphenicol respectively and grown overnight at 37˚C. A cross streak plate was generated using an individual GS5000 colony streaked across the center of an LB agar plate. Individual S17-1 λpir colonies were picked and streaked perpendicularly across the initial streak then grown at 37˚C overnight to facilitate plasmid transfer. A flame loop was used to inoculate multiple 3mL tubes of LB media with 100µg/mL carbenicillin and 20µg/mL chloramphenicol with bacterial swabs taken from the intersections of the cross streak plate, before growth at 37˚C and 225rpm overnight to select for *E. coli* incorporating both BAC and plasmid. To allow multiple generations for allelic exchange to occur, streak plates were created from the resulting growth on LB agar with 100µg/mL carbenicillin and 20µg/mL chloramphenicol, incubated overnight at 37˚C, then individual colonies picked onto an LB agar with 100µg/mL carbenicillin and 20µg/mL chloramphenicol replication plate and grown at 37˚C overnight. Next, individual colonies were inoculated into 3mL tubes of LB media with 20µg/mL chloramphenicol and grown overnight at 37˚C 225 rpm to remove carbenicillin selective pressure. To select for colonies without pGS284-based shuttle plasmid a flame loop was used to inoculate high-

sucrose/no-salt LB (10g tryptone, 5g yeast extract, 50g sucrose per liter water) media from LB media with 20µg/mL chloramphenicol tubes and grown at 30˚C 225 rpm overnight. High-sucrose/no-salt LB agar streak plates were used to generate individual colonies and grown at 30˚C overnight. Individual colonies were picked and plated on LB Agar with 20µg/mL chloramphenicol and on LB agar with 100µg/mL Carbenicillin for growth at 37˚C overnight. Colonies which grew on chloramphenicol, and thus incorporated the CyCMV BAC, but not on carbenicillin, and thus had not retained the pGS284-based shuttle plasmid, were selected and screened by colony PCR.

3.4.11 Insertion of Transgenes

The *CyUS2* to *Cy116* spanning region of CyCMV Mauritius was amplified using the primers 5’ ATGCgcggccgcTCGACTTGCACCTCATTCACATAGC-3’ and 5’-ATGCggatccCCTCTCTGCGGCTCATGCGG-3’, containing added NotI and BamHI 5’ tails (lowercase) and incorporated onto a blunt cloning vector (Invitrogen). The central AvrII site was changed to an AscI site by digesting with AvrII, treating with calf intestinal phosphatase, and ligating with the self-dimerizing oligonucleotide 5’-CTAGTGGCGCGCCA-3’. DNA was then extracted using NotI and BamHI, and ligated into pGS284 digested with BglII and NotI, grown in S17-1 λpir *E. coli.*

NTR and GPF transgenes were constructed from previously described vaccine constructs [345]. NTR, a codon optimized NeTaRev fusion protein, was inserted into pEF1/V5-His-A digesting with KpnI and NotI, while GPF a GagPol fusion protein, was PCR amplified from a GagPol fusion construct using the primers 5’-actgtctagagcggccgcAGTGTGCTGGAATTCGCCC-3’ and 5’-actgggtaccagatctGGATATCTGCAGAATTCGCCC-3’ then inserted into pEF1/V5-His-A digesting with KpnI and NotI. Both constructs were amplified to include the EF1α promoter and Bgh polyadenylation tail using the primers 5’-actgggcgcgccGGATACCCCCTAGAGCCCC-3’ and 5’-actgggcgcgcCACAGAGAGGAATCTTTGCAGC-3’ with 5’ AscI containing tails. Each transgene construct was digested and ligated separately into the AscI site of the above described pGS284-

CyUS2-Cy116 vector. NTR and GPF transgenes were transferred into the *loxP* corrected CyCMV Mauritius BAC by allelic exchange and expression confirmed by western blot.

3.4.12 Western Blot

DNA of *loxP* corrected CyCMV Mauritius with NTR or GPF incorporated was transfected into a confluent monolayer of Telo-RF cells using Lipofectamine 3000 in a six well tissue culture plate. Cell monolayers were monitored and maintained 2-3 weeks until viral cytopathic effect was observed across more than 50% of the cell monolayer and virus was passaged as above. Cells were maintained for 5-10 days until viral cytopathic effect was observed in greater than 80% of cells. The cell monolayer was disrupted with a cell scraper and homogenized by pipetting, then separated into 1.5mL aliquots and centrifuged 10 600 × *g* for five minutes. Cell pellets were resuspended in RIPA buffer (50mM Tris HCL, 150 mM NaCl, 1% Triton X100, 0.5% Sodium Deoxycholate, 0.1% SDS, 1mM EDTA, pH adjusted to 7.4) with 1× protease inhibitor cocktail (Sigma P-2714) added and incubated on ice for thirty minutes to lyse before 4˚C centrifugation at 20 817 × *g*. The supernatant was collected and quantitated with a Bradford protein assay (Bio-Rad) to determine protein concentration and then frozen at -20˚C until used.

The protein lysate was diluted to 5µg of protein with NuPAGE LDS loading dye (Life Technologies) and NuPAGE reducing agent (Invitrogen) in 25uL, and then denatured on a thermocycler at 70˚C for ten min, followed by freezing at -20˚C overnight. Samples were run at 200 V for 30-45 min in a precast TGX SDS-PAGE gel (Bio-Rad), and transferred to a PVDF membrane at 20V for seven minutes using an iBlot apparatus (Invitrogen) with PVDF iBlot Gel transfer stack (Life Technologies).

Non-specific binding to the membrane was blocked by rocking gently for one hour with 40mL blocking buffer (Odyssey Blocking Buffer; Li-Cor). The membrane was stained for Gag (NIH AIDS Research & Reference Reagent Program: KK59 at 1:1000) or Nef (NIH AIDS Research & Reference Reagent Program: clone 17.2 at 1:1000) and IE-1 (mouse anti- Rhesus CMV IE-1 clone 2A1.2 at

1:750; generously provided by Drs. Scott Hansen and Louis Picker of Oregon Health and Science University) in blocking buffer with 0.1% Tween 20 at room temperature for one hour while gently rocking. Membrane fragments were washed three times for five minutes each gently rocking, with Tris buffered Saline (TBS; 50mM Tris, 150 mM NaCl) + 0.1% Tween 20 (TBST). Membranes were then incubated gently rocking in the dark with secondary antibody (Li-Cor; IRDye 800CW Goat (polyclonal) anti-mouse IgG (H + L) at 1:15000) in blocking buffer + 0.1% Tween 20 and washed a further three times with TBST and once with TBS. Fluorescent bands were visualized using an Odyssey CLx (Li-Cor).

3.5 Results

3.5.1 CyCMV Mauritius was Cloned as a BAC

The linear CyCMV Mauritius [278] was cloned as a circular BAC. As the BAC vector was too large to insert directly by homologous recombination into CyCMV Mauritius, two intermediate viral constructs were created to facilitate the insertion of the vector between *CyUS1* and *CyUS2* in a stepwise approach (Fig**ure 3-1**). Initially a smaller *eGFP* construct with two *loxP* sites was inserted into CyCMV Mauritius to produce CyCMV Mauritius-eGFP. Successfully recombined CyCMV Mauritius-eGFP virus was then separated from wildtype CyCMV Mauritius by successive rounds of plaque purification screening for eGFP fluorescence. PCR amplification with primers internal to the *eGFP* region confirmed construction of the CyCMV Mauritius-eGFP intermediate and indicated a single nucleotide C to T transversion at position six of the inverted repeat region of one of the two inserted *loxP* sites (**Figure 3-2**).

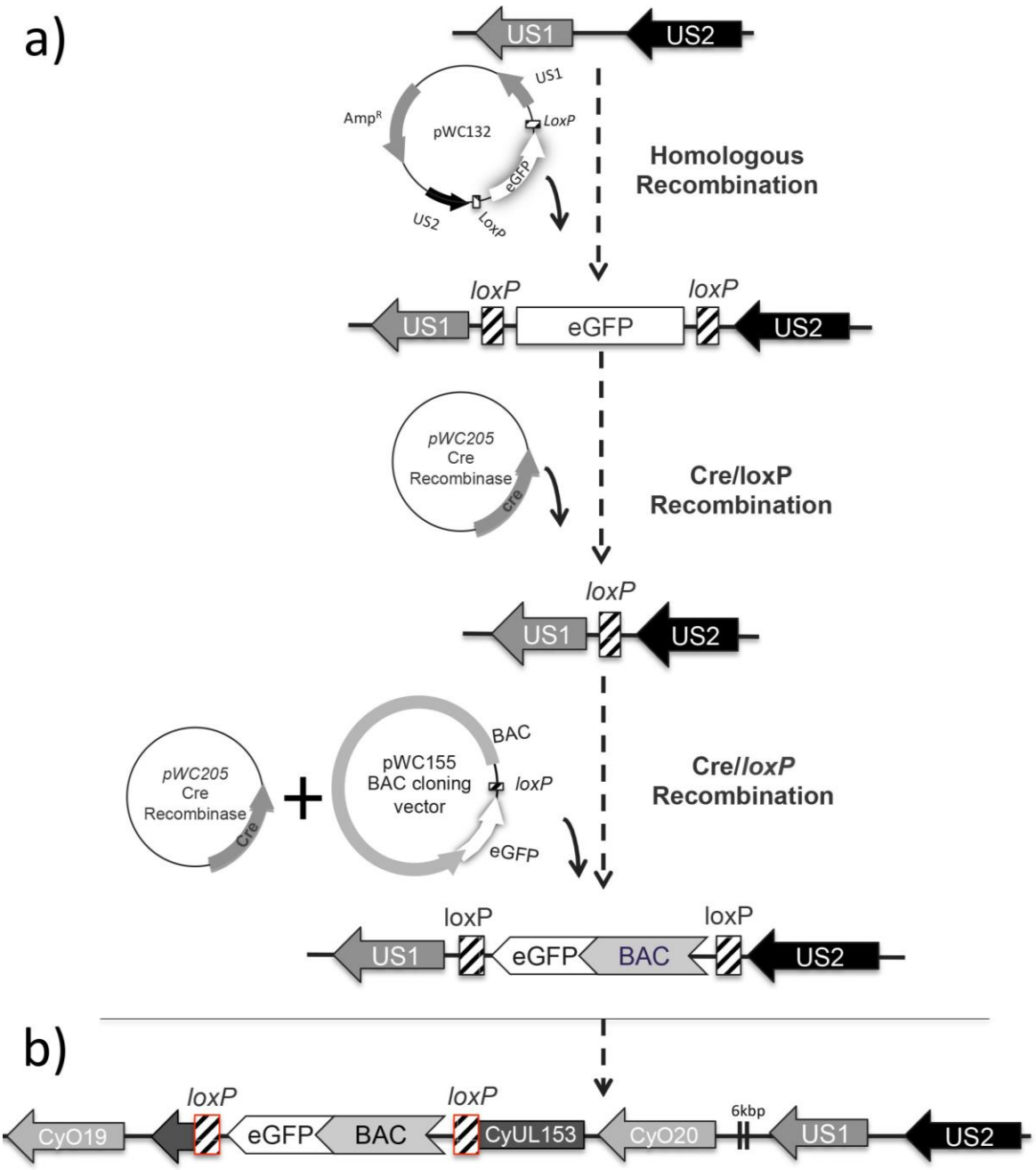


Figure 3-1: Cloning scheme to incorporate CyCMV Mauritius into a BAC. A) Planned cloning scheme for the generation of CyCMV Mauritius BAC. CyCMV Mauritius virus was cultured in cells with a modified pW132 plasmid homologous to the CyUS1-CyUS2 genome region. Homologous recombination incorporates *eGFP* and two flanking *loxP* sites into some of the CyCMV Mauritius resulting in the first intermediate: CyCMV Mauritius-eGFP. After plaque purification CyCMV Mauritius-eGFP was cultured in cells with Cre expression induced. Cre recombination excises eGFP and recombines the two *loxP* sites into one, producing CyCMV Mauritius-LoxP. After further plaque purification CyCMV Mauritius-LoxP is cultured in Cre expressing cells with pWC155, the BAC vector. Cre-*loxP* recombination inserts the BAC vector into CyCMV Mauritius-LoxP yielding CyCMV Mauritius BAC. **B)** BAC DNA integrated at a cryptic *loxP* site in CyUL153 ~6kbp upstream of the intended insertion site. The CyUL153 ORF was disrupted and both flanking *loxP* sites differed from *loxP* 511.

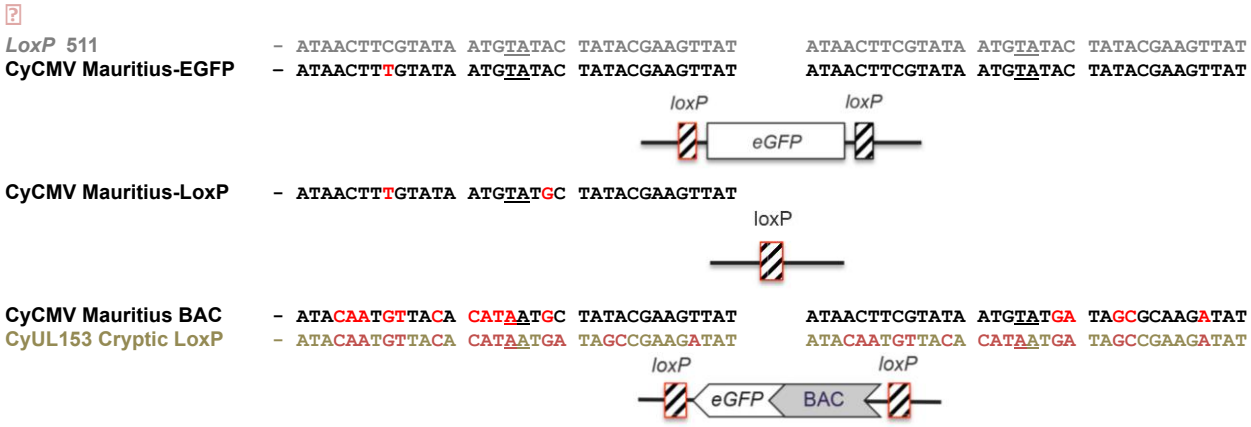


Figure 3-2: Polymorphisms on *LoxP* sites. *LoxP* 511sequences (grey; top) were introduced into CyCMV Mauritius to facilitate further cloning steps. Two *loxP* sites were introduced into the first intermediate CyCMV Mauritius-EGFP. These sites were recombined into the single *loxP* site of the second intermediate CyCMV Mauritius-LoxP. To produce CyCMV Mauritius BAC the single *loxP* 511 sites on the BAC plasmid integrated with a cryptic *loxP* site in the *CyUL153* ORF (beige; bottom; depicted in duplicate for comparison). In all *loxP* 8 bp spacer regions are separated from 13 bp recognition regions. Alterations from *loxP* 511 are in red. The central TA, the splice site of *loxP*, is underlined.

The second intermediate, CyCMV Mauritius-loxP, was generated from CyCMV Mauritius-eGFP by culturing in Cre recombinase-expressing cells to facilitate the excision of eGFP and recombination of the two inserted *loxP* sites into a single *loxP* **(Figure 3-1)**. In a second round of plaque purification CyCMV Mauritius-loxP was purified from CyCMV Mauritius-eGFP by the absence of eGFP fluorescence.

The single *loxP* site of CyCMV Mauritius-loxP was sequenced and retained the C to T transversion at position six from CyCMV Mauritius-eGFP, as well the site acquired an A to G transition at position seven of the spacer region (**Figure 3-2**). The A to G transition has been shown to improve Cre-*loxP* recombination efficiency [538, 539] therefore we proceeded using CyCMV Mauritius-loxP to generate CyCMV Mauritius BAC.

Cre-*loxP* recombination inserted the entire BAC vector including an integrated eGFP into CyCMV Mauritius-loxP, however, sequencing confirmed that the insertion occurred ~6kbp upstream of the

intended site between genes *CyO19* and *CyO20* at a *loxP*-like site in the *CyUL153* ORF (**Figure 3-1b**).

As BAC vectors have been shown to induce instability in surrounding DNA during viral growth, the number of sequential plaque purifications utilized for the eGFP-expressing BAC was minimized and multiple clones were maintained [486]. Circularized viral DNA from two clones was extracted and transformed into *E. coli* to recover self-replicating BAC.

3.5.2 Sequence Analysis of CyCMV Mauritius BAC

Complete sequencing of two clones of CyCMV Mauritius BAC (**GenBank:** MF918486 and MF918487) confirmed the insertion of BAC DNA within the *CyUL153* ORF (**Figure 3-1b**). The sequence confirmed multiple single nucleotide polymorphisms (SNPs) differentiating the BAC flanking *loxP* sites from *loxP* 511 within both CyCMV Mauritius BAC clones (**Figure 3-2**). As the T to A transversion at position four of the spacer region in one *loxP* site has been shown to prevent Cre-*loxP* recombination [538, 539], BAC recombination was used to correct both *loxP* sites on CyCMV Mauritius BAC clone 1 within bacterial culture.

The plaque purification process and repeated passaging of CyCMV Mauritius necessary to create the BAC introduced a number of mutations that differentiated both CyCMV Mauritius BAC clones from wildtype virus. In the CyCMV Mauritius BAC, 789 and 886 total polymorphisms were identified within reading frames of Clone 1 and Clone 2 respectively (**Figure 3-3**). Of these, 682 and 719 respectively were SNPs, while the remainder were insertions or deletions.

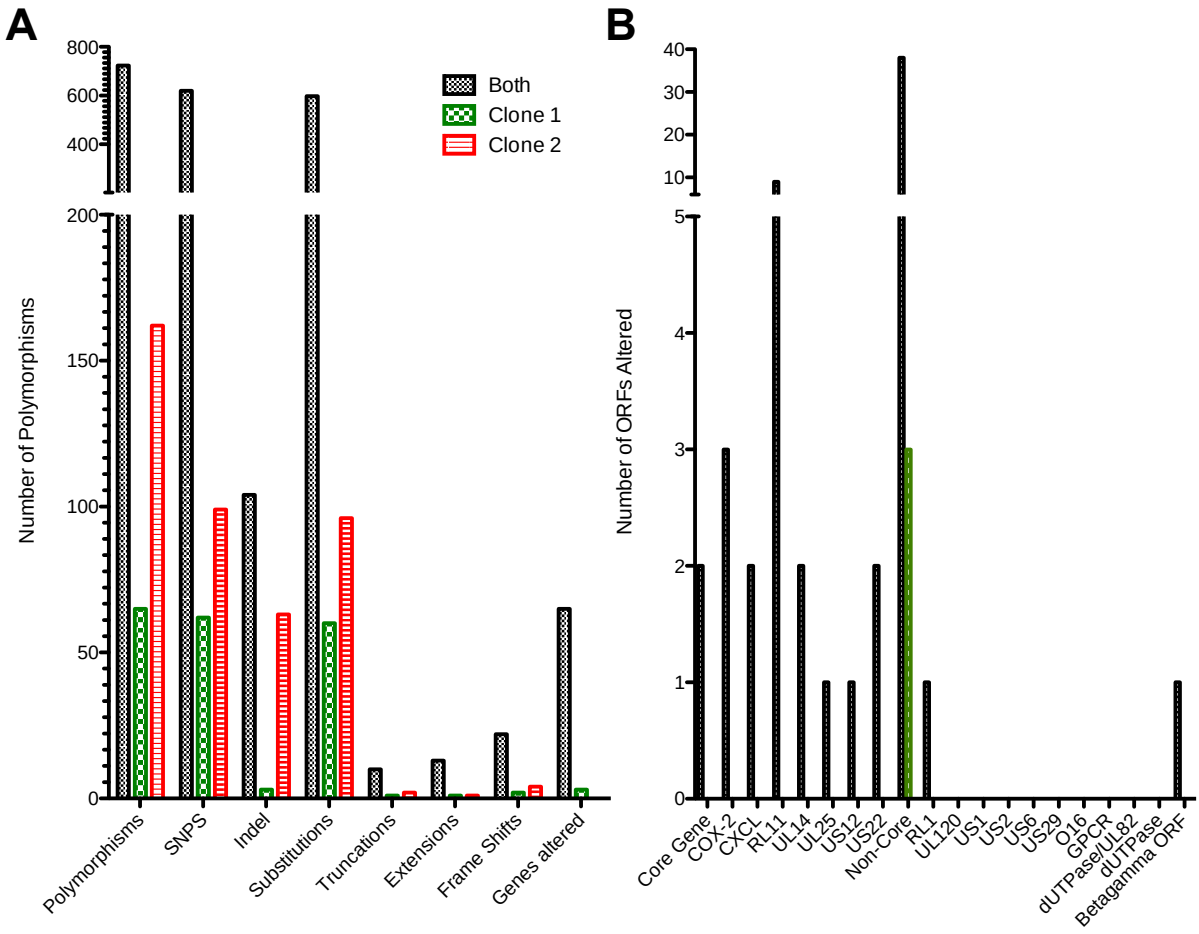


Figure 3-3: Unintentional mutations across the CyCMV Mauritius BAC clones. Mutations were introduced while passaging CyCMV Mauritius BAC and intermediates throughout the cloning process. In black are mutations common to both clones, in red or green are mutations unique to one clone. A) Total number of polymorphisms, and subcategorization by type of mutation introduced. Categories are not exclusive. B) CyCMV Mauritius BAC genes altered from CyCMV Mauritius virus, grouped by herpesvirus gene family, in which one or more amino acids were altered.

The 64 and 60 genes altered in clone one and clone two respectively were compared with respect to gene function if known (**Table 3-1**). Only two core genes were altered in the two clones, *CyUL48* with SNP resulting in a Ser to Asn substitution and *CyUL102* with the insertion of a serine. The majority of genes altered were of the Non-Core gene family, which consists predominantly of species-specific CMV genes of unknown function. A number of RL11 genes, the putative COX-2 homologue genes, and both secreted CXC5 chemokine like glycoproteins, *CyUL146* and *CyUL147*,

contained multiple disruptive mutations. Genes altered in only one of the two clones, or altered differently in both clones were all of the Non-core or RL11 families. To assess the fitness of the BAC clones in view of the mutations a quantitative assay of *in* vitro viral growth was undertaken.

Table 3-1: Genes altered from wildtype in CyCMV Mauritius BAC Clones

Clone	Mutated Gene	Gene Family	Polymorphism type	Mutation Effect	Gene Necessity
1 & 2	CyUL87	Betagamma ORF	SNP (transversion)	K -> N	Essential for Growth[1,2]
1 & 2	CyUL47	Core Gene	SNP (transition)	S -> N	Effect on viral titre in HCMV[3]
1 & 2	CyUL102	Core Gene	Insertion	S -> SS	Essential in HCMV[1]
1 & 2	CyCOX2A	COX-2	Deletion & SNPs	Multiple	Non-Essential[2]
1 & 2	CyCOX2B	COX-2	Frame Shift and SNPs	Multiple	Non-Essential[2]
1 & 2	CyCOX2C	COX-2	Frame Shift and SNPs	Multiple	Non-Essential[2]
1 & 2	CyUL146	CXCL	Multiple	Multiple	Non-Essential[1]
1 & 2	CyUL147	CXCL	Multiple	Multiple	Non-Essential[1,2]
1 & 2	CyTRL1	RL1	Multiple Substitutions	Multiple	Cell Tropism Factor[2]
1 & 2	Cy20	RL11	Large Deletion (48 nt)	Deletion	
1 & 2	CyRL11Q(cy22)	RL11	SNP X 5	Multiple	Non-Essential[2]
1 & 2	CyRL11R(cy23)	RL11	Multiple	Multiple	
1 & 2*	CyRL11S(cy24)	RL11	Multiple	Multiple	Non-Essential[2]
1 & 2	CyUL6	RL11	Large Deletion (591 nt)	Extension	Non-Essential[1,2]
1 & 2	CyUL7	RL11	Multiple	Multiple	Non-Essential[1,2]
1 & 2	CyUL9a	RL11	Large Deletion (684 nt)	Extension	
1 & 2	CyUL9b	RL11	Large deletion (363 nt)	Extension	
1 & 2*	CyUL9c	RL11	Multiple	Multiple	
1 & 2	CyUL9d	RL11	Multiple	Multiple	
1 & 2	CyUL11	RL11	Multiple	Multiple	Effect on Viral Growth in HCMV[1]
1 & 2	CyUL14	UL14	SNP (transversion)	D -> E	Non-Essential in RhCMV/Moderate Growth Defects in HCMV[1,2]
1 & 2	CyUL25	UL25	SNP (transversion)	H -> Q	Non-Essential[1,2]
1 & 2	CyUL141	UL14	SNPs	Multiple	Non-Essential[2]
1 & 2*	CyUL153	RL11	Multiple	Multiple	
1 & 2	CyUS18	US12	SNP (transition)	T -> A	Non-Essential[1]
1 & 2	CyUL26	US22	SNP X 4	Multiple	Severe Growth Defect in HCMV[1]/Non-Essential[2]
1	Cy02	Non-Core	SNP (transversion)	Multiple	

Clone	Mutated Gene	Gene Family	Polymorphism type	Mutation Effect	Gene Necessity
1	Cy03	Non-Core	SNP (transversion)	Multiple	
1 & 2*	Cy28	Non-Core	Multiple Substitutions	Multiple	
1 & 2	Cy46	Non-Core	Multiple Substitutions	Multiple	
1 & 2	Cy49	Non-Core	SNP (transition)	I -> V	
1 & 2*	Cy61	Non-Core	Multiple	Multiple	
2	Cy62	Non-Core	SNP (transversion)	Multiple	
1 & 2	Cy92	Non-Core	SNP (transversion)	Multiple	
1 & 2*	Cy93	Non-Core	Multiple Substitutions	Multiple	
1	Cy107	Non-Core	SNP (transversion)	R -> S	
1	Cy146	Non-Core	Multiple	Multiple	
1 & 2*	Cy165	Non-Core	Multiple Substitutions	Multiple	
1 & 2	Cy168	Non-Core	SNP (transversion)	R -> G	
1 & 2	Cy171	Non-Core	Multiple Substitutions	Multiple	
1 & 2	Cy172	Non-Core	Multiple Substitutions	Multiple	
1 & 2	Cy174	Non-Core	Multiple	Multiple	
1 & 2	Cy181	Non-Core	Large Deletion (195 nt)	Extension	
1 & 2	Cy182	Non-Core	Large Deletion (327 nt)	Extension	
1 & 2	Cy183	Non-Core	Multiple	Multiple	
1 & 2	CyO1(Cy06)	Non-Core	SNP (transition)	S -> G	
1 & 2	CyO11	Non-Core	SNP (transition)	T -> A	
1 & 2	CyO13 (Cy189)	Non-Core	11 SNPs	Multiple	
1 & 2	CyO14 (Cy190)	Non-Core	SNP (transversion)	S -> R	
1 & 2	CyO15 (Cy191)	Non-Core	SNP (transversion)	S -> R	
1 & 2	CyO8	Non-Core	Multiple	Multiple	
1 & 2	CyRL11B(Cy09)	Non-Core	Insertion (tandem repeat)	Frame Shift	
1 & 2	CyUL116	Non-Core	Deletion	Frame Shift	Non-Essential[2]
1 & 2	CyUL128_ex1	Non-Core	Multiple	Multiple	Cell Tropism Factor[1,5]
1 & 2*	CyUL128_ex2	Non-Core	Insertion	Multiple	Cell Tropism Factor[1,5]
1 & 2	CyUL130	Non-Core	Deletion	Extension	Cell Tropism Factor[1,5]
1 & 2	CyUL131A	Non-Core	Multiple Substitutions	Multiple	Cell Tropism Factor[5]
1 & 2	CyUL132	Non-Core	Multiple	Multiple	Effect on Viral Growth in HCMV/Non-Essential in RhCMV[1,2]
1 & 2	CyUL147A	Non-Core	SNP X3	Multiple	Hypervariable Region/ Frequent Mutations[4]
1 & 2	CyUL148	Non-Core	Multiple	Multiple	Non-Essential[2]
1	CyUL19	Non-Core	SNP X3	Multiple	Non-Essential[1,2]
1 & 2	CyUL20	Non-Core	Multiple	Multiple	Effect on Viral Growth in HCMV/Non-Essential in RhCMV[1,2]

Clone	Mutated Gene	Gene Family	Polymorphism type	Mutation Effect	Gene Necessity
1 & 2	CyUL21A	Non-Core	Deletion (tandem repeat)	Extension	Cell Tropism factor/Moderate Growth Defects in HCMV[6]
1 & 2	CyUL42	Non-Core	SNP (transversion)	R -> T/ Truncation	Non-Essential[1,2]
1 & 2	CyUL74A	Non-Core	Deletion	Multiple	

Footnotes:

* indicates that mutations in specified gene is different in each clones
1: Based on study with HCMV Towne strain [51]
2: Based on study with RhCMV 68-1 [71]
3: Based on study of HCMV UL47 protein function [72]
4: Based on study of HCMV UL146 through UL147A region [73]
5: Based on study of HCMV Ul131A region [74]
6: Based on study of HCMV UL21a protein [75]

3.5.3 Growth Analysis of CyCMV Mauritius Reconstituted from BAC as Infectious Virus

The CyCMV Mauritius BACs were transfected into cell culture then passaged and maintained as viruses. Infections with both CyCMV Mauritius BACs were observed in the cynomolgus macaque derived MSF-T cell line [534], Rhesus macaque derived Telo-RF cell line [535] and the human derived MRC-5 cell line [504]. In all three cases, robust infection with typical cytopathic effect comparable to infection with wildtype CyCMV was observed.

The growth rate of the CyCMV Mauritius BACs were quantitatively compared against wildtype low passage CyCMV Mauritius in duplicate with single cycle growth curves across a range of initial MOI (**Figure 3-4**). At 0.0025 MOI and 0.025 MOI, viral burst size was comparable in both CyCMV Mauritius BAC clones and wildtype CyCMV Mauritius, though in Clone 2 this peak was reached later, at day ten instead of day eight (**Figure 3-4a, Figure 3-4B**). At 0.25 MOI the decrease in plaque forming units (PFU) in the media characteristic of single cycle CMV growth was not observed at two days post inoculation with any virus (**Figure 3-4C**), indicating an oversaturation with inoculum. Despite the common observation that BAC integration slows CMV viral growth in previous studies

[401, 481], measured values for wildtype CyCMV Mauritius and both CyCMV Mauritius BAC clone titers were not distinguishable beyond standard error at day eight or day ten.

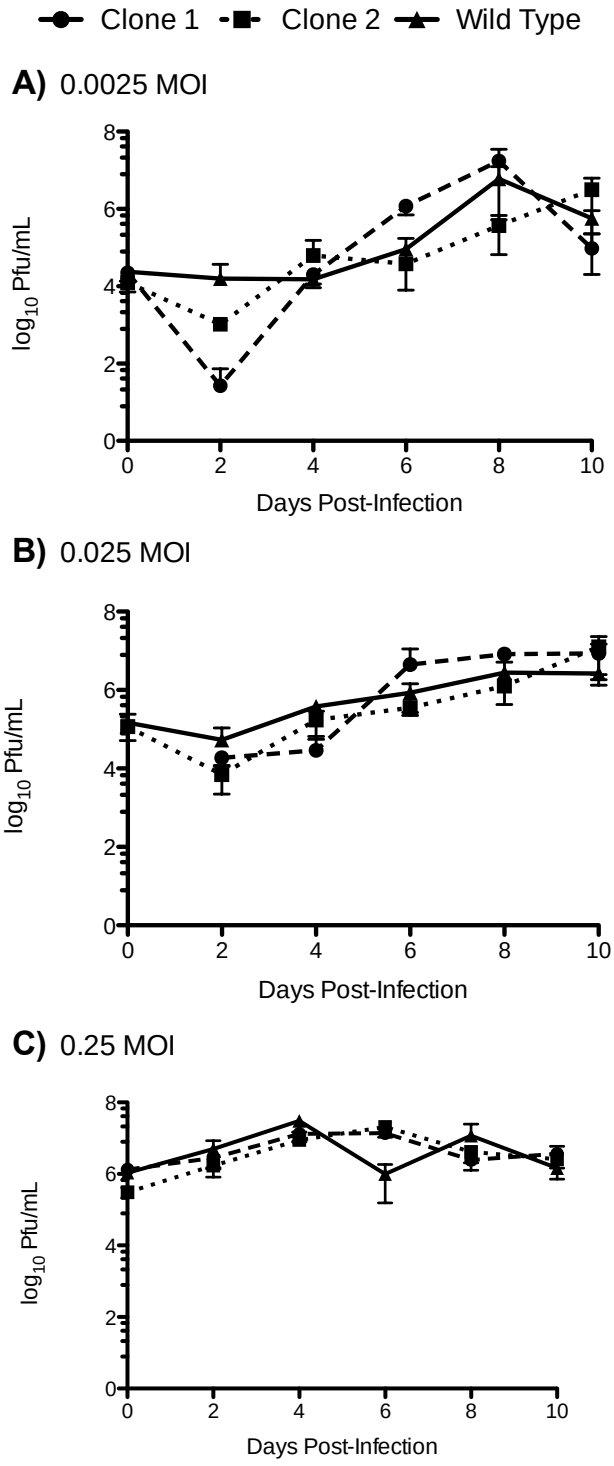


Figure 3-4: Single-cycle growth curve of CyCMV Mauritius BAC clones compared to wildtype CyCMV strain at three MOI. CyCMV Mauritius BAC and clones were inoculated into Telo-RF cells at 0.0025 MOI (A) 0.025 MOI (B) and 0.25 MOI (C). Viral cultures were stopped at two, four, six, eight and ten days in duplicate and titred. Similar growth kinetics were observed between all viruses. Bars represent standard error.

3.5.4 Expression of SIVmac239 derived GPF and NTR Transgenes

One structural SIV vaccine construct, a Gag-Pol fusion (GPF), and one non-structural SIV vaccine construct, a Nef-Tat-Rev fusion (NTR) were integrated by BAC recombination into CyCMV Mauritius BAC Clone 1. Both constructs were codon-optimized for expression and modified to remove SIV gene functionality (**Table 3-2**). To avoid difficulties in BAC recombination GPF and NTR transgenes were both placed behind an EF1α promoter, as was used with RhCMV 68-1 and RhCMV 68-1.2 [185, 187, 200, 244]

Table 3-2: Inactivating mutations in SIV vaccine construct

SIV ORF	Mutation Location	Mutation Function/Domain
Gag	G2A	Delete Myristoylation Signal
	ΔFS	Remove Gag-Pol Ribosomal Frameshift
Pol	D25A	Inactivate Protease
	H12A / H16A C40A / C43A	Disrupt Integrase Zinc-Binding HHCC Domain
	D64A / D116A / E152A	Disrupt Integrase Catalytic Domain
Nef	G2A	Delete Myristoylation Signal
	E93	Remove Premature Stop Codon
	D143G / LM194AA	Inactivate Nef
Tat	C21G	Inactivate Tat
	Δ80-84 (Arg-rich region)	Disrupt RNA Binding and Nuclear Localization
	Juxtapose AA1-79 and AA85-129	Disrupt Tertiary Structure
Rev	KR39DL	Inactivating mutation
	Δ42-46	Disrupt RNA Binding and Nuclear Localization
	IE81DL	Inactivating mutations

GPF and NTR transgenes were integrated into CyCMV Mauritius BAC Clone 1 between *CyUS3* and *Cy215* by BAC recombination. Positive bacterial colonies were selected based on colony PCR, and Sanger sequencing was used to confirm the location and integrity of the transgenes. The presence of each transgene was confirmed within whole cell lysate of infected tissue culture. Expression of each transgene was strong, at similar levels of expression to those seen using the VZV-SIV constructs (**Figure 3-5)**. The macaque CMV IE-1 gene was used as a housekeeping gene to quantitate transgene expression relative to CMV gene production. Both GPF (**Figure 3-5A**) and NTR (**Figure 3-5B**) proteins blotted at their expected size.

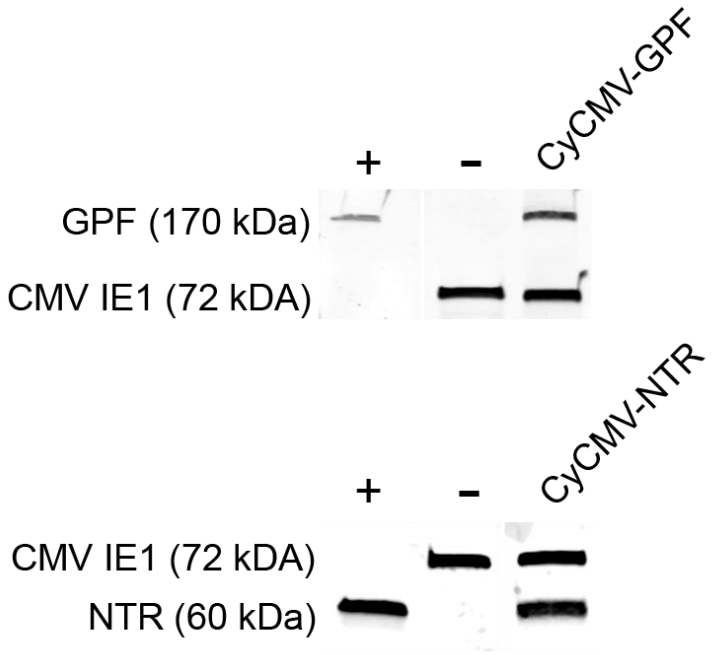


Figure 3-5: Expression of SIV transgene constructs by CyCMV Mauritius BAC in cell culture. CyCMV Mauritius BAC-GPF and CyCMV Mauritius BAC-NTR were grown and passaged as virus in Telo-RF cell culture. Cells were harvested when greater than 80% of the cell monolayer showed CMV cytopathic effect. 5uG whole protein lysate were added per well of VZV_{Oka} expressing GPF or NTR (positive control), CyCMV Mauritius BAC-GPF, or CyCMV Mauritius BAC-NTR to a western blot. Blots were stained against CMV IE-1 as a housekeeping gene, and SIV Gag (A) or SIV Nef (B).

3.6 Discussion

Macaque monkeys, including rhesus macaques and cynomolgus macaques, are important models in biomedical research. The increased usage of cynomolgus macaques has highlighted important differences between the two as research models of infectious diseases [428, 494] (See Section 1.4.2.2). Even within each species the differences between Indian and Chinese rhesus macaques, or Filipino and Mauritian cynomolgus macaques, can be meaningful for research outcomes [428, 430]. As part of a plan to compare CMV to VZV as vaccine vectors in the cynomolgus macaque model and in light of recent findings supporting the use of CMV as a vaccine vector [185, 187, 200, 244, 391-393] we generated a CMV BAC for use within the cynomolgus macaque model.

Our group has previously isolated and characterized two strains of CyCMV, CyCMV Ottawa from macaques of Filipino origin [289], and CyCMV Mauritius from macaques of Mauritian origin [278] (See Chapter 2). CMVs are large, slow growing viruses that are difficult to work with by traditional molecular means (See Section 1.4.4.1). The construction of a CyCMV Mauritius BAC preserves the CyCMV Mauritius genome, preventing the accumulation of mutations throughout repeated laboratory passaging, and facilitates the use of BAC recombination for powerful site-specific alterations. Without a CyCMV BAC or another equivalent molecular tool, detailed studies into CyCMV and CyCMV as a vaccine vector are not practical within the cynomolgus macaque model.

3.6.1 Construction of CyCMV Mauritius as a BAC

As a BAC, CyCMV Mauritius can be preserved, faithfully replicated, and modified with high precision. To minimize the chance of polymorphisms or deletions during BAC construction, a site was chosen to insert the BAC vector between ORFs *CyUS1* and *CyUS2* near the more stable center of the linear CyCMV Mauritius virus (**Figure 3-1a**). Instead the BAC construct was found inserted 6kbp upstream of the selected site due to the interaction of Cre recombinase with a nearby cryptic *loxP* site within the *CyUL153* ORF (**Figure 3-1b**). Though RhCMV 68-1 and other HCMV strains constructed as BACs contain multiple cryptic *loxP* sites, this is the first time this phenomenon has been described in the literature.

Structurally similar to *loxP* but differing by more than 5 mismatches, cryptic *loxP* sites can interact with Cre to introduce nicks, mutations, or rarely to recombination with other similar *loxP* sites [540]. CyCMV Mauritius contains a cryptic *loxP* site with high similarity to the CyCMV Mauritius BAC *loxP* sequences at the site of insertion in the *CyUL153* ORF (**Figure 3-2**).

The transfer plasmid utilized in the original homologous recombination was prepared from the pWC132 plasmid used to construct RhCMV 68-1 BAC [401]. However, the 1445 bp *CyUS1* to *CyUS2* homology region needed for this plasmid was cloned from CyCMV Ottawa and not CyCMV Mauritius due to the availability of genetic material. CyCMV Ottawa, isolated from a Filipino

cynomolgus macaque, is closely related to CyCMV Mauritius [278, 289]. Across the *CyUS1-CyUS2* region the consensus sequences of the two viruses are 92.5% identical, differing by a 55bps gap present in CyCMV Mauritius near the start of *CyUS1* relative to CyCMV Ottawa and several SNPs. While it is possible that differences between the CyCMV Ottawa sequence utilized and CyCMV Mauritius may have been factors affecting the stability of the insert – including the *loxP* sites – they did not prevent the initial integration of *eGFP* within CyCMV Mauritius.

The plaque purified CyCMV-eGFP construct upon sequencing was found to contain a SNP in one of the two eGFP-flanking *loxP* sites (**Figure 3-2**). This alteration may have been caused during recombination or through viral replication. Regions of the viral genome immediately surrounding inserted extraneous DNA are known to be unstable [486]. Regardless of this fact, essentially a genetic bottleneck was introduced throughout the plaque purification process by the serial isolation of individual viral plaques and the C to T transversion was selected. This SNP did not prevent subsequent Cre-*loxP* recombination to excise the *eGFP*, but a second SNP was observed in the single *loxP* site of the CyCMV Mauritius-LoxP intermediate (**Figure 3-2**). It is possible that further mutation of the *loxP* site was facilitated by the presence of this previous mutation.

Though a number of *loxP* variants are utilized in cloning to favor integration or excision of DNA [538, 539], the same *loxP* 511 sequence was utilized in all cloning plasmids. This sequence has a reduced recombination efficiency with wildtype *loxP* and increased efficiency with other *loxP* 511 sequences [539, 541]. Cre mediated recombination efficiency drops between non-matching *loxP* and single or double stranded nicks accumulate during abortive recombination events [538-540, 542]. This mechanism may have contributed to the accumulation of further mutations in the *loxP* sites throughout (**Figure 3-2**).

Mutations that increase the similarity of the *loxP* between *CyUS1* and *CyUS2* with the *CyUL153* cryptic *loxP* site increased the efficiency of recombination between the sites [540]. If the BAC inserted first at the intended *loxP* site and subsequently moved, this may have been facilitated by the SNPs. However since polymorphisms on the BAC flanking *loxP* sites were predominantly restricted

to the DNA originating from CyCMV Mauritius-LoxP, it can be assumed most SNPs accumulated on *loxP* sites in the virus prior to integration of the BAC, and it can be speculated that the BAC integrated directly into the cryptic *loxP* site. As identical alterations were seen in both *loxP* sites of CyCMV Mauritius BAC, changes to the BAC flanking *loxP* sites occurred prior to plaque purification where the two clones were separated.

The BAC vector facilitates stable replication of the viral genome within bacteria but is unnecessary during subsequent replication as a virus and, because of its size, can cause instability in the surrounding DNA [486]. It should therefore be removed immediately upon reconstitution of the virus. Removal of the BAC vector increases in importance with the introduction of viral transgenes into the CyCMV Mauritius genome. The CMV viral capsid has a total packaging capacity of around 230kb [286]. CyCMV Mauritius is 217kb, leaving little room for transgenes while the 9kb BAC vector remains inserted. For both these reasons two *loxP* sites border the BAC vector to facilitate its removal early in viral growth by Cre/*loxP* recombination [481, 489].

Regardless of the cause, the *loxP* sites on both CyCMV Mauritius BAC clones were not capable of efficient Cre mediated recombination. With the exception of a single mutation known to increase the efficiency of Cre-*loxP* recombination, the induced mutations have all been demonstrated to decrease Cre-*loxP* recombination efficiency [538, 539] One mutation, the central T to A, is expected prevent recombination entirely [538-540, 543]. A substantial benefit of a CyCMV Mauritius BAC is the relative ease of site-specific modification compared to CyCMV Mauritius virus. BAC recombination was utilized to correct the *loxP* sites to the initially intended *loxP* 511 sequence and repair the ability for Cre-*loxP* mediated excision of the BAC vector during viral growth.

3.6.2 Fitness of CyCMV Mauritius BAC:

Due to the multiple plaque purification steps necessary to produce a BAC, some degree of attenuation may have resulted from the mutations occurring throughout the CyCMV Mauritius BAC genomes with the notable exception of a relative paucity of mutations in genes coding for core

proteins reflecting the increased functional importance of these genes. (**Figure 3-3; Table 3-1**). Attenuation defined as a shift towards a less fit or less virulent state is usually observed when viral culture occurs in novel circumstances such as repeated *in vitro* passaging. With CMV significant attenuation is often observed in laboratory strains of simian and human CMV as a result of repeated passaging in fibroblast cell culture [288, 544]. In particular the mutation to a non-functional state or deletion of genes in the *UL128-UL131* region is observed in most laboratory CMV strains passaged on fibroblast cell lines [517]. Following the isolation of RhCMV 68-1, there was both an inversion and a deletion in this region of the virus that resulted in the removal of *UL128* and two-thirds of *UL130* [200, 288, 402]. In the CyCMV Mauritius BAC all three genes, *UL128*, *UL130* and *UL131A*, are present but only 70% nucleotide pairwise identity was preserved across the *UL128-UL131* region relative to CyCMV Mauritius. Based on predicted amino acid sequences the accumulated mutations result in 62.9%, 52.3% and 61% amino acid identity respectively with the three genes of wildtype CyCMV Mauritius.

Despite missing *UL128* and most of *UL130*, RhCMV 68-1 is still able to infect its native host, rhesus macaque monkeys, including some rhesus endothelial and epithelial cell lines [379, 403] (See section 1.3.2.1). However, the virus is unable to infect cynomolgus macaques [291, 379]. It was recently shown that repair of the *UL128-UL131* region in the derivative RhCMV 68-1.2 BAC facilitated viral replication within cynomolgus macaques, as did the repair of the *UL36* homologue [379]. Repair of *UL36*, without the restoration of the pentameric complex, facilitates the infection of kidney epithelial cells in both rhesus and cynomolgus macaques, resulting in viral shedding through urine [379]. The CyCMV *UL36* homologue (*CyUL36*), a putative inhibitor of caspase-8 induced apoptosis [545] is present and unaltered in both CyCMV Mauritius BAC clones.

Only two single amino acid changes occurred in the core genes. *CyUL102* and *CyUL47*, contained conserved changes in single amino acids unlikely to affect gene function. Core genes show the highest level of conservation between macaque CMV strains [278]. Members of this gene family have key functions and selective pressure is maintained throughout growth in cell culture. By contrast Non-

Core family genes, which are unclassified into other families, and RL11 family genes, which encode transmembrane glycoproteins, are both highly variable between CMV strains [282, 546]. The majority of polymorphisms in the CyCMV Mauritius BAC clones occurred in genes of these families.

The genes *CyUL146* and *CyUL147*, which are homologous to the host gene *CXCL5*, were heavily mutated and may no longer be functional. Both genes are non-essential to growth *in vitro* and therefore lacked selective pressure during passaging [519]. *In vivo* the genes are involved in host immune modulation but each gene is absent in several other macaque CMV strains. Notably RhCMV 68-1 lacks a *UL146* homologue, RhCMV 180.92 lacks homologues to both *UL146* and *UL147*, and wildtype strain RhCMV 19262 is truncated in both genes [379].

Both RhCMV strains contain a homologue of COX-2. This gene is absent from CyCMV Ottawa [278, 291] but present in CyCMV Mauritius. The gene acquired mutations, though CyCMV Mauritius did retain all copies of *CyUS28*, including a homologue of HCMV *US28* (*CyUS28e*) that induces host expression of the COX-2 pathway [387, 526].

After categorizing the mutations introduced through the creation of the CyCMV Mauritius BAC clones it remained important to characterize their fitness as viruses. To do this CyCMV Mauritius was reconstituted as a virus and passaged in multiple cell types.

Natural CMV infection is restricted to the native CMV host species [268]. This restriction occurs after viral entry, which is unaffected in cell culture by host cell origin, though productive replication is impaired between most divergent hosts [268, 291]. There is some ability, however, for primate CMVs to move between hosts with limited infection [268, 294]

RhCMV 68-1 is unable to infect cynomolgus macaques [291]. It has been suggested that this is linked to the absence of a *UL36* homologue, and a missing *UL128-UL131* region [379]. Despite this RhCMV 68-1 is able to grow in cynomolgus macaque fibroblast cell lines [379], and if the *UL128-UL131* region is repaired, is also able to grow in human cell lines[484]. CyCMV Mauritius by contrast required no repair to grow on a variety of primate cell lines. Wildtype CyCMV Mauritius grows and proliferates in human [407], rhesus macaque[534], and cynomolgus macaque derived

fibroblast cell lines [534]. Growth of wildtype CMV is comparable in all three cell lines, though growth is faster in rhesus macaque derived Telo-RF cells due to the faster rate of cell proliferation [534].

Both CyCMV Mauritius BAC clones maintained this breadth of *in vitro* host range. CyCMV Mauritius BAC clones were grown and maintained on Cynomolgus macaque derived MSF-T, Rhesus macaque-derived Telo-RF and human-derived MRC-5 cell lines. CMV cytopathic effect was observed after infection, and virus was successfully passaged multiple times without a decrease in viral potency. Additionally, growth rates of CyCMV Mauritius BAC clones did not differ significantly from that of wildtype CyCMV Mauritius (**Figure 3-4**).

We can conclude that viral fitness was maintained for both of the CyCMV Mauritius BAC clones with the sole caveat that these assays were all performed on fibroblast cell lines. Based on the attenuation in the *UL128-UL131* region of both CyCMV Mauritius BAC clones it is likely that they, are to some extent fibroblast adapted [200]. Whether the two CyCMV BAC clones can generate virus with the ability to replicate on non-fibroblast cell lines, such as endothelial or epithelial cell lines, has not been tested due to the unavailability of such lines. The relative ease with which one can use BAC recombination to manipulate the CyCMV Mauritius BAC allows for the possibility of correcting regions of the virus analogous to the work of B.J. Burwitz et al. who reinserted *UL128-UL131* to generate RhCMV 68 1.2 [379].

3.6.3 CyCMV Mauritius BAC as a SIV Vaccine Vector:

In the natural host, CMV infection results in lifelong intermittent and frequent low-level viremia [309, 312]. The resulting adaptive immune response is robust. In healthy primates 4-5% of all peripheral T cell populations and 10% of the memory T cell population is directed against CMV antigens[312, 342]. Notably, unlike other chronic diseases such as HIV or Hepatitis C, the immune response to CMV is not exhausted but persists in the host [340, 341]. CMV and other ancient herpesviruses establish symbiotic relationships with their host where continued persistent or

reactivating infection occurs with little pathology [257, 258]. The host is unable to eliminate CMV but is able to prevent high level viremia, organ damage, and central nervous system infection [258, 309].

Increasingly CMV is thought of as a potential vaccine vector. The host response induced by CMV is that desired for an effective vaccine: vigorous CD8+ effector memory responses, like those elicited by CMV, have been correlated with protection against infection in the yellow fever and smallpox vaccines [220] and are thought likely to be protective against HIV [185, 388], TB [388, 547], Hepatitis C [388, 548], and various cancers [389]. The potential for a T effector memory cell-based vaccine against SIV, the simian equivalent of HIV, has been demonstrated previously using similar herpesvirus vectors; VZV [345, 549], rhesus rhadinovirus [399], and RhCMV [187, 200, 244] all provided partial protection, a reduction in set-point viral load, or clearing of detectable virus in some animals. Our group has previously looked at protective efficacy of a human VZV vectored vaccine containing two engineered SIV transgenes against SIV in cynomolgus macaques [345].

One application of the CyCMV Mauritius BAC is the construction of a CMV vectored vaccine candidate for use in cynomolgus macaques. In this pursuit our group introduced the two SIV transgene constructs into separate CyCMV Mauritius BAC constructs.

GPF and NTR were expressed in CyCMV Mauritius BAC-GPF and CyCMV Mauritius BAC-NTR infected cell cultures respectively (**Figure 3-5**). Expression of each transgene was strong, at similar levels of expression to those seen using the VZV-SIV constructs (**Figure 3-5**). Both constructs were stable in CyCMV Mauritius BAC derived virus and continued to express after viral passaging. None-the-less as there is no selective pressure within CMV to maintain expression of either transgene, repeated passaging in cell culture should be minimized. There remains a need to examine the CyCMV vector growth characteristics and long-term expression of the transgenes in the cynomolgus macaque model and to characterize the immune response generated.

Animal models of CMV have been developed as vectors for candidate vaccines against tuberculosis [390], Ebola virus [391-393], herpes simplex virus-1 [394], and SIV [185, 187, 200,

244]. Most relevant to this work, trials utilizing RhCMV 68-1-based SIV vaccines curiously protected only half vaccinated animals from disease [185, 187, 200, 244]. As the correlates of protection remain unknown in these studies, it is important to recall that protection observed in rhesus macaques does not always correspond to protection in humans, as was the case with AD5 vectored SIV vaccines [32, 46, 247-253].

The RhCMV 68-1 based SIV vaccine exhibited a number of anomalies arising from attenuated genes within the virus, including CD8+ T cell targeting of MHC-II and MHC-E epitopes [200, 245, 385]. These are not predictive of protection as the responses occur in both protected and unprotected animals, but demonstrate the impact subtle differences in the CMV genome can have on the vaccine phenotype.. The CyCMV Mauritius BAC presented here will provide the first opportunity to study CMV as a vaccine vector in macaques that does not rely on RhCMV 68-1. This is particularly important since there remains a number of CMV genes for which the function is unknown, and the impact of mutations are not always appreciated Comparisons between the immune response to RhCMV and this vector will inform the selection and the attenuation strategy of a future human CMV vector. CyCMV Mauritius BAC similarly expands the field of research to include cynomolgus macaques, which though similar to rhesus macaques respond differently to a number of infectious diseases including SIV [493, 494].

In conclusion CyCMV Mauritius has been cloned as a full-length BAC capable of Cre mediated excision. The genome underwent minor attenuation as a result of the cloning process, but few key genes were disrupted. CyCMV Mauritius BAC was successfully reconstituted as a virus and retained wildtype growth kinetics in multiple cell lines. Moving forward CyCMV Mauritius BAC is an ideal tool alongside RhCMV 68-1.2 to investigate and understand the mechanisms of specificity within CMVs of closely related species. The newly synthesized BAC has a large viral packaging capacity, and with the removal of the BAC vector by Cre/*loxP* recombination early in viral growth, is capable of facilitating and expressing large transgenes, as demonstrated by CyCMV Mauritius BAC-GPF and CyCMV Mauritius BAC-NTR. CyCMV Mauritius expands the examination of CMV as a vaccine

vector into the cynomolgus macaque model where it can be compared directly to other vectors including human VZV_{oka}, which replicates in cynomolgus macaques but not rhesus macaques [475, 550, 551]. The CyCMV Mauritius BAC facilitates a meticulous approach to the study and utilization of CMV in the increasingly utilized and currently underdeveloped cynomolgus macaque animal model.

3.7 Acknowledgements

This work was supported by the Canadian Institutes for Health Research (CIHR, grant number 107743), a CIHR Doctoral research award to A.K.M, an Ontario HIV Treatment Network (OHTN) Junior Investigator Development Award to D.O.W, an OHTN postdoctoral fellowship to A.P.A, and an OHTN Career Scientist Award to K.S.M. Dr. Kelly S. MacDonald is the recipient of the H.G Sellers Endowed Research Chair of the University of Manitoba.

3.8 Supplementary Materials

Supplementary Materials 1: Synthesized DNA fragments utilized to correct the upstream *loxP* site, 1968bp long

ATCGAGATCTTCCGAAACCAATACCGCACGATATTTATGAACTTGAGGCGTATACACAGCACTGCTAGTCAGTT
CGGTTTCCACAAACACCTTTTCGGTAGTTGCACAGACAGGACCGCATAACACAGTAGGCCTCGGAGGAGCGACC
TGCACTTCAGTTTTATCATCCTGATTGTTGTTAGTACTCGGGTCCCAGGTGTTTACCACCTCCCACTCCACTCT
CTGGTTCGCTTCATCTTCTTGGTGTCTTCGGCTCCAATGCGATCGCCGCCGTACTTCGTAACGATCTTTTTGAC
AACAATGGCTGCAGTCGTGTCGGCACCTGCACCATGGTCCGCAGTCACAGCACACCTCGTAACACAGTCTGCAG
GTGAACGGCCAGTCGTAAATCAGGTAGATGAGGCCCACTAGCGATGTCATTCCGCCGGCTCCACATACCAGCAC
GCCAAGAGGCCACATGTCTAAAACACTCCCAGCAGAATTGGTGTTACCCGGCGATGTAGCGTTCATGTTGCGGT
AGCTGCTCGTCGTCATCCCTATATTTCGGCGGTAACTCACTTTTCTGGTTCTGTAGGTTCAGACGAACTATGAT
GAAATATACATTTGCTCGTAACCAGTTGTCAATCTTCTTCTCTCCGGACGATATTTGACAATTAGCTCTTCTCT
GTCGTCCGAATGGCGCTGTCGGTAGTACCTCATTCTTCTCATTCTTCCCTGTCCAAACATTAATGCCACGACAA
CGGTAATAACTAGTGCCAAGGCCCACACGGCATGCGTAGCTGGGGCTGCAACATCTTGAAGGTACACATTAACG
GTAGTGCTAGATGTTGGTGGTGCAGAACTAACAAACGTACAAAAGTCAGGTGTACTGTTGGGATAGCTAGTGCC
GTTTCCAGATGTAACTGTAAGATCAAAGGCCTCTGGTTTTGGACTGTTGTCTTGTGTCTTTAGAACATAACTTC
GTATAGTATACATTATACGAAGTTATATTCGATGCGGCCGCAAGGGGTTCGCGTCAGCGGGTGTTGGCGGGTGT
CGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCA
CAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGAT

CGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACG
CCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATT
CGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTAAGATACATTGATGAGTTTGGA
CAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAC
CATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGT
GGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCTAGAGTCGCGGCCGCTT
TACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTGATCGCG
CTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAGCACGGGGCCGT
CGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCTTG
AAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAG
CTTGTGCCCCAGGATGTTGCCGTCCTGCTAGCGCGGCCGCAGAT

Supplementary Materials 2: Synthesized DNA fragments utilized to correct the downstream *loxP* site, 1845bp long

AGTTGCGGCCGCTAGCAGTCTGGGACCACGGTCCCACTTGTATTGTCGATCAGACTATCAGCGTGAGACTACGA
TTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTCGTAACCTGTAGAACGGAGTAACCTCGGTGTGCG
GTTGTATGCCTGCTGTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGCGCACGGTTATGTGGACAAAA
TACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCAAGTGTGTCGCTGTCGACGAGCT
CGCGAGCTCGGACATGAGGTTGCCCCGTATTCAGTGTCGCTGATTTGTATTGTCTGAAGTTGTTTTTACGTTAA
GTTGATGCAGATCAATTAATACGATACCTGCGTCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCAC
GTTGTGATATGTAGATGATAATCATTATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCG
ACCTCGCGGGTTTTCGCTATTTATGAAAATTTTCCGGTTTAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAAT
GTTTTTATTTAAAATACCCTCTGAAAAGAAAGGAAACGACAGGTGCTGAAAGCGAGCTTTTTGGCCTCTGTCGT
TTCCTTTCTCTGTTTTTGTCCGTGGAATGAACAATGGAAGTCCGAGCTCATCGCTAATAACTTCGTATAGTATA
CATTATACGAAGTTATAACATTAAGGTGGACTTACTACAATTATAGCAAAGGGTACCGTAATGATGTTTTGTCT
GATGAGAACCTGACGTTCTGCAAAGTTCATGTTCAGTTGGTTTTCCAGTCTTGTTATGCATTAAAACCCAATAT
GTATGATGATGAGTGGTAATAGAAAAATTAGTTGCCGTCAGGTTTACAGTTTCTCCAACAGTCGCATTAACTTT
TACTAATGTGTATCCGAGAACGGCAGCTGTGGATGTACGAGTTGTGGTAGTATTTACAGTACTGTTGATTGAAG
TTATGTTAATGTTATGTGTACTCGTTGGTGTAGTAGTAATAGCTGAACTTGATGTGGTAATGGTTGTGTTTCCA
CTTGTATTAGTAAATATTTCGCTACTATTTGAAGATGTAACGGTAGCTGCAGATGAAGTAGATATATTTGTTGT
AATTGAGGTATCCGTACTAGATGTATTTACAGTACCGGATGATGTTACTGTATTCTGAGTTTGAGTGGGTGTTG
ATACGTTTGCGGTGGTTTCTGAAGTGGATAGCGCTGTACTGGTTATTGTAGATGTTACACTCATATTGGTAGTG
GATATAGGTGATGTGCTTATGAAGGTAGATGTTGCATTAGTAGTAGTTGCTACTGTTGATGTATTGTTTGCAGC
GGCATAACGAGAAAGTATGGCGATGAAAATGTAGATAGTACGATCCTTTATAAGGTAGCTTATTCTGAAATGCA
TATTGGAATTTTATTCACACGCCGTCTTGATGTTTCGACGTACGTGTCTTGTAGTTTCACAGCCGTTACACCTA
AGTGATCATGACTCCGTCTGGAACACCTTTATAGGTCGTTGGTACAATAAAATGCGTCATTGCTACGTGCTTCA

GATTTCCACTTGGATGTTCTTCAGGTTGAAAGTTATTTAAAAACTTGGAGCCTTGCGGGTTCTAATAACATACA
TGAGTTAGATAATATAATACATGCAAATAAATATTTGCTAATGCATAATAAAACTGAATACATACTCTAAGAAT
TAGGATAGTTTTGTATATTCTGGATGTCGTACCGATCCATGTCAAATCTCTAGAGAAGATCTAGACAGT

Chapter Four: Discussion and Future Directions

4.1 CyCMV Mauritius is a Necessary Model of HCMV

As an overall objective, this thesis set out to prepare a Mauritius strain of CyCMV as a vaccine vector for use in cynomolgus macaques in order to model the use of HCMV vectored vaccines against HIV. The use of CyCMV Mauritius-GPF and CyCMV Mauritius-NTR as vaccines against SIV is intended to inform the design of future HCMV vectored HIV vaccines. It is however, important to recall that by necessity all research of this type utilizes a triple removed model with a simian viral vector expressing SIV transgenes in a simian host. This has implications for the inferences that can be drawn, and makes the use of concurrent simian models to predict results especially useful.

First, in the case of the immunogen, SIV is not the same as HIV. Though closely related, and in many ways analogous, the two retroviruses possess distinct differences at the level of genomics, proteomics, and pathogenicity (See Section 1.4.3). Careful consideration of the appropriate SIV strain for both transgene and challenge virus is required, taking into account co-receptor utilization, tropism, neutralizing capacity, heterogeneity of challenge, and relative virulence of the virus. Furthermore, comparison to HIV in the context of these differences must accompany any non-human primate vaccine trial.

Second, the analogue of the human vector, macaque CMV, must be taken into account with respect to its similarities and differences to HCMV. The species specificity of CMVs preclude not only the direct testing of HCMV in a macaque model, and the use of CyCMV or RhCMV as a vector for a HIV vaccine in humans, but also limit the examination of CyCMV and RhCMV in rhesus and cynomolgus macaques respectively [291, 379] (See Section 1.4.2.2). Consequently the studies presented here, the immunogenicity trial to come (See Section 4.5.1), and a proposed SIV challenged trial in the cynomolgus macaque model (See Section 4.5.2) will serve to support or reject the existing data from the RhCMV 68-1-rhesus macaque model system in order to more confidently move towards the development of HCMV as a vector in HIV and other vaccines.

At present, there remain several substantial unanswered questions regarding the nature of protection in the RhCMV 68-1-rhesus macaque model. While only half of macaques were protected from SIV disease progression across a number of studies [185, 187, 244], if such protection could be extrapolated to humans with confidence this would be sufficient to significantly alter the course of the HIV epidemic and could potentially be further augmented in combination with other approaches (See Section 1.1.1.3). However, since the mechanism and correlates of protection with RhCMV 68-1 based

vaccines are not known, such confidence is premature until either a mechanism is demonstrated or human data is forthcoming.

T_{EM} responses have long been suspected as the mechanism of protection in herpesvirus based vaccines [185, 275, 345, 392, 393, 552], but recently it has been suggested that protection against SIV from an RhCMV 68-1 vectored vaccine may be mediated by CD8+ T cells responding to MHC-II and MHC-E presented peptides [200, 245]. It is argued that while MHC-I is downregulated during both CMV and SIV infection, MHC-E is reciprocally unregulated providing an alternative channel for CTL targeting of infected cells [245, 404]. If this is true it presents problems extrapolating from the RhCMV 68-1 model to a HCMV vectored HIV vaccine. It is not known how to induce CD8+ T cell targeting of MHC-II and MHC-E in HCMV. Recently it has been shown that the proposed mechanism, attenuation in the genes of the pentameric complex [200, 245], does not result in a similar phenotype in HCMV strains tested to date [386]. Additionally MHC-E heterogeneity in the rhesus macaque model is far greater than that observed in either humans or cynomolgus macaques [404, 438-440]. MHC-II and MHC-E targeting, which are observed in both protected and unprotected RhCMV 68-1 vaccinated macaques alike, may be the additive result of any of a combination of differences between RhCMV 68-1 and other CMV strains, or simply a unique feature of the rhesus macaque model itself.

To reiterate the key differences, unlike RhCMV 68-1, which has been passaged as a laboratory strain since 1968 [400], CyCMV Mauritius is as close to a primary isolate as can be achieved. Overall CyCMV Mauritius has fewer deleted and altered genes than RhCMV 68-1 when compared to wildtype macaque CMV isolates or HCMV [278, 379]. This does not inherently mean CyCMV Mauritius will make a better model of HCMV. However, CyCMV Mauritius will allow us to compare the protective efficacy of a vaccine made from a closer to wildtype CMV virus to a highly passaged virus like RhCMV 68-1, which we know has resulted in a narrowed cell tropism, abnormal MHC restriction and other functional losses in RhCMV 68-1 as a result of numerous gene deletions and inversions [200, 245, 288, 379].

To apply either macaque-macaque CMV vaccine model to HCMV it is necessary to understand the mechanism of protection observed with macaque CMV based vaccines, whether that protection is linked to T_{EM}/T_{RM} responses, MHC-II and MHC-E targeting CD8+ T cells, or other factors. A HCMV vaccine vector will need to be selected, and constructed, to ensure it provides similar protection against HIV in humans. If the protection generated by RhCMV 68-1 based SIV vaccines is due to oddities of the virus [200, 245], HCMV should be adapted to incorporate RhCMV 68-1 like attenuations, if possible. If protection is instead mediated by T_{EM}/T_{RM} as originally suspected, there will be greater breadth available in constructing a HCMV vectored HIV vaccine. Thus CyCMV

Mauritius facilitates the exploration of viral determinants of protection in both a second virus, and a second animal model.

The macaque challenge model, the third variable, is not the same as natural infection in humans, in neither the artificial challenge nor the biology of the model. Though perhaps the closest analogue of the human immune system outside of the apes, at least 25 million years of evolution has shaped immunologic, morphologic, and genetic distinctions between humans and macaques [425-427] (See section 1.4). In addition, the circumstances of challenge even when a multi-low dose mucosal challenge is used do not replicate natural human infection. Without the presence of additional risk factors such as sexually transmitted infections, sexual HIV transmission in humans occurs via cell bound or free virus [29, 115] only after hundreds of exposures on average [154, 157, 158] (See Section 1.1.1.3), a rate that cannot be practically replicated in a macaque challenge.

In retrospect it may not be surprising that even while initial macaque-SIV vaccine challenge trials using an AD5 vectored SIV vaccine showed complete [247, 248], or partial protection [249-252], the subsequent human trials did not [32, 46, 253]. After the failure of the STEP trials, further macaque studies designed to examine the narrow breadth of CD8+ T cell responses in human vaccine recipients and increased susceptibility of previously AD5 seropositive individuals were able to replicate the human results [254, 553]. The lesson learned from these experiments is therefore not only that data from a single animal model does not always predict the human outcome, but also that relying on a single measured outcome without an understanding of mechanism of protection is not prudent. It has become accepted that in addition to protective efficacy, secondary correlates of immune protection must be considered, and additional controls that allow the elucidation of vector effects, differing routes of exposure and dose of challenge are worth examining before extrapolation to humans can be confidently made.

Cynomolgus macaques differ from rhesus macaques by 1.2 million years of evolution, producing as overall genetic divergence of 0.4% between the species [426] (See Section 1.4.2.2). These changes are sufficient to prevent cross infection of cynomolgus macaques with RhCMV 68-1 [291, 379], and result in lower peak and set-point viral loads in cynomolgus macaques after SIV challenge [437, 463]. Because genes possibly relevant to protection from SIV differ between rhesus and cynomolgus macaques, including TRIM5 [426, 443] and MHC-E [404], there is benefit in comparing CMV vaccine vectors in the two models.

Importantly, cynomolgus macaques, and in particular Mauritian cynomolgus macaques, are available to a number of researchers who do not have access to rhesus macaques. Since cynomolgus macaques and not rhesus macaques are susceptible to infection with human VZV_{Oka} [475, 550, 551], the herpesvirus utilized by our group as the vector in a similar SIV vaccine trial [345], this includes

the possibility of synergistic vaccine trials including CyCMV Mauritius as one of multiple herpesvirus vaccine vectors.

Through the investigations of CMV vectored vaccines using the cynomolgus macaque and CyCMV models, the development of CyCMV Mauritius-GPF and CyCMV Mauritius-NTR facilitates the further exploration of HCMV vectored vaccines against HIV in a manner not previously possible.

4.2Development of CyCMV Mauritius as a BAC

A key development of the work presented in this thesis is the completion of CyCMV Mauritius as a BAC. CyCMV Mauritius BAC is of substantial value to the scientific community beyond just its use as a SIV vaccine vector. This macaque model of CMV is also key to the study of congenital CMV and vaccination against the virus itself (see Section 1.4.4). The CyCMV Mauritius BAC provides the ability for repeated precise manipulations of a minimally attenuated CMV genome that can be used within the cynomolgus macaque model. Unlike with other editing techniques such as CRISPR/Cas9, as a BAC any number of genes can be knocked out, or transgenes inserted through sequential BAC recombination including those inhibitory to viral replication (See section 1.4.4.1). The effects of multiple genes on CMV growth, immunogenicity, or virulence can be studied with relative ease in cynomolgus macaques using CyCMV Mauritius BAC.

4.2.1 Construction of CyCMV Mauritius BAC

It is possible that CyCMV Mauritius will be one of the last CMV viruses cloned as a BAC by the methodology utilized in this thesis. CyCMV Mauritius was produced by a single homologous recombination to introduce a floxed *eGFP* into the virus, followed by a series of sequential Cre/*loxP* recombination steps to insert the BAC DNA (See Section 3.4). This process relied on three stages of plaque purification, two consisting of five rounds and one of two rounds. Each round of plaque purification introduced a genetic bottleneck that could select for mutations in the virus. A number of polymorphisms occurred within ORFs of CyCMV Mauritius due to this repeated plaque purification (See Section 3.5.2).

Instead of the method utilized in this paper, CRISPR/Cas9 aided recombination has recently been used to incorporate the BAC vector directly into the full genome of the aphaherpesvirus pseudorabies in a single recombination step [554]. This approach simplified the cloning process, reducing the cumulative number of rounds of plaque purification required to create a BAC. However, the pseudorabies BAC still required seven rounds of plaque purification and produced a virus with slowed growth kinetics and numerous attenuations [554]. Though faster than the methodology employed in this thesis it is distinctly possibly the resulting pseudorabies virus is less fit then it could

be if produced with alternative methods. In cloning CyCMV Mauritius BAC the majority of plaque purification rounds were employed to generate the intermediates CyCMV Mauritius-eGFP and CyCMV Mauritius-LoxP. Only two rounds of plaque purification occurred to isolate CyCMV Mauritius BAC. The BAC vector is large; its presence in the genome slows the growth of the virus and produces instability in the viral genome [401, 481, 486, 488] (See section 1.4.4.1). Plaque purification after the BAC has been added will preferentially select for faster growing viral variants that no longer produce or encode for genes unnecessary *in vitro*. Therefore plaque purification was minimized most stringently at this step for CyCMV Mauritius BAC.

CRISPR/Cas9 is not currently a preferential method of BAC construction. To compete with a step-wise recombination approach the single-recombination CRISPR/Cas9 method will need to require fewer rounds of plaque purification prior to transformation into *E. coli*. Still, as CRISPR/Cas9 mediated homologous recombination is used more frequently in other studies, and as the methodology continues to improve, it is likely that direct addition of the BAC vector by CRISPR/Cas9 will be favoured over traditional methods of homologous recombination.

After the plaque purification is complete, the methodology utilized to incorporate the virus as a BAC is no longer meaningful. Both viruses replicate faithfully in *E. coli* and can be edited by BAC recombination to correct unintentionally introduced mutations as easily as they are edited to introduce new and desired changes. BAC recombination was used in this manner with CyCMV Mauritius BAC clone 1 to correct the *loxP* sites flanking the BAC vector and later to introduce the SIV transgenes that facilitate the study of CMV as a vaccine vector.

4.2.2 Integration of the BAC Vector at a Cryptic *loxP* Site

The integration of the BAC vector at a cryptic *loxP* site in the *CyUL153* ORF instead of between CyUS1 and CyUS2 was unexpected. *LoxP* is comprised of two 13bp recognition regions that are inverted repeats, and a central 8bp asymmetrical spacer region. The probability of the 34bp *loxP* sequence occurring by chance in the 217 kb CyCMV Mauritius genome is near zero [540, 555]. As was expected, the sequenced CyCMV Mauritius genome did not contain a wildtype *loxP* sequence. It did however contain a number of cryptic *loxP* sequences – *loxP* like sequences that can interact with Cre [540, 555], something we identified only after observing mutations in the *loxP* sequences and the exotopic integration of the BAC vector.

Cre-loxP recombination is mediated when a Cre homotetramer ring binds two aligned *loxP* regions [543]. Most binding is along the DNA backbone, with nucleotide to protein contact occurring only along a 6 bp stretch in each spacer region [543]. Two separate *loxP* sequences are nicked in the spacer region at the central TA and a Holliday junction forms between the DNA strands [539, 556].

Cre-*loxP* recombination is tolerant to a number of changes along the recognition or spacer region, though typically the central TA is essential [538-540, 543]. A variety of *loxP* sequences are employed in molecular cloning to favour desired outcomes. The wildtype *loxP* sequence favours the excision of DNA over integration [539, 555]. CyCMV Mauritius BAC was constructed using *loxP* 511 which, compared to wildtype *loxP*, contains an A rather than G at position 7 of the 8bp spacer region [539] (**Figure 3-2**; Section 3.5.1). *LoxP* 511 has a high rate of recombination with other l*oxP* 511 sequences but a reduced rate of recombination with wildtype *loxP* [539].

Cryptic *loxP* sequences differ substantially from wildtype *loxP*. Using the criteria of Semprini et al. the CyCMV Mauritius genome was explored for three *loxP* like patterns [540]. Pattern 1 looks for wildtype recognition sequences with a tolerance for up to 10 mismatches separated by any 8bp (ATAACTTCGTATA NNNNNNNN TATACGAAGTTAT) [540]. Pattern 2 is weighted towards DNA base pairs in direct contact with the enzyme (ATNACNNCNTATA NNNNNNNN TATANGNNGTNAT) and allows five additional mismatches [540]. Pattern 3 looks for the TATA immediately surrounding the spacer region only, with no additional mismatches allowed (TATA NNNNNNNN TATA) [540]. Using the program Geneious 8.1.3 to search the genome following these criteria no sites were identified that matched all three patterns. CyCMV Mauritius contained 31 cryptic *loxP* sites matching pattern one, 9 matching pattern two, and 7 matching pattern three. By contrast RhCMV68-1 contains 26 sites matching pattern one, 15 matching pattern two, and 11 matching pattern three. RhCMV 68-1 was successfully cloned as a BAC using *loxP* methodology without any published issues regarding cryptic *loxP* sites [401]. It may be that issues did occur, but resulting clones were discarded, or that cryptic *loxP* sites have not caused unintended recombination in other BACs. CyCMV Mauritius contains two cryptic *loxP* sites in the *CyUL153* gene, the pattern one cryptic *loxP* site where the BAC integrated and a pattern two cryptic *loxP* site 118 bp downstream. *CyUS2* contains a pattern one cryptic loxP site roughly 400 bp upstream of the intended integration site. No notable recombination appears to have occurred between the other 46 cryptic l*oxP* sites of CyCMV Mauritius in the presence of Cre.

One of the mutations to the *loxP* sequence of CyCMV Mauritius-loxP, a reversion to a G at position 7 of the spacer region, increased the similarity of the intended *loxP* site and the *CyUL153* cryptic *loxP* site (**Figure 3-2**; Section 3.5.1). Similarity between cryptic *loxP* sites is correlated with increased efficiency of recombination between them [538-540, 555]. As discussed above it is possible this similarity facilitated recombination between the *loxP* site and the cryptic *loxP* site after it was first introduced (See Section 3.6.1). This same mutation, and the C to T transversion at position 8 of the recognition sequence, would have decreased the efficiency of recombination between the intended *loxP* site and the *loxP* 511 on the BAC plasmid pWC155 [538]. While recombination efficiency

would not have been higher between the *CyUL153* cryptic *loxP* site and the *loxP* 511 on the BAC plasmid, with a drop in efficiency at the intended site chance could account for a direct integration at the cryptic *loxP* site. This is, however, would have been quite an unlikely event.

The insertion of the BAC DNA in the *CyUL153* ORF is not expected to disadvantage the BAC. *CyUL153* is a transmembrane gene of the RL11 family, which are not well conserved between CMV strains. Viral growth curves show no notable difference between CyCMV Mauritius BAC and wildtype CyCMV Mauritius (**Figure 3-5**; See Section 3.5.3). Viral fitness, essential to contrast the CyCMV Mauritius-cynomolgus macaque model to the highly mutated RhCMV 68-1-rhesus macaque model, was maintained.

4.2.3 CyCMV Mauritius BAC is Excisable

Currently the *loxP* corrected CyCMV Mauritius BAC is an excisable BAC. This is necessary, as the removal of BAC DNA from CyCMV Mauritius-NTR and CyCMV Mauritius-GPF as the vaccines are reconstituted into viral form prevents genome instability. To remove the BAC DNA and prevent instability the CyCMV Mauritius BAC constructs must be infected into Cre expressing cell culture. Removal of the BAC DNA to produce infectious virus by Cre-*loxP* recombination can be a further source of mutation. Plaque purification to select for recombined virus can require up to 5 passages in the presence of Cre [489]. Introducing a *cre* transgene into the BAC between the two flanking loxP sites can reduce this. A single passage in cell culture may be sufficient to remove a *cre* expressing self-excising BAC from the virus [489]. Cre can interact with cryptic *loxP* sites in the virus and be mutagenic so it is important that the *cre* gene excise itself along with the BAC [540, 555]. Similarly, it is important the *cre* gene is not behind a bacterial promoter and contains an intron to prevent any expression in *E. coli* [489].

If CyCMV Mauritius BAC is to be used to generate a library of modified CyCMV strains, potentially to study the effect of gene knockouts, BAC recombination should be used to clone *cre* into the BAC. It is less essential for the initial use of CyCMV Mauritius as a viral vector where the BAC vector will be removed once and the generated viral vaccine sequenced and carefully screened.

4.3 Development of CyCMV Mauritius as a SIV Viral Vector

BAC recombination was used to introduce two SIV vaccine constructs, NTR and GPF, between *CyUS3* and *Cy215* of CyCMV Mauritius BAC clone 1 to construct CyCMV Mauritius-NTR and CyCMV Mauritius-GPF, the two vaccine constructs that will be utilized in the cynomolgus macaque-SIV vaccine challenge model in order to evaluate a HCMV based HIV vaccine strategy (See Section 3.5.1). Two constructs were used to match recent trials using RhCMV 68-1 in the rhesus macaque-

SIV challenge model and VZV_{Oka} in the cynomolgus macaque-SIV challenge model [185, 187, 244, 345], following the same logic that SIV vaccine trials including multiple target genes in the vaccine construct have historically been more successful than trials including only antigens of a single gene [51]. Furthermore, there is additional advantage in vaccinating with a broad range of antigens including both structural and non-structural SIV components [51, 475, 557]. In a DNA prime and poxvirus boost vaccine trial the higher breadth of resulting responses controlled SIVmac251 viremia better in Rev-Tat-Nef and Gag-Pol-Env dual vaccinated rhesus macaques than macaques vaccinated with either construct alone [557].

The CyCMV Mauritius based SIV vaccine was prepared as two separate constructs to be given together, as was the RhCMV 68-1 vectored vaccine [185, 187, 244], due to size constraints of the CMV vector. CMVs have a carrying capacity of roughly 230 kbp of DNA [286]. CyCMV Mauritius is 217 kbp, 225 kbp with the BAC included. Separating the 6 kbp GPF and 4 kbp NTR constructs improves the stability of the virus and prevents the loss of DNA segments from the virus during plaque purification before the BAC vector can be removed.

The structural vaccine construct GPF encodes a fused Gag-Pol polyprotein (See Section 3.5.4; **Table 3-2**). Gag and Pol were chosen due to their early presentation on surface MHC within two hours of viral entry [31, 42] (See Section 1.1.3.1). Gag specific CTL responses are capable of targeting infected cells prior to proviral integration [31]. Pol specific CTL are able to eliminate infected cells prior to viral production [42]. Pol in particular contain regions that are highly conserved across HIV clades [46] (See Section 1.1.1.1).

NTR encodes a fused Nef-Tat-Rev polyprotein (See Section 3.5.4; **Table 3-2**). The regulatory proteins Tat and Rev were selected due to relatively high levels of conservation across HIV clades, and natural responses that are correlated with disease non-progression [51, 63, 65, 66]. (See Section 1.1.1.1). The accessory protein Nef was chosen due to its essential role in highly pathogenic infection and early expression within 6-12 hours of viral entry [42, 51] (See Section 1.1.1.1).

Similar constructs, sometimes including *env*, have been utilized in a number of vaccines, including the RhCMV 68-1 SIV vaccines [185, 187, 200, 244]. The constructs used here were redesigned from those utilized in a previous VZV vectored SIV vaccine [345]. Both constructs were codon optimized for mammalian gene expression. Genes of the HIV genome utilize codons at different ratios then host genes [60, 61] (See Section 1.1.1.1), and mammalian codon optimization increases transgene expression in mammalian cells [60]. Herpesviruses similarly exhibit unique codon usage ratios, and seemingly utilize codon usage ratios to regulate gene expression in infected cells [558, 559]. At this time we have utilized host codon optimization, which is sufficient to express the genes in infected cells (**Figure 3-5**). Herpesvirus codon optimization would be expected to initially supress transgene

expression until rescued by viral processes, though it has been suggested such an approach may result in improved antibody responses in herpesvirus vectored vaccines [560]. Similar to with the expression of viral transgenes under early rather than IE promoters, this may be the result of the timing of transgene expression and may not produce the desired immune phenotype [31, 391].

Early rather than immediate expression of transgenes in a CMV vectored Ebola vaccine in rhesus macaques timed the expression of the transgenes with peak expression of MHC downregulating genes, disrupting the establishment of CTL responses [391]. In mice, similarly timed early gene expression created an initially more robust immune response that did not persist or adopt an inflationary memory T cell phenotype [31]. As MHC presentation and a robust T_{EM}/T_{RM} response are hypothesis we wish to test in this vaccine model, the genes in both CyCMV Mauritius constructs were behind the same strong, constitutive promoter used in the RhCMV 68-1 vectored SIV vaccines [185, 187, 200, 244].

For safety reasons, amino acids were changed in the catalytic domains of Pol, Nef, Tat and Rev to inactivate the genes (**Table 3-2**). Myristoylation sequences were removed from the N terminus of *gag* and *nef* to prevent localization of the genes to the lipid membrane. As Tat is excitotoxic to nerve cells [58], the secondary structure of the protein was further disrupted to prevent any activity by splitting the gene and inverting the N and C portions. *Nef*, *tat* and *rev* were encoded as a single construct and the frame shift region of the *gag-pol* ORF was removed to translate *gag* and *pol* constitutively as a single polyprotein (See Section 1.1.3.1).

Both transgene constructs were expressed in virally infected cells (See Section 3.5.4; **Figure 3-5**). Expression levels were similar to the constructs under the CMV promoter in VZV (**Figure 3-5**). Until the evaluation of immunogenicity has been completed in cynomolgus macaques (See Section 4.5.1) it cannot be known if this is a sufficient level of expression, but as the EF1α promoter is the same as that used for similar genes in RhCMV 68-1 [185, 187, 200, 244] gene expression levels are expected to be sufficient.

The transgenes remained stable when passaged in tissue culture (See Section 3.5.4), and similar transgenes continued to stimulate immune responses in the RhCMV 68-1 SIV vaccine trials for multiple years [185, 187, 200, 244]. Still, stability of the constructs *in vivo* will need to be confirmed to ensure the transgenes, which will present a target for the immune responses against the CMV vector, are not mutated or lost in vaccinated animals (See Section 4.5.1).

4.4 Limitations of CMV as a Vaccine Vector

There are numerous potential advantages to the use of HCMV as a live vaccine vector in humans, many resulting from the persistent infection and ongoing viral replication of HCMV. These

advantages do not mitigate the safety concerns that must be addressed with the use of any live vaccine vector.

There is potential for the attenuation of an HCMV vaccine vector to address these concerns, which can be modeled using CyCMV Mauritius. It is intended, however, that CyCMV Mauritius will be used in the cynomolgus macaque model without modification beyond those introduced while cloning it as a BAC. Modifications to the virus by removing genes to reduce the risk of transmission or morbidity should be considered only after an initial trial to elucidate of the mechanism of protection.

Furthermore, CMVs are capable of recombination when two or more strains infect the same cell [277]. As CMV is highly prevalent globally (See Section 1.2.2.1) it may not be possible to eliminate the chance of an attenuated CMV vaccine recombining with a wildtype CMV strain. Though HCMV Towne and HCMV AD169 have both been given safely as vaccines [328, 330, 561-563], these strains did not contain any genes that were not already present in circulating wildtype HCMV strains. Recombination of HCMV Towne or HCMV AD169 with clinical strains adds no new risk to any population. With a HCMV based HIV vaccine reverting to a wildtype phenotype, or a clinical strain of HCMV acquiring the vaccine transgenes, the transgenes could transmit despite any attenuation of the original HCMV vaccine. This is primarily problematic if the HIV based transgenes are unsafe in any population (such as pregnant mothers) or if the wildtype form of HCMV increases the risk of HIV acquisition in any populations, as did AD5 among activated CD4+ cell infiltrated mucosa [32, 46, 253, 254] (See Section 1.1.3.1).

Even without attenuation, HCMV infection is typically asymptomatic. The primary disease risk with HCMV inoculation is not in immune competent. HCMV can be transmitted to the foetus from a HCMV seropositive chronically infected mother, but since the foetus receives protective maternal immunoglobulin, the outcome is usually benign [265, 315, 316]. However, HCMV transmission rates are far higher and severity of disease is far worse when primary HCMV infection occurs in HCMV seronegative pregnant women [316-319] (See Section 1.2.2.1). Previously seropositive mothers with developed immune responses are at reduced risk for congenital HCMV related complications, a fact which has driven the development of live attenuated HCMV strains as vaccines against congenital HCMV [265, 316, 328, 330, 331, 561-563]. As live HCMV strains do not cause complications or concern administered to these populations, it is conceivable that a HCMV vectored HIV vaccine will be similarly accepted. Since much of the world lives in circumstances where HCMV acquisition occurs naturally during early childhood, and woman are seropositive for HCMV by adulthood, the risk of vaccination with a HCMV vector is not at play.

4.4.1 Chronic Inflammation and Immune Senescence

Despite a lack of acute pathogenicity in the normal adult host, there is concern that chronic inflammation accompanying CMV infection may be linked to or associated with complications related to immune senescence in the elderly. If this is the case, then a HCMV vectored HIV vaccine, which will remain and reactivate in vaccine recipients into their old age, may be ill advised. It is true that there is a general decline in immune function with age leading to increased morbidity and mortality from infectious diseases [258]. Changes include thymic atrophy and a loss of naïve cells, reduced proliferation of immune cells, reduced clonal expansion, and changes in cytokine secretion patterns [258, 261]. There are resulting regulatory differences in a number of immune related genes and an increase in the ratio of CD4+ to CD8+ T cells [258, 261]. Interestingly there is also an expansion of memory cells and especially CD28- CD27- CCR7- T_{EM} in the elderly [258, 261].

HCMV infection is lifelong, as is inoculation with a HCMV based vaccine. Despite this there is little evidence that CMV interacts negatively with the aged immune system. The elderly rarely experience HCMV related disease [258]. The general observed decline in immune function that occurs against many common pathogens is not observed with CMV. In aged rhesus macaques, compared to adult rhesus macaques, there is no change in the number of circulating CD4+ or CD8+ T cells targeting CMV and the immune response against CMV is unchanged [258]. In humans there is an increased portion of CD8+ and CD4+ T cells targeting HCMV accompanying immune senescence [261].

CMV is a persistent infection and lifelong CMV infection does drive inflammation [258, 323]. Long-term exposure to persistent CMV infection results in systemic high expression of Th1 cytokines [309, 323]. This is linked to inflammation of the vasculature and leads to arterial stiffness and increased risk of cardiac complications later in life [309, 337, 338]. A study comparing differences in gene pathways between elderly (≥90 years old) CMV seropositive and seronegative individuals compared to CMV seronegative 30-years olds found an increase in the portion of memory B cells and CD4+ CD28- T cells [261]. An elderly CMV seropositive human cohort had increased levels of IFNγ, changes to gene regulation in several inflammatory and anti-inflammatory signalling pathways – including the docosahexaenoic acid signalling pathway, and the emergence of perforin and granzyme B expressing CD4+ T cells [261]. Yet overall gene expression differences between the elderly CMV seronegative and young CMV seronegative cohorts far outnumbered the differences between elderly CMV seropositive and young CMV seronegative individuals [261].

CMV has evolved with humans for millennia and until very recently the CMV seroprevalence rate could be assumed to be near 100% as they are in the developing world [263]. Changes to our immune

system as a result of CMV may be part of a calibrated immune response evolved in coordination with the host virome. MCMV seropositive mice are better protected against challenge with *Yersinia pestis*, *Listeria monocytogenes*, and influenza [259]. This protection can be attributed to elevated levels of TNFα, IFNγ and chronic macrophage activation resulting from CMV infection [259]. CMV seropositive mice subjected to CD4+ and CD8+ cell depletion prior to challenge with *L. monocytogenes* remained protected [259]. Chronic infection with HSV-1 or Sindbis virus did not provide the same chronic inflammation or similar protection [259]. CMV enhanced protection is not universal to all pathogens; MCMV seropositive mice were no better protected against West Nile virus, though they fared no worse than MCMV seronegative mice [259].

There are harms associated with CMV late in life, but these occur after childbearing age, the time when evolutionary selection has its primary effect. These risks include cardiovascular disease [309, 337, 338] and also potentially an increased risk of diabetes [564]. However, there may also be substantial benefits [261]. While our ancient human and pre-human ancestors did not lived as long as we expect to, CMV is part of an evolved virome. The differences in gene regulation and immune activation of CMV seronegative nonagenarians may be unhealthy compared to CMV seropositive individuals in meaningful ways – similar to the odd immune phenotypes observed in pathogen and germ free mice [259-261]. There is a symbiotic relationship between CMV and host that shapes the young immune response in ways that provides protection against some pathogens [259, 260, 262]. Changes in gene regulation relating to NKC and T cell receptor signalling that occur in CMV seronegative nonagenarians in comparison to 30-year-old controls do not occur in elderly CMV seropositive individuals [261]. In the elderly it may be that elements of immune senescence are warded off by HCMV infection rather than caused by HCMV, though it is clear if HCMV is to be used as a vaccine vector that this field will need further examination.

4.5 Future Directions

4.5.1 Immunogenicity the CyCMV Mauritius Vaccine Constructs

Prior to a cynomolgus macaque-SIV challenge trial designed to evaluate the efficacy of the developed CyCMV Mauritius vaccine constructs (See Section 4.5.2) it will be necessary to evaluate the *in vivo* infectivity of the constructs, stability of the transgenes, and immunogenicity of the vaccines in the cynomolgus macaque model. The BAC vector will be removed from CyCMV Mauritius-GPF, CyCMV Mauritius-NTR and CyCMV Mauritius BAC by Cre/*loxP* recombination, and the virus sequenced to confirm fitness following plaque purification.

Rhesus macaques given RhCMV 68-1 vaccine constructs were inoculated with 5×10^6 PFU of each vector [185, 244]. The CyCMV Mauritius vaccine constructs will be given subcutaneously in similar dosage. Both CyCMV Mauritius vaccine constructs and transgene free BAC derived CyCMV Mauritius will be individually grown to high titre on fibroblast cell lines to produce sufficient virus for a trial.

Eight healthy, adult male Mauritian origin cynomolgus macaques screened for known restriction factors, including TRIM5α (See Section 1.1.1.1) and protective MHC haplotypes, will be divided into three groups and MHC haplotype matched. A control group (n=2) will be given a saline mock inoculation, a control group (n=2) will be given 1×10^7 PFU each of transgene free CyCMV Mauritius derived from CyCMV Mauritius BAC, and the vaccine group (n=4) will be given a 1:1 mixture of CyCMV Mauritius-GPF and CyCMV Mauritius-NTR at 5×10^6 PFU each (1×10^7 PFU total).

A protocol similar to those described with RhCMV 68-1 in cynomolgus macaques [291, 379] and in rhesus macaques [185] will be used to confirm the infectivity of the CyCMV Mauritius vaccine constructs. Fresh urine will be collected from the macaques' cage pans weekly. The urine will be filter sterilized and centrifuged at 900 ×*g* and 4˚C for 30 minutes to remove debris. CyCMV virus will be pelleted from the urine at 20 000 ×*g* and 4˚C for 1 hour then mixed with cell media and co-cultured over MRC-5 cells. MRC-5 cultures will be maintained until CMV CPE is observed or until 8 weeks have passed. As all cynomolgus macaques are CMV seropositive the infected cells will be lysed and probed by western blot for SIV transgenes expression rather than for expression of CMV specific genes.

Quantitative PCR (qPCR) using primer sets for both SIV transgenes and the attenuated CyCMV Mauritius BAC *UL128-UL130* region, as well as a control CyCMV primer set, will be performed weekly on whole blood, serum, and urine co-cultures to detect CyCMV Mauritius DNA. Importantly this will confirm infection with the CyCMV Mauritius vaccine constructs separate from endogenous CMV in the absence of viral shedding.

Viral outgrowth and qPCR will be continued throughout the duration of the infectivity study to confirm the persistence and stability of the viral vector and the stability throughout.

The strength and phenotype of the garnered immune response will be assayed in blood and mucosal tissue. One hypothesis for the benefit of T_{EM} and T_{RM} responses is their abundance as effector cells already located in the periphery, at the portals of entry, and the sites of initial spread for SIV [244]. It is therefore important to characterize the systemic immune response as well as the mucosal response. Whole blood will be collected and separated into Serum and PBMCs to assay the systemic immune response as previously described [185, 476]. Mucosal samples will be collected

through rectal swabs, rectal biopsies and bronchoalveolar lavages (BALs) then processed to isolate lymphocytes as previously described [185, 475]. Draining lymph node biopsies will be taken as sites of expansion along at several timepoints. Biopsies will be homogenized and processed to isolate lymphocytes for intracellular cytokine staining.

PBMCs or isolated lymphocytes will be stimulated with Nef, Tat, Rev, Gag or Pol overlapping 15-mer polypeptide pools and co-stimulated with CD28 and CD49d (with a negative control co-stimulated in the absence of peptide pools) for one hour before protein transport is suspended with Brefeldin A and incubated for a further 11 hours. After stimulation cells will be stained with fluorescent linked antibodies and analyzed by flow cytometry. Staining for CD3 (T cell receptor), CD4 and CD8 will be used to differentiation CD4+ T cells and CD8+ T cells. CCR7 and CD28 will be used to differentiate CD28+/CCR7+ central memory and naive cells, CD28-/CCR7- T_{EM} and T_{RM}, and CD28+/CCR7- transitional effector memory cells [185, 244, 565]. CD45RA will be used to differentiate CD45RA+ naïve cells from CD45RA- central memory cells [565, 566]. Intracellular CD69, a marker of recent activation, will be used to confirm cells were recently activated in response to the polypeptide stimulation [323]. Intracellular staining for IFNγ and TNFα will then be used to measure cell activation in response to the SIV peptides [244, 390]. Responses to Vif and Env peptides, which will be useful after SIV challenge to distinguish responses to vaccine antigens from non-vaccine antigens, can be included as additional negative controls.

Though the transgenes are not designed to provoke antibody responses, antibodies develop naturally against multiple non-surface CMV encoded genes [309, 366]. A RhCMV 68-1 vectored Ebola vaccine saw antibody mediated protection against the included transgenes [391] (See Section 1.3.1). Plasma will be analyzed by Enzyme Linked Immunosorbent Assay (ELISA) as in Marzi et al. [391] and SIV western blot as done previously in our lab [476] to detect the presence of transgene binding antibody. A B cell Enzyme Linked Immunospot assay (ELISPOT) will be performed to titre SIV transgene specific plasma cell concentrations in whole blood [391]. If antibody is detected serum can be used in a neutralization assay to SIV [567].

In most RhCMV 68-1 based vaccine trials the immune response continues to expand long after the initial vaccination [185, 187, 200, 244, 391-394]. Ideally the CyCMV Mauritius-SIV vaccinated macaques should be followed for a full year to ascertain the durability of the immune response and the stability of the viral construct, however economics dictate that this may not be possible so the quantitative CyCMV viral load with be followed to a minimum of 12 weeks after which the animals will be necropsied so that both virologic and immunologic analysis of the tissue can be conducted..

4.5.2 Proposed CyCMV Mauritius Vaccine Trial

Provided an acceptable immune phenotype is observed in vaccinated cynomolgus macaques throughout the CyCMV Mauritius vaccine infectivity trial (See Section 4.5.1), the study will be followed by a SIV challenge trial. A multi-arm study with group size of twelve will be employed, providing an adequate sample size to show 50% protection from disease among the vaccinated animals using α=0.05 β=0.2. A similar level of protection was seen in previous herpesvirus based vaccines studies [185, 244, 345, 476, 568]. From previous studies we know that sterilizing immunity is unlikely, though it will be assessed [185, 187, 244, 345]. The primary outcome will be the proportion of animals protected from disease progression. The secondary outcomes will be the difference in peak and set-point viral loads and composite differences in cumulative viral load over time between vaccine recipients and controls.

The vaccinated, empty vector control, and mock vaccinated adult male Mauritian cynomolgus macaques will all be subject to an increasing low-dose intrarectal SIV challenge. The challenge will be delayed a minimum of 20 weeks post vaccination to allow time for the immune response from the CyCMV Mauritius vaccine to stabilize. Exact timing post vaccination should be determined cognizant of the results of the infectivity study.

The macaques will be challenged intrarectally on a weekly basis with molecularly tagged SIVmac239 swarm [569]. As a tagged swarm the SIVmac239 clone still represents a consistent, robust challenge in each animal but determination of transmitted/founder variants is possible to identify multiple infections [569]. Such a challenge virus will also match the SIVmac239 based transgenes used to construct the vaccine.

Twice weekly blood draws will be used to track SIV plasma viral load by nested qPCR. An initial detection of greater than 30 copies of SIV per milliliter blood (or above the then current limit of detection) on two consecutive occasions will be used to confirm SIV acquisition and subsequent viral challenge will be stopped for animals that have tested positive. PCR positive animals will continue to be monitored with weekly qPCR tests for viral load, as well as a continuation of the tests of the immunogenicity study (See Section 4.5.1) to correlate the immune response with protection. These include flow cytometric staining of lymphocytes [185, 475], ELISA/western blot to detect antibody against SIV [185, 476], ELISPOT for SIV specific plasma cells [391], CyCMV vaccine outgrowth and qPCR of CyCMV transgenes [185]. Immune responses against SIV Env and Vif will be used to examine the infected macaques' immune responses to non-vaccine antigens (See Section 4.5.1). Assays of CyCMV Mauritius vector persistence (qPCR) and shedding (viral outgrowth from urine)

will look for correlation between the persistence and replication of the vaccine vector to protection from SIV.

Protection or lack of protection from SIV acquisition will be examined initially by number of challenges and cumulative viral dose to initial infection. Protection from SIV disease will further look at peak viral load and set-point viral load of SIV in vaccinated versus control macaques. SIV positive animals who develop a stable set-point viral load below the limit of detection with or without viral blips, as has been seen in multiple RhCMV vectored SIV vaccines [185, 187, 200, 244], will be considered protected from disease progression and should be followed as long as is logistically possible to ensure a return to viremia does not occur.

4.5.3 Moving Beyond the Macaque Model

As stated above, the use of CyCMV Mauritius-GPF and CyCMV Mauritius-NTR as vaccines against SIV in the cynomolgus macaque-SIV challenge model is intended to inform the design of any future HCMV vectored HIV vaccine (See Section 4.1). Due to species specificity of the virus, any CMV-based vector for a HIV vaccine in humans will require a HCMV. Numerous strains of HCMV have been isolated and cultured. Two early isolates, HCMV Towne and HCMV AD169, are considered laboratory strains due to a high level of attenuation and limited ability to persist *in vivo* [561]. HCMV strains previously tested as vaccine should be the initial candidates considered as vectors to develop a vaccine. Both AD169 and the Towne strain exist as readily modifiable BACs [481, 531]. However, due to the level of attenuation in both strains, introduced through repeated passaging in cell culture, HCMV Towne and HCMV AD169 do not induce the same immune responses that less attenuated clinical isolates of CMV do, like the Merlin or Toldeo strains [331, 561, 570].

Which HCMV strain is to be used, and how it is to be modified, will ultimately depend upon the results of the CyCMV Mauritius based SIV vaccine trial and how they relate to those of the RhCMV 68-1 based vaccine trials. What factors mediate protection, be they T_{EM} responses, MHC-II and MHC-E targeting CD8+ T cells, or other factors, will influence the design of a HCMV vaccine vector. Already attempts have been made to graft RhCMV 68-1-like mutations onto a HCMV Towne/Toledo chimera [386]. Considering the large CMV genome still includes a number of genes without known function, this is premature. Only once the mechanism of protection is better understood can an HCMV strain be responsibly mutated, likely aided by the direct comparison of CyCMV Mauritian and RhCMV 68-1.

Furthermore, the SIV antigens used in CyCMV Mauritius-NTR and CyCMV Mauritius-GPF, alongside those used in the RhCMV 68-1 vectored vaccines, can only aid in the design of antigenic

HIV transgenes for a HCMV vectored HIV vaccine. Unlike the transgenes of the CyCMV Mauritius based vaccines, which are based off a single viral clone, SIVmac239. The antigens ultimately utilized in a HCMV vectored HIV vaccine will likely need to include founder sequence regions from multiple HIV clades to reasonably protect from a broad range of global HIV strains.

Additionally, neither CyCMV Mauritius-SIV vaccines, nor the later RhCMV 68-1 vectored SIV vaccines included Env, the primary antibody target. It is possible that such an approach will provide sufficient protection, but more likely a successful HCMV based HIV vaccine approach will include a paired antibody based HIV vaccine to mediate protection. For example, a co-administered vaccine generating a promising HIV-1 neutralizing antibody against envelope is a possibility as a prime boost vaccine strategy.

4.6 Conclusion

This thesis set out to prepare a Mauritius strain of CyCMV as a SIV vaccine for use in cynomolgus macaques in order to evaluate the use of HCMV vectored vaccines against HIV. As promising as the RhCMV 68-1 vectored vaccines have been in the rhesus macaque model, unanswered questions remain surrounding the mechanism of protection. It is not known why only half of the animals vaccinated with the RhCMV 68-1 SIV vaccine supress viremia [185, 187, 200, 244]. It is not understood what the implications are of MHC-I, MHC-II and MHC-E epitope targeting by CD8+ T cells in RhCMV 68-1 vaccinated rhesus macaques, only that this may not apply outside the model to HCMV [200, 245, 386, 404]. It is not understood whether inflationary T cell and antibody responses can be derived from the same CMV vector [391]. These questions do not temper the outstanding optimism the success of the RhCMV 68-1 based vaccines have brought to the field, but they do dictate the direction it must take. Whether the results of the RhCMV vaccines are due to characteristics of CMV based vaccines as a whole and will apply in human trials, or are due to eccentricities of a single heavily passaged virus cannot be answered using RhCMV 68-1 alone.

In this thesis we presented CyCMV Mauritius: a strain of CMV isolated from a cynomolgus macaque of Mauritian origin. We sequenced the virus, and compared it in the context of HCMV and the highly attenuated and well-used RhCMV 68-1 model. We cloned CyCMV Mauritius as a BAC, maintaining the virus as a minimally passaged viral clone, and we prepared from this BAC two CyCMV Mauritius vectored SIV vaccine constructs for evaluation in the cynomolgus macaque-SIV vaccine model.

CyCMV Mauritius-GPF and CyCMV Mauritius-NTR provide the first opportunity in macaques to contrast multiple CMV vaccine vector strains. As the first CMV based SIV vaccines that have been generated with a focus on the retention of virus fitness they present the first opportunity to examine

CMV based vaccination in a near wildtype strain. As the first CMV based vaccine constructed using cynomolgus macaque based CMV they present the first opportunity to study CMV vectored SIV vaccines in the cynomolgus macaque model. These are valuable molecular tool that facilitate the further understanding of herpesviruses as vaccine vectors and their potential application in humans.

It will be years, if not decades, before HCMV is used as a vaccine vector against human diseases. The constructs prepared in this thesis cannot be used directly in human trials, but they will serve in animal model trial to inform on the mechanisms of action through which HCMVs, as a highly evolved viral symbiote, thoroughly manipulate their host immune system. In doing so this work is a necessary and fruitful step to further the understanding of HCMV as a vaccine vector and, perhaps, to unlocking the path to an effective HIV vaccine.

Chapter Five: References

1. Gallo, R.C., et al., *Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS.* Science, 1984. **224**(4648): p. 500-3.
2. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.
3. Kuritzkes, D.R. and B.D. Walker, *HIV-1: Pathogenesis, clinical manifestations and treatment* in *Feilds Virology*, D.M. Knipe, et al., Editors. 2007, Lippincott, Williams & Wilkins: Philadelphia. p. 2187-2214.
4. Kuritzkes, D.R. and B.D. Walker, *HIV and Their Replication*, in *Feilds Virology*, D.M. Knipe and P. Howley, Editors. 2007, Lippincott, Williams & Wilkins: Philadelphia, PA. p. 2187-2214.
5. Frankel, A.D. and J.A. Young, *HIV-1: fifteen proteins and an RNA.* Annu Rev Biochem, 1998. **67**: p. 1-25.
6. Freed, E.O. and M.A. Martins, *HIVs and Their Replication*, in *Fields Virology*, D.M. Knipe, et al., Editors. 2007, Lippincott, Williams & Wilkins: Philadelphia. p. 2107-2185.
7. Gonda, M.A., et al., *Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus.* Science, 1985. **227**(4683): p. 173-7.
8. Li, Y., et al., *Complete nucleotide sequence, genome organization, and biological properties of human immunodeficiency virus type 1 in vivo: evidence for limited defectiveness and complementation.* J Virol, 1992. **66**(11): p. 6587-600.
9. Burdick, R.C., et al., *Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes.* PLoS Pathogens, 2017. **13**(8): p. e1006570.
10. McDonald, D., et al., *Visualization of the intracellular behavior of HIV in living cells.* J Cell Biol, 2002. **159**(3): p. 441-52.
11. Chinen, J. and W.T. Shearer, *Molecular virology and immunology of HIV infection.* Journal of Allergy and Clinical Immunology, 2002. **110**(2): p. 189-198.
12. Pace, M.J., et al., *HIV reservoirs and latency models.* Virology, 2011. **411**(2): p. 344-54.
13. Laird, G.M., et al., *Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay.* PLoS Pathog, 2013. **9**(5): p. e1003398.
14. Chun, T.W., et al., *In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency.* Nat Med, 1995. **1**(12): p. 1284-90.
15. Zhang, Z., et al., *Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells.* Science, 1999. **286**(5443): p. 1353-7.
16. Van Lint, C., S. Bouchat, and A. Marcello, *HIV-1 transcription and latency: an update.* Retrovirology, 2013. **10**: p. 67-67.
17. Chun, T.W., et al., *Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13193-7.
18. Maartens, G., C. Celum, and S.R. Lewin, *HIV infection: epidemiology, pathogenesis, treatment, and prevention.* Lancet, 2014. **384**(9939): p. 258-71.

19. Sundquist, W.I. and H.-G. Kräusslich, *HIV-1 Assembly, Budding, and Maturation.* Cold Spring Harbor Perspectives in Medicine, 2012. **2**(7): p. a006924.
20. Mohammadi, P., et al., *24 Hours in the Life of HIV-1 in a T Cell Line.* PLOS Pathogens, 2013. **9**(1): p. e1003161.
21. van't Wout, A.B., et al., *Macrophage-tropic variants initiate human immunodeficiency virus type 1 infection after sexual, parenteral, and vertical transmission.* J Clin Invest, 1994. **94**(5): p. 2060-7.
22. Zhu, T., et al., *Genotypic and phenotypic characterization of HIV-1 patients with primary infection.* Science, 1993. **261**(5125): p. 1179-81.
23. Deng, H., et al., *Identification of a major co-receptor for primary isolates of HIV-1.* Nature, 1996. **381**(6584): p. 661-6.
24. Dragic, T., et al., *HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.* Nature, 1996. **381**(6584): p. 667-73.
25. Moore, J.P., et al., *The CCR5 and CXCR4 Coreceptors—Central to Understanding the Transmission and Pathogenesis of Human Immunodeficiency Virus Type 1 Infection.* AIDS Research and Human Retroviruses, 2004. **20**(1): p. 111-126.
26. Margolis, L. and R. Shattock, *Selective transmission of CCR5-utilizing HIV-1: the 'gatekeeper' problem resolved?* Nat Rev Microbiol, 2006. **4**(4): p. 312-7.
27. Endres, M.J., et al., *CD4-independent infection by HIV-2 is mediated by fusin/CXCR4.* Cell, 1996. **87**(4): p. 745-56.
28. Feng, Y., et al., *HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.* Science, 1996. **272**(5263): p. 872-7.
29. McMichael, A.J., et al., *The immune response during acute HIV-1 infection: clues for vaccine development.* Nat Rev Immunol, 2010. **10**(1): p. 11-23.
30. Allers, K., et al., *Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation.* Blood, 2011. **117**(10): p. 2791-9.
31. Sacha, J.B., et al., *Gag-specific CD8(+) T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression.* Journal of Immunology, 2007. **178**(5): p. 2746-2754.
32. Duerr, A., et al., *Extended follow-up confirms early vaccine-enhanced risk of HIV acquisition and demonstrates waning effect over time among participants in a randomized trial of recombinant adenovirus HIV vaccine (Step Study).* J Infect Dis, 2012. **206**(2): p. 258-66.
33. Takeda, A., C.U. Tuazon, and F.A. Ennis, *Antibody-enhanced infection by HIV-1 via Fc receptor-mediated entry.* Science, 1988. **242**(4878): p. 580.
34. Guzzo, C., et al., *Virion incorporation of integrin α4β7 facilitates HIV-1 infection and intestinal homing.* Science Immunology, 2017. **2**(11): p. eaam7341.
35. Tremblay, M.J., J.F. Fortin, and R. Cantin, *The acquisition of host-encoded proteins by nascent HIV-1.* Immunol Today, 1998. **19**(8): p. 346-51.
36. Martin, G. and M.J. Tremblay, *HLA-DR, ICAM-1, CD40, CD40L, and CD86 are incorporated to a similar degree into clinical human immunodeficiency virus type 1 variants expanded in natural reservoirs such as peripheral blood mononuclear cells and human lymphoid tissue cultured ex vivo.* Clinical Immunology, 2004. **111**(3): p. 275-285.
37. Arthur, L.O., et al., *Cellular proteins bound to immunodeficiency viruses: implications for pathogenesis and vaccines.* Science, 1992. **258**(5090): p. 1935-8.

38. Staprans, S.I., et al., *Enhanced SIV replication and accelerated progression to AIDS in macaques primed to mount a CD4 T cell response to the SIV envelope protein.* Proc Natl Acad Sci U S A, 2004. **101**(35): p. 13026-31.
39. Levitz, L., et al., *Conservation of HIV-1 T cell epitopes across time and clades: validation of immunogenic HLA-A2 epitopes selected for the GAIA HIV vaccine.* Vaccine, 2012. **30**(52): p. 7547-60.
40. Rousseau, C.M., et al., *HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load.* J Virol, 2008. **82**(13): p. 6434-46.
41. Jacks, T., et al., *Characterization of ribosomal frameshifting in HIV-1 gag-pol expression.* Nature, 1988. **331**: p. 280-283.
42. Sacha, J.B., et al., *Pol-specific CD8+ T cells recognize simian immunodeficiency virus-infected cells prior to Nef-mediated major histocompatibility complex class I downregulation.* J Virol, 2007. **81**(21): p. 11703-12.
43. Cromer, D., et al., *Estimating the in-vivo HIV template switching and recombination rate.* Aids, 2016. **30**(2): p. 185-92.
44. Cuevas, J.M., et al., *Extremely High Mutation Rate of HIV-1 In Vivo.* PLOS Biology, 2015. **13**(9): p. e1002251.
45. Eriksson, S., et al., *Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies.* PLoS Pathog, 2013. **9**(2): p. e1003174.
46. Buchbinder, S.P., et al., *Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial.* Lancet, 2008. **372**(9653): p. 1881-1893.
47. Chang, H.C., et al., *HIV-1 Tat protein exits from cells via a leaderless secretory pathway and binds to extracellular matrix-associated heparan sulfate proteoglycans through its basic region.* AIDS, 1997. **11**(12): p. 1421-1431.
48. Frankel, A.D. and C.O. Pabo, *Cellular uptake of the tat protein from human immunodeficiency virus.* Cell, 1988. **55**(6): p. 1189-1193.
49. Helland, D.E., et al., *Transcellular transactivation by the human immunodeficiency virus type 1 tat protein.* J Virol, 1991. **65**(8): p. 4547-9.
50. Arya, S., et al., *Trans-activator gene of human T-lymphotropic virus type III (HTLV-III).* Science, 1985. **229**(4708): p. 69-73.
51. Ferrantelli, F., A. Cafaro, and B. Ensoli, *Nonstructural HIV proteins as targets for prophylactic or therapeutic vaccines.* Curr Opin Biotechnol, 2004. **15**(6): p. 543-56.
52. Caputo, A., et al., *HIV-1 Tat-based vaccines: an overview and perspectives in the field of HIV/AIDS vaccine development.* Int Rev Immunol, 2009. **28**(5): p. 285-334.
53. Fanales-Belasio, E., et al., *HIV-1 Tat addresses dendritic cells to induce a predominant Th1-type adaptive immune response that appears prevalent in the asymptomatic stage of infection.* J Immunol, 2009. **182**(5): p. 2888-97.
54. Fanales-Belasio, E., et al., *Native HIV-1 Tat protein targets monocyte-derived dendritic cells and enhances their maturation, function, and antigen-specific T cell responses.* J Immunol, 2002. **168**(1): p. 197-206.
55. Secchiero, P., et al., *Extracellular HIV-1 tat protein up-regulates the expression of surface CXC-chemokine receptor 4 in resting CD4+ T cells.* J Immunol, 1999. **162**(4): p. 2427-31.

56. Huang, L., et al., *Tat Protein Induces Human Immunodeficiency Virus Type 1 (HIV-1) Coreceptors and Promotes Infection with both Macrophage-Tropic and T-Lymphotropic HIV-1 Strains.* Journal of Virology, 1998. **72**(11): p. 8952-8960.
57. Gibellini, D., et al., *Selective up-regulation of functional CXCR4 expression in erythroid cells by HIV-1 Tat protein.* Clinical and Experimental Immunology, 2003. **131**(3): p. 428-435.
58. Hui, L., et al., *Role of endolysosomes in HIV-1 Tat-induced neurotoxicity.* ASN Neuro, 2012. **4**(4): p. 243-52.
59. Felber, B.K., C.M. Drysdale, and G.N. Pavlakis, *Feedback regulation of human immunodeficiency virus type 1 expression by the Rev protein.* J Virol, 1990. **64**(8): p. 3734-41.
60. Li, M., et al., *Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11.* Nature, 2012. **491**(7422): p. 125-8.
61. Haas, J., E.C. Park, and B. Seed, *Codon usage limitation in the expression of HIV-1 envelope glycoprotein.* Curr Biol, 1996. **6**(3): p. 315-24.
62. Pollard, V.W. and M.H. Malim, *The HIV-1 Rev protein.* Annu Rev Microbiol, 1998. **52**: p. 491-532.
63. Li, G., et al., *An integrated map of HIV genome-wide variation from a population perspective.* Retrovirology, 2015. **12**(1): p. 18.
64. Butt, et al., *Sequence Conservation and Antibody Cross-Recognition of Clade B Human Immunodeficiency Virus (HIV) Type 1 Tat Protein in HIV-1-Infected Italians, Ugandans, and South Africans.* The Journal of Infectious Diseases, 2003. **188**(8): p. 1171-1180.
65. Cao, J., et al., *Evolution of CD8+ T cell immunity and viral escape following acute HIV-1 infection.* J Immunol, 2003. **171**(7): p. 3837-46.
66. van Baalen, C.A., et al., *Human immunodeficiency virus type 1 Rev- and Tat-specific cytotoxic T lymphocyte frequencies inversely correlate with rapid progression to AIDS.* J Gen Virol, 1997. **78 (Pt 8)**: p. 1913-8.
67. Ensoli, B., et al., *Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation.* Journal of Virology, 1993. **67**(1): p. 277-287.
68. Bellino, S., et al., *The presence of anti-Tat antibodies in HIV-infected individuals is associated with containment of CD4+T-cell decay and viral load, and with delay of disease progression: results of a 3-year cohort study.* Retrovirology, 2014. **11**(1): p. 49.
69. Anderson, J.L. and T.J. Hope, *Recent insights into HIV accessory proteins.* Current Infectious Disease Reports, 2003. **5**(5): p. 439-450.
70. Sakai, H., et al., *Function of human immunodeficiency virus type 1 Vpu protein in various cell types.* J Gen Virol, 1995. **76 (Pt 11)**: p. 2717-22.
71. Deora, A. and L. Ratner, *Viral protein U (Vpu)-mediated enhancement of human immunodeficiency virus type 1 particle release depends on the rate of cellular proliferation.* J Virol, 2001. **75**(14): p. 6714-8.
72. Andrew, A. and K. Strebel, *HIV-1 Vpu targets cell surface markers CD4 and BST-2 through distinct mechanisms.* Mol Aspects Med, 2010. **31**(5): p. 407-17.

73. Laguette, N., et al., *Evolutionary and Functional Analyses of the Interaction between the Myeloid Restriction Factor SAMHD1 and the Lentiviral Vpx Protein.* Cell host & microbe, 2012. **11**(2): p. 205-217.
74. Macreadie, I.G., et al., *Expression of HIV-1 nef in yeast: the 27 kDa Nef protein is myristylated and fractionates with the nucleus.* Yeast, 1993. **9**(6): p. 565-73.
75. Shugars, D.C., et al., *Analysis of human immunodeficiency virus type 1 nef gene sequences present in vivo.* Journal of Virology, 1993. **67**(8): p. 4639-4650.
76. Kaminchik, J., et al., *Genetic characterization of human immunodeficiency virus type 1 nef gene products translated in vitro and expressed in mammalian cells.* J Virol, 1991. **65**(2): p. 583-8.
77. Forshey, B.M. and C. Aiken, *Disassembly of Human Immunodeficiency Virus Type 1 Cores In Vitro Reveals Association of Nef with the Subviral Ribonucleoprotein Complex.* Journal of Virology, 2003. **77**(7): p. 4409-4414.
78. Ovod, V., et al., *Immunological variation and immunohistochemical localization of HIV-1 Nef demonstrated with monoclonal antibodies.* Aids, 1992. **6**(1): p. 25-34.
79. Kienzle, N., et al., *Expression and cellular localization of the Nef protein from human immunodeficiency virus-1 in stably transfected B-cells.* Arch Virol, 1992. **124**(1-2): p. 123-32.
80. Greenway, A.L., et al., *Human immunodeficiency virus type 1 Nef binds to tumor suppressor p53 and protects cells against p53-mediated apoptosis.* J Virol, 2002. **76**(6): p. 2692-702.
81. Varin, A., et al., *Exogenous Nef protein activates NF-kappa B, AP-1, and c-Jun N-terminal kinase and stimulates HIV transcription in promonocytic cells. Role in AIDS pathogenesis.* J Biol Chem, 2003. **278**(4): p. 2219-27.
82. Sol-Foulon, N., et al., *HIV-1 Nef-induced upregulation of DC-SIGN in dendritic cells promotes lymphocyte clustering and viral spread.* Immunity, 2002. **16**(1): p. 145-55.
83. Blagoveshchenskaya, A.D., et al., *HIV-1 Nef downregulates MHC-I by a PACS-1- and PI3K-regulated ARF6 endocytic pathway.* Cell, 2002. **111**(6): p. 853-66.
84. Lundquist, C.A., et al., *Nef-mediated downregulation of CD4 enhances human immunodeficiency virus type 1 replication in primary T lymphocytes.* J Virol, 2002. **76**(9): p. 4625-33.
85. Williams, M., et al., *Direct binding of human immunodeficiency virus type 1 Nef to the major histocompatibility complex class I (MHC-I) cytoplasmic tail disrupts MHC-I trafficking.* J Virol, 2002. **76**(23): p. 12173-84.
86. Stoddart, C.A., et al., *Human immunodeficiency virus type 1 Nef-mediated downregulation of CD4 correlates with Nef enhancement of viral pathogenesis.* J Virol, 2003. **77**(3): p. 2124-33.
87. Chowers, M.Y., et al., *Optimal infectivity in vitro of human immunodeficiency virus type 1 requires an intact nef gene.* Journal of Virology, 1994. **68**(5): p. 2906-2914.
88. Learmont, J.C., et al., *Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. A report from the Sydney Blood Bank Cohort.* N Engl J Med, 1999. **340**(22): p. 1715-22.
89. Learmont, J., et al., *Long-term symptomless HIV-1 infection in recipients of blood products from a single donor.* The Lancet, 1992. **340**(8824): p. 863-867.
90. Daniel, M.D., et al., *Protective effects of a live attenuated SIV vaccine with a deletion in the nef gene.* Science, 1992. **258**(5090): p. 1938-41.

91. Fukazawa, Y., et al., *Lymph node T cell responses predict the efficacy of live attenuated SIV vaccines.* Nat Med, 2012. **18**(11): p. 1673-81.
92. Johnson, R.P. and R.C. Desrosiers, *Protective immunity induced by live attenuated simian immunodeficiency virus.* Curr Opin Immunol, 1998. **10**(4): p. 436-43.
93. Whitney, J.B. and R.M. Ruprecht, *Live attenuated HIV vaccines: pitfalls and prospects.* Curr Opin Infect Dis, 2004. **17**(1): p. 17-26.
94. Fuchs, S.P. and R.C. Desrosiers, *Promise and problems associated with the use of recombinant AAV for the delivery of anti-HIV antibodies.* Molecular Therapy. Methods & Clinical Development, 2016. **3**: p. 16068.
95. Nakashima, M., et al., *Mapping Region of Human Restriction Factor APOBEC3H Critical for Interaction with HIV-1 Vif.* J Mol Biol, 2017. **429**(8): p. 1262-1276.
96. Simon JH, et al., *Virion incorporation of human immunodeficiency virus type-1 Vif is determined by intracellular expression level and may not be necessary for function.* Virology, 1998. **248**: p. 182–187.
97. Lake, J.A., et al., *The role of Vif during HIV-1 infection: interaction with novel host cellular factors.* J Clin Virol, 2003. **26**(2): p. 143-52.
98. Doyle, T., C. Goujon, and M.H. Malim, *HIV-1 and interferons: who's interfering with whom?* Nat Rev Microbiol, 2015. **13**(7): p. 403-13.
99. Li, G., M. Bukrinsky, and R.Y. Zhao, *HIV-1 viral protein R (Vpr) and its interactions with host cell.* Curr HIV Res, 2009. **7**(2): p. 178-83.
100. Mahalingam, S., et al., *Functional Analysis of the Simian Immunodeficiency Virus Vpx Protein: Identification of Packaging Determinants and a Novel Nuclear Targeting Domain.* Journal of Virology, 2001. **75**(1): p. 362-374.
101. van Weringh, A., et al., *HIV-1 modulates the tRNA pool to improve translation efficiency.* Mol Biol Evol, 2011. **28**(6): p. 1827-34.
102. Faria, N.R., et al., *HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations.* Science, 2014. **346**(6205): p. 56-61.
103. Campbell-Yesufu, O.T. and R.T. Gandhi, *Update on human immunodeficiency virus (HIV)-2 infection.* Clin Infect Dis, 2011. **52**(6): p. 780-7.
104. Hendriks, J.C., et al., *The treatment-free incubation period of AIDS in a cohort of homosexual men.* Aids, 1993. **7**(2): p. 231-9.
105. Giesecke, J., et al., *Incubation time of AIDS: progression of disease in a cohort of HIV-infected homo- and bisexual men with known dates of infection.* Scand J Infect Dis, 1990. **22**(4): p. 407-11.
106. Hallett, T.B., et al., *Embracing different approaches to estimating HIV incidence, prevalence and mortality.* AIDS, 2014. **28 Suppl 4**: p. S523-32.
107. Murray, C.J., et al., *Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.* Lancet, 2014. **384**(9947): p. 1005-70.
108. UNAIDS. *UNAIDS Fact Sheet November 2016*. 2016 [cited 2017 February 14, 2017]; Available from: http://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf.
109. (UNAIDS), J.U.N.P.o.H.A., *Global AIDS Update 2016*. 2016, UNAIDS: Geneva, Switzerland.

110. Chun, T.-W., et al., *Decay of the HIV Reservoir in Patients Receiving Antiretroviral Therapy for Extended Periods: Implications for Eradication of Virus.* The Journal of Infectious Diseases, 2007. **195**(12): p. 1762-1764.
111. Siliciano, J.D., et al., *Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells.* Nat Med, 2003. **9**(6): p. 727-8.
112. Whitney, J.B., et al., *Rapid Seeding of the Viral Reservoir Prior to SIV Viremia in Rhesus Monkeys.* Nature, 2014. **512**(7512): p. 74-77.
113. Fiebig, E.W., et al., *Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection.* AIDS, 2003. **17**(13): p. 1871-9.
114. Keele, B.F., et al., *Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection.* Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7552-7.
115. Sodora, D.L., et al., *Vaginal transmission of SIV: assessing infectivity and hormonal influences in macaques inoculated with cell-free and cell-associated viral stocks.* AIDS Res Hum Retroviruses, 1998. **14 Suppl 1**: p. S119-23.
116. Oberle, C.S., et al., *Tracing HIV-1 transmission: envelope traits of HIV-1 transmitter and recipient pairs.* Retrovirology, 2016. **13**(1): p. 62.
117. Joseph, S.B., et al., *Bottlenecks in HIV-1 transmission: insights from the study of founder viruses.* Nat Rev Micro, 2015. **13**(7): p. 414-425.
118. Salazar-Gonzalez, J.F., et al., *Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection.* The Journal of Experimental Medicine, 2009. **206**(6): p. 1273.
119. Geijtenbeek, T.B., et al., *DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells.* Cell, 2000. **100**(5): p. 587-97.
120. Moir, S., et al., *B Cells of HIV-1–Infected Patients Bind Virions through Cd21–Complement Interactions and Transmit Infectious Virus to Activated T Cells.* The Journal of Experimental Medicine, 2000. **192**(5): p. 637-646.
121. Nourmohammad, A., J. Otwinowski, and J.B. Plotkin, *Host-Pathogen Coevolution and the Emergence of Broadly Neutralizing Antibodies in Chronic Infections.* PLoS Genet, 2016. **12**(7): p. e1006171.
122. Koup, R.A., et al., *Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome.* J Virol, 1994. **68**(7): p. 4650-5.
123. Brenchley, J.M., et al., *CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract.* J Exp Med, 2004. **200**(6): p. 749-59.
124. Mattapallil, J.J., et al., *Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection.* Nature, 2005. **434**(7037): p. 1093-7.
125. Veazey, R.S., et al., *Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection.* Science, 1998. **280**(5362): p. 427-31.
126. Schuetz, A., et al., *Initiation of ART during early acute HIV infection preserves mucosal Th17 function and reverses HIV-related immune activation.* PLoS Pathog, 2014. **10**(12): p. e1004543.
127. Trono, D., et al., *HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals.* Science, 2010. **329**(5988): p. 174-80.

128. Levesque, M.C., et al., *Polyclonal B Cell Differentiation and Loss of Gastrointestinal Tract Germinal Centers in the Earliest Stages of HIV-1 Infection.* PLOS Medicine, 2009. **6**(7): p. e1000107.
129. Keating, S.M., E.S. Jacobs, and P.J. Norris, *Soluble mediators of inflammation in HIV and their implications for therapeutics and vaccine development.* Cytokine Growth Factor Rev, 2012. **23**(4-5): p. 193-206.
130. Weiss, R.A., et al., *Variable and conserved neutralization antigens of human immunodeficiency virus.* Nature, 1986. **324**(6097): p. 572-5.
131. Laird, M.E., et al., *Importance of the V1/V2 Loop Region of Simian-Human Immunodeficiency Virus Envelope Glycoprotein gp120 in Determining the Strain Specificity of the Neutralizing Antibody Response.* Journal of Virology, 2008. **82**(22): p. 11054-11065.
132. Gray, E.S., et al., *The Neutralization Breadth of HIV-1 Develops Incrementally over Four Years and Is Associated with CD4(+) T Cell Decline and High Viral Load during Acute Infection.* Journal of Virology, 2011. **85**(10): p. 4828-4840.
133. Haynes, B.F. and M.J. McElrath, *Progress in HIV-1 vaccine development.* Curr Opin HIV AIDS, 2013. **8**(4): p. 326-32.
134. Simek, M.D., et al., *Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm.* J Virol, 2009. **83**(14): p. 7337-48.
135. Tomaras, G.D., et al., *Polyclonal B cell responses to conserved neutralization epitopes in a subset of HIV-1-infected individuals.* J Virol, 2011. **85**(21): p. 11502-19.
136. Zhou, T., et al., *Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01.* Science, 2010. **329**(5993): p. 811-7.
137. Corti, D., et al., *Analysis of Memory B Cell Responses and Isolation of Novel Monoclonal Antibodies with Neutralizing Breadth from HIV-1-Infected Individuals.* PLOS ONE, 2010. **5**(1): p. e8805.
138. McLellan, J.S., et al., *Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9.* Nature, 2011. **480**(7377): p. 336-43.
139. Walker, L.M., et al., *Broad and Potent Neutralizing Antibodies from an African Donor Reveal a New HIV-1 Vaccine Target.* Science (New York, N.y.), 2009. **326**(5950): p. 285-289.
140. Bonsignori, M., et al., *Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors.* J Virol, 2011. **85**(19): p. 9998-10009.
141. Walker, L.M., et al., *Broad neutralization coverage of HIV by multiple highly potent antibodies.* Nature, 2011. **477**(7365): p. 466-70.
142. Huang, J., et al., *Broad and potent neutralization of HIV-1 by a gp41-specific human antibody.* Nature, 2012. **491**(7424): p. 406-12.
143. Morris, L., et al., *Isolation of a Human Anti-HIV gp41 Membrane Proximal Region Neutralizing Antibody by Antigen-Specific Single B Cell Sorting.* PLOS ONE, 2011. **6**(9): p. e23532.
144. Haynes, B.F., et al., *Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies.* Science, 2005. **308**(5730): p. 1906-8.

145. Verkoczy, L., et al., *Autoreactivity in an HIV-1 broadly reactive neutralizing antibody variable region heavy chain induces immunologic tolerance.* Proc Natl Acad Sci U S A, 2010. **107**(1): p. 181-6.
146. Verkoczy, L., et al., *Rescue of HIV-1 broad neutralizing antibody-expressing B cells in 2F5 VH x VL knockin mice reveals multiple tolerance controls.* J Immunol, 2011. **187**(7): p. 3785-97.
147. Easterhoff, D., et al., *Boosting of HIV envelope CD4 binding site antibodies with long variable heavy third complementarity determining region in the randomized double blind RV305 HIV-1 vaccine trial.* PLoS Pathog, 2017. **13**(2): p. e1006182.
148. Liao, H.X., et al., *Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus.* Nature, 2013. **496**(7446): p. 469-76.
149. Haynes, B.F., et al., *B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study.* Nat Biotechnol, 2012. **30**(5): p. 423-33.
150. Jardine, J.G., et al., *Minimally Mutated HIV-1 Broadly Neutralizing Antibodies to Guide Reductionist Vaccine Design.* PLOS Pathogens, 2016. **12**(8): p. e1005815.
151. Morison, L., *The global epidemiology of HIV/AIDS.* British Medical Bulletin, 2001. **58**(1): p. 7-18.
152. Cohen, M.S., et al., *The spread, treatment, and prevention of HIV-1: evolution of a global pandemic.* The Journal of Clinical Investigation, 2008. **118**(4): p. 1244-1254.
153. Etemad, B., et al., *Characterization of HIV-1 envelopes in acutely and chronically infected injection drug users.* Retrovirology, 2014. **11**.
154. Hladik, F. and M.J. McElrath, *Setting the stage: host invasion by HIV.* Nat Rev Immunol, 2008. **8**(6): p. 447-57.
155. Aceijas, C., et al., *Global overview of injecting drug use and HIV infection among injecting drug users.* Aids, 2004. **18**(17): p. 2295-303.
156. Lehman, D.A. and C. Farquhar, *Biological mechanisms of vertical human immunodeficiency virus (HIV-1) transmission.* Rev Med Virol, 2007. **17**(6): p. 381-403.
157. Gray, R.H., et al., *Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda.* Lancet, 2001. **357**(9263): p. 1149-53.
158. Boily, M.C., et al., *Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies.* Lancet Infect Dis, 2009. **9**(2): p. 118-29.
159. Baggaley, R.F., R.G. White, and M.C. Boily, *HIV transmission risk through anal intercourse: systematic review, meta-analysis and implications for HIV prevention.* Int J Epidemiol, 2010. **39**(4): p. 1048-63.
160. Jiang, J., et al., *Pre-Exposure Prophylaxis for the Prevention of HIV Infection in High Risk Populations: A Meta-Analysis of Randomized Controlled Trials.* PLOS ONE, 2014. **9**(2): p. e87674.
161. Samji, H., et al., *Closing the Gap: Increases in Life Expectancy among Treated HIV-Positive Individuals in the United States and Canada.* PLOS ONE, 2013. **8**(12): p. e81355.
162. Rasmussen, L.D., et al., *Time trends for risk of severe age-related diseases in individuals with and without HIV infection in Denmark: a nationwide population-based cohort study.* The Lancet HIV. **2**(7): p. e288-e298.

163. Group, I.S.S., et al., *Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection.* N Engl J Med, 2015. **373**(9): p. 795-807.
164. UNAIDS, *GLOBAL REPORT: UNAIDS report on the global AIDS epidemic 2013*. 2013, Joint United Nations Programme on HIV/AIDS (UNAIDS): Geneva, Switzerland. p. 194.
165. Schwartländer, B., et al., *Towards an improved investment approach for an effective response to HIV/AIDS.* The Lancet. **377**(9782): p. 2031-2041.
166. Stover, J., et al., *How can we get close to zero? The potential contribution of biomedical prevention and the investment framework towards an effective response to HIV.* PLoS One, 2014. **9**(11): p. e111956.
167. Foundation, T.H.J.K.F. and UNAIDS, *Financing the Response to HIV in Low- and Middle-Income Countries: International Assistance from Donor Governments in 2014.* 2015.
168. Harmon, T.M., et al., *Exploring the Potential Health Impact and Cost-Effectiveness of AIDS Vaccine within a Comprehensive HIV/AIDS Response in Low- and Middle-Income Countries.* PLoS One, 2016. **11**(1): p. e0146387.
169. Marcus, J.L., et al., *Potential Interventions to Support Adherence to HIV Preexposure Prophylaxis (PrEP): A Systematic Review.* HIV medicine, 2014. **15**(7): p. 385-395.
170. Grant, R.M., et al., *Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men.* The New England journal of medicine, 2010. **363**(27): p. 2587-2599.
171. Baeten, J.M., et al., *Antiretroviral Prophylaxis for HIV-1 Prevention among Heterosexual Men and Women.* The New England journal of medicine, 2012. **367**(5): p. 399-410.
172. Thigpen, M.C., et al., *Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana.* N Engl J Med, 2012. **367**(5): p. 423-34.
173. Van Damme, L., et al., *Preexposure Prophylaxis for HIV Infection among African Women.* The New England journal of medicine, 2012. **367**(5): p. 411-422.
174. Abdool Karim, Q., et al., *Effectiveness and Safety of Tenofovir Gel, an Antiretroviral Microbicide, for the Prevention of HIV Infection in Women.* Science (New York, N.Y.), 2010. **329**(5996): p. 1168-1174.
175. Klatt, N.R., et al., *Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women.* Science, 2017. **356**(6341): p. 938-945.
176. van de Wijgert, J. and S. McCormack, *Vaginal dysbiosis and pre-exposure prophylaxis efficacy.* Lancet HIV, 2017. **4**(10): p. e427-e429.
177. Arnaout, R.A., et al., *A simple relationship between viral load and survival time in HIV-1 infection.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(20): p. 11549-11553.
178. Langford, S.E., J. Ananworanich, and D.A. Cooper, *Predictors of disease progression in HIV infection: a review.* AIDS Research and Therapy, 2007. **4**: p. 11-11.
179. Quinn, T.C., et al., *Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1.* New England Journal of Medicine, 2000. **342**(13): p. 921-929.
180. Alchin, D.R., *HIV vaccine development: an exploratory review of the trials and tribulations.* Immunol Res, 2014. **60**(1): p. 35-7.

181. McMichael, A.J., *HIV VACCINES.* Annual Review of Immunology, 2006. **24**(1): p. 227-255.
182. Persaud, D., et al., *Absence of detectable HIV-1 viremia after treatment cessation in an infant.* N Engl J Med, 2013. **369**(19): p. 1828-35.
183. Ledford, H., *HIV rebound dashes hope of cure.* Nature, 2014.
184. Check Hayden, E., *Infant's vanquished HIV leaves doctors puzzled.* Nature, 2013.
185. Hansen, S.G., et al., *Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge.* Nat Med, 2009. **15**(3): p. 293-9.
186. Plotkin, S.A. and S.A. Plotkin, *Correlates of Vaccine-Induced Immunity.* Clinical Infectious Diseases, 2008. **47**(3): p. 401-409.
187. Hansen, S.G., et al., *Immune clearance of highly pathogenic SIV infection.* Nature, 2013. **502**(7469): p. 100-4.
188. Cardo, D.M., et al., *A Case–Control Study of HIV Seroconversion in Health Care Workers after Percutaneous Exposure.* New England Journal of Medicine, 1997. **337**(21): p. 1485-1490.
189. Tsai, C.C., et al., *Effectiveness of postinoculation (R)-9-(2-phosphonylmethoxypropyl) adenine treatment for prevention of persistent simian immunodeficiency virus SIVmne infection depends critically on timing of initiation and duration of treatment.* J Virol, 1998. **72**(5): p. 4265-73.
190. Tsai, C.C., et al., *Prevention of SIV infection in macaques by (R)-9-(2-phosphonylmethoxypropyl)adenine.* Science, 1995. **270**(5239): p. 1197-9.
191. Ramos, A., et al., *Intersubtype Human Immunodeficiency Virus Type 1 Superinfection following Seroconversion to Primary Infection in Two Injection Drug Users.* Journal of Virology, 2002. **76**(15): p. 7444-7452.
192. Jost, S., et al., *A patient with HIV-1 superinfection.* N Engl J Med, 2002. **347**(10): p. 731-6.
193. Altfeld, M., et al., *HIV-1 superinfection despite broad CD8+ T-cell responses containing replication of the primary virus.* Nature, 2002. **420**(6914): p. 434-9.
194. Koelsch, K.K., et al., *Clade B HIV-1 superinfection with wild-type virus after primary infection with drug-resistant clade B virus.* Aids, 2003. **17**(7): p. F11-6.
195. Kaul, R., et al., *Late seroconversion in HIV-resistant Nairobi prostitutes despite pre-existing HIV-specific CD8(+) responses.* Journal of Clinical Investigation, 2001. **107**(3): p. 341-349.
196. Kaul, R., et al., *HIV-1-specific mucosal CD8+ lymphocyte responses in the cervix of HIV-1-resistant prostitutes in Nairobi.* J Immunol, 2000. **164**(3): p. 1602-11.
197. MacDonald, K.S., et al., *Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection.* J Infect Dis, 2000. **181**(5): p. 1581-9.
198. Gamble, L.J. and Q.L. Matthews, *Current progress in the development of a prophylactic vaccine for HIV-1.* Drug Design, Development and Therapy, 2011. **5**: p. 9-26.
199. Picker, L.J., S.G. Hansen, and J.D. Lifson, *New paradigms for HIV/AIDS vaccine development.* Annu Rev Med, 2012. **63**: p. 95-111.
200. Hansen, S.G., et al., *Cytomegalovirus vectors violate CD8+ T cell epitope recognition paradigms.* Science, 2013. **340**(6135): p. 1237874.

201. McMichael, A.J. and T. Hanke, *HIV vaccines 1983-2003.* Nat Med, 2003. **9**(7): p. 874-80.
202. Richman, D.D., et al., *Rapid evolution of the neutralizing antibody response to HIV type 1 infection.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(7): p. 4144-4149.
203. Wei, X., et al., *Antibody neutralization and escape by HIV-1.* Nature, 2003. **422**(6929): p. 307-12.
204. Moore, P.L., et al., *Nature of Nonfunctional Envelope Proteins on the Surface of Human Immunodeficiency Virus Type 1.* Journal of Virology, 2006. **80**(5): p. 2515-2528.
205. Zhu, P., et al., *Distribution and three-dimensional structure of AIDS virus envelope spikes.* Nature, 2006. **441**(7095): p. 847-52.
206. Schiller, J. and B. Chackerian, *Why HIV Virions Have Low Numbers of Envelope Spikes: Implications for Vaccine Development.* PLOS Pathogens, 2014. **10**(8): p. e1004254.
207. McCoy, L.E. and R.A. Weiss, *Neutralizing antibodies to HIV-1 induced by immunization.* J Exp Med, 2013. **210**(2): p. 209-23.
208. Pace, C.S., et al., *Bispecific antibodies directed to CD4 domain 2 and HIV envelope exhibit exceptional breadth and picomolar potency against HIV-1.* Proc Natl Acad Sci U S A, 2013. **110**(33): p. 13540-5.
209. Stiegler, G., et al., *A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1.* AIDS Res Hum Retrovir, 2001. **17**.
210. Gautam, R., et al., *A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges.* Nature, 2016. **533**(7601): p. 105-109.
211. Moldt, B., et al., *Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo.* Proc Natl Acad Sci U S A, 2012. **109**(46): p. 18921-5.
212. Shingai, M., et al., *Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques.* J Exp Med, 2014. **211**(10): p. 2061-74.
213. Nishimura, Y., et al., *Determination of a Statistically Valid Neutralization Titer in Plasma That Confers Protection against Simian-Human Immunodeficiency Virus Challenge following Passive Transfer of High-Titered Neutralizing Antibodies.* Journal of Virology, 2002. **76**(5): p. 2123-2130.
214. Bricault, C.A., et al., *A Multivalent Clade C HIV-1 Env Trimer Cocktail Elicits a Higher Magnitude of Neutralizing Antibodies than Any Individual Component.* J Virol, 2014.
215. McGuire, A.T., et al., *Specifically modified Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice.* Nat Commun, 2016. **7**: p. 10618.
216. Tenbusch, M., et al., *Risk of immunodeficiency virus infection may increase with vaccine-induced immune response.* J Virol, 2012. **86**(19): p. 10533-9.
217. Hu, H., et al., *Preferential infection of human Ad5-specific CD4 T cells by HIV in Ad5 naturally exposed and recombinant Ad5-HIV vaccinated individuals.* Proceedings of

the National Academy of Sciences of the United States of America, 2014. **111**(37): p. 13439-13444.
218. Martins, M.A., et al., *Rare Control of SIVmac239 Infection in a Vaccinated Rhesus Macaque.* AIDS Res Hum Retroviruses, 2017.
219. Siegrist, C.-A., *Vaccine immunology*, in *Vaccines*, S. Plotkin, W. Orenstein, and P. Offit, Editors. 2013, Elsevier: Philadelphia, PA. p. 17–36.
220. Miller, J.D., et al., *Human effector and memory CD8+ T cell responses to smallpox and yellow fever vaccines.* Immunity, 2008. **28**(5): p. 710-22.
221. Gilbert, P.B., et al., *Statistical Interpretation of the RV144 HIV Vaccine Efficacy Trial in Thailand: A Case Study for Statistical Issues in Efficacy Trials.* The Journal of Infectious Diseases, 2011. **203**(7): p. 969-975.
222. Pugach, P., et al., *A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene.* J Virol, 2015. **89**(6): p. 3380-95.
223. Sanders, R.W., et al., *A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies.* PLOS Pathogens, 2013. **9**(9): p. e1003618.
224. Schnepp, B.C. and P.R. Johnson, *Adeno-Associated Virus Delivery of Broadly Neutralizing Antibodies.* Current opinion in HIV and AIDS, 2014. **9**(3): p. 250-256.
225. Fuchs, S.P., et al., *AAV-Delivered Antibody Mediates Significant Protective Effects against SIVmac239 Challenge in the Absence of Neutralizing Activity.* PLOS Pathogens, 2015. **11**(8): p. e1005090.
226. Bischof, G.F., et al., *Use of a gamma-2 herpesvirus as a vector to deliver antibodies to rhesus monkeys.* Gene Ther, 2017.
227. Plotkin, S.L. and S.A. Plotkin, *A short history of vaccination.* 2013: p. 1-13.
228. Hofmann-Lehmann, R., et al., *Live attenuated, nef-deleted SIV is pathogenic in most adult macaques after prolonged observation.* Aids, 2003. **17**(2): p. 157-66.
229. Franchini, G., et al., *Poxvirus-based vaccine candidates for HIV: two decades of experience with special emphasis on canarypox vectors.* Expert Rev Vaccines, 2004. **3**(4 Suppl): p. S75-88.
230. Vaccari, M., et al., *CD4+ T-cell loss and delayed expression of modulators of immune responses at mucosal sites of vaccinated macaques following SIVmac251 infection.* Mucosal Immunology, 2008. **1**: p. 497.
231. Manrique, M., et al., *DNA-MVA vaccine protection after X4 SHIV challenge in macaques correlates with day-of-challenge antiviral CD4+ cell-mediated immunity levels and postchallenge preservation of CD4+ T cell memory.* AIDS Res Hum Retroviruses, 2008. **24**(3): p. 505-19.
232. Hel, Z., et al., *Containment of simian immunodeficiency virus infection in vaccinated macaques: correlation with the magnitude of virus-specific pre- and postchallenge CD4+ and CD8+ T cell responses.* J Immunol, 2002. **169**(9): p. 4778-87.
233. Draper, S.J., M.G. Cottingham, and S.C. Gilbert, *Utilizing poxviral vectored vaccines for antibody induction—Progress and prospects.* Vaccine, 2013. **31**(39): p. 4223-4230.
234. Pialoux, G., et al., *A prime-boost approach to HIV preventive vaccine using a recombinant canarypox virus expressing glycoprotein 160 (MN) followed by a recombinant glycoprotein 160 (MN/LAI). The AGIS Group, and l'Agence Nationale de Recherche sur le SIDA.* AIDS Res Hum Retroviruses, 1995. **11**(3): p. 373-81.

235. Kristensen, N.N., et al., *Cytokine production by virus-specific CD8(+) T cells varies with activation state and localization, but not with TCR avidity.* J Gen Virol, 2004. **85**(Pt 6): p. 1703-12.
236. Paliard, X., et al., *Simultaneous production of IL-2, IL-4, and IFN-gamma by activated human CD4+ and CD8+ T cell clones.* J Immunol, 1988. **141**(3): p. 849-55.
237. Conlon, K., et al., *Comparison of lymphokine secretion and mRNA expression in the CD45RA+ and CD45RO+ subsets of human peripheral blood CD4+ and CD8+ lymphocytes.* Eur J Immunol, 1995. **25**(3): p. 644-8.
238. Borrow, P., et al., *Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection.* Journal of Virology, 1994. **68**(9): p. 6103-6110.
239. McDermott, A.B. and R.A. Koup, *CD8(+) T cells in preventing HIV infection and disease.* Aids, 2012. **26**(10): p. 1281-92.
240. Schmitz, J.E., et al., *Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes.* Science, 1999. **283**(5403): p. 857-60.
241. Jin, X., et al., *Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques.* J Exp Med, 1999. **189**(6): p. 991-8.
242. Douek, D., M. Roederer, and R. Koup, *Emerging concepts in the immunopathogenesis of AIDS.* Annual review of medicine, 2009. **60**: p. 471-484.
243. Zhang, J.Y., et al., *PD-1 up-regulation is correlated with HIV-specific memory CD8+ T-cell exhaustion in typical progressors but not in long-term nonprogressors.* Blood, 2007. **109**(11): p. 4671-8.
244. Hansen, S.G., et al., *Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine.* Nature, 2011. **473**(7348): p. 523-7.
245. Hansen, S.G., et al., *Broadly targeted CD8(+) T cell responses restricted by major histocompatibility complex E.* Science, 2016. **351**(6274): p. 714-20.
246. Haynes, B.F., et al., *Immune-Correlates Analysis of an HIV-1 Vaccine Efficacy Trial.* New England Journal of Medicine, 2012. **366**(14): p. 1275-1286.
247. Shiver, J.W., et al., *Replication-incompetent adenoviral vaccine vector elicits effective anti-immunodeficiency-virus immunity.* Nature, 2002. **415**(6869): p. 331-5.
248. Bogers, W.M., et al., *Systemic neutralizing antibodies induced by long interval mucosally primed systemically boosted immunization correlate with protection from mucosal SHIV challenge.* Virology, 2008. **382**(2): p. 217-25.
249. Demberg, T., et al., *A replication-competent adenovirus-human immunodeficiency virus (Ad-HIV) tat and Ad-HIV env priming/Tat and envelope protein boosting regimen elicits enhanced protective efficacy against simian/human immunodeficiency virus SHIV89.6P challenge in rhesus macaques.* J Virol, 2007. **81**(7): p. 3414-27.
250. Hidajat, R., et al., *Correlation of vaccine-elicited systemic and mucosal nonneutralizing antibody activities with reduced acute viremia following intrarectal simian immunodeficiency virus SIVmac251 challenge of rhesus macaques.* J Virol, 2009. **83**(2): p. 791-801.
251. Wilson, N.A., et al., *Vaccine-induced cellular immune responses reduce plasma viral concentrations after repeated low-dose challenge with pathogenic simian immunodeficiency virus SIVmac239.* J Virol, 2006. **80**(12): p. 5875-85.

252. Casimiro, D.R., et al., *Attenuation of simian immunodeficiency virus SIVmac239 infection by prophylactic immunization with dna and recombinant adenoviral vaccine vectors expressing Gag.* J Virol, 2005. **79**(24): p. 15547-55.
253. Hammer, S.M., et al., *Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine.* N Engl J Med, 2013. **369**(22): p. 2083-92.
254. Qureshi, H., et al., *Low-dose penile SIVmac251 exposure of rhesus macaques infected with adenovirus type 5 (Ad5) and then immunized with a replication-defective Ad5-based SIV gag/pol/nef vaccine recapitulates the results of the phase IIb step trial of a similar HIV-1 vaccine.* J Virol, 2012. **86**(4): p. 2239-50.
255. Rerks-Ngarm, S., et al., *Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand.* N Engl J Med, 2009. **361**(23): p. 2209-20.
256. Liao, H.X., et al., *Vaccine induction of antibodies against a structurally heterogeneous site of immune pressure within HIV-1 envelope protein variable regions 1 and 2.* Immunity, 2013. **38**(1): p. 176-86.
257. Grose, C., *Pangaea and the Out-of-Africa Model of Varicella-Zoster Virus Evolution and Phylogeography.* J Virol, 2012. **86**(18): p. 9558-65.
258. Cicin-Sain, L., et al., *Cytomegalovirus-specific T cell immunity is maintained in immunosenescent rhesus macaques.* J Immunol, 2011. **187**(4): p. 1722-32.
259. Barton, E.S., et al., *Herpesvirus latency confers symbiotic protection from bacterial infection.* Nature, 2007. **447**(7142): p. 326-9.
260. Tao, L. and T.A. Reese, *Making Mouse Models That Reflect Human Immune Responses.* Trends Immunol, 2017. **38**(3): p. 181-193.
261. Kuparinen, T., et al., *Cytomegalovirus (CMV)-dependent and -independent changes in the aging of the human immune system: a transcriptomic analysis.* Exp Gerontol, 2013. **48**(3): p. 305-12.
262. Francois, S., et al., *NK cells improve control of friend virus infection in mice persistently infected with murine cytomegalovirus.* Retrovirology, 2013. **10**(1): p. 58.
263. Cannon, M.J., D.S. Schmid, and T.B. Hyde, *Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection.* Rev Med Virol, 2010. **20**(4): p. 202-13.
264. Dollard, S.C., S.D. Grosse, and D.S. Ross, *New estimates of the prevalence of neurological and sensory sequelae and mortality associated with congenital cytomegalovirus infection.* Reviews in Medical Virology, 2007. **17**(5): p. 355-363.
265. Bale, J.F., Jr., *Cytomegalovirus infections.* Semin Pediatr Neurol, 2012. **19**(3): p. 101-6.
266. Bate, S.L., S.C. Dollard, and M.J. Cannon, *Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988-2004.* Clin Infect Dis, 2010. **50**(11): p. 1439-47.
267. Staras, S.A., et al., *Seroprevalence of cytomegalovirus infection in the United States, 1988-1994.* Clin Infect Dis, 2006. **43**(9): p. 1143-51.
268. Schleiss, M.R., *Developing a Vaccine against Congenital Cytomegalovirus (CMV) Infection: What Have We Learned from Animal Models? Where Should We Go Next?* Future Virol, 2013. **8**(12): p. 1161-1182.
269. Vochem, M., et al., *Cytomegalovirus infection of preterm infants by breast milk.* Monatsschrift Kinderheilkunde, 1997. **145**(6): p. 619-625.

270. Josephson, C.D., et al., *Blood transfusion and breast milk transmission of cytomegalovirus in very low-birth-weight infants: a prospective cohort study.* JAMA Pediatr, 2014. **168**(11): p. 1054-62.
271. Balcells, C., et al., *Vertically transmitted cytomegalovirus infection in newborn preterm infants.* Journal of Perinatal Medicine, 2016. **44**(5): p. 485-490.
272. Murph, J.R. and J.F. Bale, Jr., *The natural history of acquired cytomegalovirus infection among children in group day care.* Am J Dis Child, 1988. **142**(8): p. 843-6.
273. Handsfield, H.H., et al., *Cytomegalovirus Infection in Sex Partners: Evidence for Sexual Transmission.* The Journal of Infectious Diseases, 1985. **151**(2): p. 344-348.
274. Mocarski, E.S., T. Shenk, and R. Pass, *Cytomegaloviruses*, in *Feilds Virology*, D.M. Knipe and P. Howley, Editors. 2007, Lippincott, Williams & Wilkins: Philadelphia, PA. p. 2702-2757.
275. Hansen, S.G., et al., *Evasion of CD8+ T cells is critical for superinfection by cytomegalovirus.* Science, 2010. **328**(5974): p. 102-6.
276. Shenk, T.E. and M.F. Stinski, eds. *Human Cytomegalovirus*. Current Topics in Microbiology and Immunology, ed. R.W. Compans, et al. Vol. 325. 2008, Springer: Berlin.
277. Faure-Della Corte, M., et al., *Variability and recombination of clinical human cytomegalovirus strains from transplantation recipients.* J Clin Virol, 2010. **47**(2): p. 161-9.
278. Russell, J.N., et al., *A novel strain of cynomolgus macaque cytomegalovirus: implications for host-virus co-evolution.* BMC Genomics, 2016. **17**(1): p. 277.
279. Murphy, E., et al., *Reevaluation of human cytomegalovirus coding potential.* Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13585-90.
280. Davison, A.J., et al., *The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome.* J Gen Virol, 2003. **84**(Pt 1): p. 17-28.
281. Marsh, A.K., *Characterization and Development of a Novel Cytomegalovirus for Evaluation as an HIV Vaccine Viral Vector in a Non-Human Primate Model*, in *Department of Immunology*. 2013, University of Toronto: Toronto. p. 289.
282. Malouli, D., et al., *Reevaluation of the coding potential and proteomic analysis of the BAC-derived rhesus cytomegalovirus strain 68-1.* J Virol, 2012. **86**(17): p. 8959-73.
283. Jean Beltran, P.M. and I.M. Cristea, *The life cycle and pathogenesis of human cytomegalovirus infection: lessons from proteomics.* Expert Rev Proteomics, 2014. **11**(6): p. 697-711.
284. Varnum, S.M., et al., *Identification of Proteins in Human Cytomegalovirus (HCMV) Particles: the HCMV Proteome.* Journal of Virology, 2004. **78**(20): p. 10960-10966.
285. Kalejta, R.F., *Tegument proteins of human cytomegalovirus.* Microbiol Mol Biol Rev, 2008. **72**(2): p. 249-65, table of contents.
286. Saydam, O., D.L. Glauser, and C. Fraefel, *Construction and packaging of herpes simplex virus/adeno-associated virus (HSV/AAV) Hybrid amplicon vectors.* Cold Spring Harb Protoc, 2012. **2012**(3): p. 352-6.
287. Mcvoy, M.A. and S.P. Adler, *Human Cytomegalovirus DNA Replicates after Early Circularization by Concatemer Formation, and Inversion Occurs within the Concatemer.* Journal of Virology, 1994. **68**(2): p. 1040-1051.
288. Hansen, S.G., et al., *Complete sequence and genomic analysis of rhesus cytomegalovirus.* J Virol, 2003. **77**(12): p. 6620-36.

289. Marsh, A.K., et al., *Genomic sequencing and characterization of cynomolgus macaque cytomegalovirus.* J Virol, 2011. **85**(24): p. 12995-3009.
290. Davison, A., et al., *Comparatve Genomics of Primate Cytomegaloviruses*, in *Cytomegaloviruses: From Molecular Pathogenesis to Intervention*, M.J. Reddehase, Editor. 2013, Caister Academic Publishing: Norfolk UK. p. 1-22.
291. Marsh, A.K., et al., *Examining the species-specificity of rhesus macaque cytomegalovirus (RhCMV) in cynomolgus macaques.* PLoS One, 2015. **10**(3): p. e0121339.
292. Powers, C. and K. Fruh, *Rhesus CMV: an emerging animal model for human CMV.* Med Microbiol Immunol, 2008. **197**(2): p. 109-15.
293. Schleiss, M.R., *Animal models of congenital cytomegalovirus infection: an overview of progress in the characterization of guinea pig cytomegalovirus (GPCMV).* J Clin Virol, 2002. **25 Suppl 2**: p. S37-49.
294. Leendertz, F.H., et al., *Novel cytomegaloviruses in free-ranging and captive great apes: phylogenetic evidence for bidirectional horizontal transmission.* J Gen Virol, 2009. **90**(Pt 10): p. 2386-94.
295. Sinzger, C., M. Digel, and G. Jahn, *Cytomegalovirus Cell Tropism*, in *Human Cytomegalovirus*, T. Shenk and S. M, Editors. 2008, Springer-Verlag: Berlin Heidelberg. p. 63-83.
296. Soderberg-Naucler, C., K.N. Fish, and J.A. Nelson, *Reactivation of latent human cytomegalovirus by allogeneic stimulation of blood cells from healthy donors.* Cell, 1997. **91**(1): p. 119-26.
297. Fish, K.N., et al., *Human Cytomegalovirus Persistently Infects Aortic Endothelial Cells.* Journal of Virology, 1998. **72**(7): p. 5661-5668.
298. Rue, C.A., et al., *A cyclooxygenase-2 homologue encoded by rhesus cytomegalovirus is a determinant for endothelial cell tropism.* J Virol, 2004. **78**(22): p. 12529-36.
299. van der Strate, B.W.A., et al., *Dissemination of Rat Cytomegalovirus through Infected Granulocytes and Monocytes In Vitro and In Vivo.* Journal of Virology, 2003. **77**(20): p. 11274-11278.
300. Grefte, A., et al., *Presence of human cytomegalovirus (HCMV) immediate early mRNA but not ppUL83 (lower matrix protein pp65) mRNA in polymorphonuclear and mononuclear leukocytes during active HCMV infection.* J Gen Virol, 1994. **75 (Pt 8)**: p. 1989-98.
301. Gerna, G., et al., *Human Cytomegalovirus Replicates Abortively in Polymorphonuclear Leukocytes after Transfer from Infected Endothelial Cells via Transient Microfusion Events.* Journal of Virology, 2000. **74**(12): p. 5629-5638.
302. Rice, G.P., R.D. Schrier, and M.B. Oldstone, *Cytomegalovirus infects human lymphocytes and monocytes: virus expression is restricted to immediate-early gene products.* Proceedings of the National Academy of Sciences of the United States of America, 1984. **81**(19): p. 6134-6138.
303. Wang, D. and T. Shenk, *Human cytomegalovirus virion protein complex required for epithelial and endothelial cell tropism.* Proc Natl Acad Sci U S A, 2005. **102**(50): p. 18153-8.
304. Li, G. and J.P. Kamil, *Viral Regulation of Cell Tropism in Human Cytomegalovirus.* J Virol, 2015. **90**(2): p. 626-9.

305. Wille, P.T., et al., *Human Cytomegalovirus (HCMV) Glycoprotein gB Promotes Virus Entry In Trans Acting as the Viral Fusion Protein Rather than as a Receptor-Binding Protein.* mBio, 2013. **4**(3): p. e00332-13.
306. Ogawa-Goto, K., et al., *Microtubule Network Facilitates Nuclear Targeting of Human Cytomegalovirus Capsid.* Journal of Virology, 2003. **77**(15): p. 8541-8547.
307. Kalejta, R.F., *Functions of Human Cytomegalovirus Tegument Proteins Prior to Immediate Early Gene Expression*, in *Human Cytomegalovirus*, T.E. Shenk and M.F. Stinski, Editors. 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 101-115.
308. Arcangeletti, M.-C., et al., *Human cytomegalovirus reactivation from latency: validation of a "switch" model in vitro.* Virology Journal, 2016. **13**(1): p. 179.
309. La Rosa, C. and D.J. Diamond, *The immune response to human CMV.* Future Virol, 2012. **7**(3): p. 279-293.
310. Reeves, M. and J. Sinclair, *Aspects of Human Cytomegalovirus Latency and Reactivation*, in *Human Cytomegalovirus*, T. Shenk and S. M, Editors. 2008, Springer-Verlag: berlin Heidelberg. p. 297-314.
311. Cheung, A.K., et al., *Viral gene expression during the establishment of human cytomegalovirus latent infection in myeloid progenitor cells.* Blood, 2006. **108**(12): p. 3691-9.
312. Sylwester, A.W., et al., *Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects.* J Exp Med, 2005. **202**(5): p. 673-85.
313. Boppana, S.B. and W.J. Britt, *Synopsis of Clinical Aspects of Human Cytomegalovirus Disease*, in *Cytomegaloviruses: From Molecular Pathogenesis to Intervention*, M.J. Reddehase, Editor. 2013, Caister Academic Press: Norfolk, UK. p. 1-26.
314. Gail, J.D., *Infectious Diseases Society of America and Centers for Disease Control: Summary of a Workshop on Surveillance for Congenital Cytomegalovirus Disease.* Reviews of Infectious Diseases, 1991. **13**(2): p. 315-329.
315. Stagno, S., et al., *Primary cytomegalovirus infection in pregnancy. Incidence, transmission to fetus, and clinical outcome.* JAMA, 1986. **256**(14): p. 1904-8.
316. Schleiss, M.R. and T.C. Heineman, *Progress toward an elusive goal: current status of cytomegalovirus vaccines.* Expert Rev Vaccines, 2005. **4**(3): p. 381-406.
317. Fowler, K.B., et al., *The outcome of congenital cytomegalovirus infection in relation to maternal antibody status.* N Engl J Med, 1992. **326**(10): p. 663-7.
318. Fowler, K.B., S. Stagno, and R.F. Pass, *Interval between births and risk of congenital cytomegalovirus infection.* Clin Infect Dis, 2004. **38**(7): p. 1035-7.
319. Fowler, K.B., *Maternal Immunity and Prevention of Congenital Cytomegalovirus Infection.* Jama, 2003. **289**(8): p. 1008.
320. Kimberlin, D.W., et al., *Effect of ganciclovir therapy on hearing in symptomatic congenital cytomegalovirus disease involving the central nervous system: a randomized, controlled trial.* J Pediatr, 2003. **143**(1): p. 16-25.
321. Oliver, S.E., et al., *Neurodevelopmental outcomes following ganciclovir therapy in symptomatic congenital cytomegalovirus infections involving the central nervous system.* J Clin Virol, 2009. **46 Suppl 4**: p. S22-6.

322. Lanzieri, T.M., et al., *Cytomegalovirus infection among infants in California neonatal intensive care units, 2005–2010.* Journal of Perinatal Medicine, 2014. **42**(3): p. 393-399.
323. Dunn, H.S., et al., *Dynamics of CD4 and CD8 T cell responses to cytomegalovirus in healthy human donors.* J Infect Dis, 2002. **186**(1): p. 15-22.
324. Takano, T., M. Zielenska, and L.E. Becker, *Cytomegalovirus encephalitis in a child with adenosine deaminase-deficient severe combined immune deficiency: a neuropathologic study.* Neuropediatrics, 1998. **29**(3): p. 151-4.
325. Orlikowski, D., et al., *Guillain-Barre syndrome following primary cytomegalovirus infection: a prospective cohort study.* Clin Infect Dis, 2011. **52**(7): p. 837-44.
326. Wrzolek, M.A., et al., *Opportunistic infections of the central nervous system in children with HIV infection: report of 9 autopsy cases and review of literature.* Clin Neuropathol, 1995. **14**(4): p. 187-96.
327. Boppana, S. and W.J. Britt, *Synopsis of clinical aspects of human cytomegalovirus disease*, in *Cytomegaloviruses: Drom Molecular Pathogenesis to Intervention*, M.J. Reddehase, Editor. 2013, Caister Academic Press: Norfolk, UK. p. 1-25.
328. Sung, H. and M.R. Schleiss, *Update on the current status of cytomegalovirus vaccines.* Expert Rev Vaccines, 2010. **9**(11): p. 1303-14.
329. Plotkin, S.A., et al., *Candidate cytomegalovirus strain for human vaccination.* Infect Immun, 1975. **12**(3): p. 521-7.
330. Plotkin, S.A., et al., *Multicenter trial of Towne strain attenuated virus vaccine in seronegative renal transplant recipients.* Transplantation, 1994. **58**(11): p. 1176-8.
331. Adler, S.P., et al., *Immunity induced by primary human cytomegalovirus infection protects against secondary infection among women of childbearing age.* J Infect Dis, 1995. **171**(1): p. 26-32.
332. Boehme, K.W., M. Guerrero, and T. Compton, *Human cytomegalovirus envelope glycoproteins B and H are necessary for TLR2 activation in permissive cells.* J Immunol, 2006. **177**(10): p. 7094-102.
333. Xagorari, A. and K. Chlichlia, *Toll-Like Receptors and Viruses: Induction of Innate Antiviral Immune Responses.* The Open Microbiology Journal, 2008. **2**: p. 49-59.
334. Reddehase, M.J., *Antigens and immunoevasins: opponents in cytomegalovirus immune surveillance.* Nat Rev Immunol, 2002. **2**(11): p. 831-44.
335. Lockridge, K.M., et al., *Primate cytomegaloviruses encode and express an IL-10-like protein.* Virology, 2000. **268**(2): p. 272-80.
336. Kotenko, S.V., et al., *Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10).* Proc Natl Acad Sci U S A, 2000. **97**(4): p. 1695-700.
337. Wall, N.A., et al., *Cytomegalovirus seropositivity is associated with increased arterial stiffness in patients with chronic kidney disease.* PLoS One, 2013. **8**(2): p. e55686.
338. Firth, C., et al., *Cytomegalovirus infection is associated with an increase in systolic blood pressure in older individuals.* Qjm, 2016. **109**(9): p. 595-600.
339. Moss, P., *CMV-driven immunosenescence : mechanisms and management*, in *6th International congenital CMV conference / 16th International CMV/betaherpesvirus workshop*. 2017: Noordwijkerhout, the Netherlands.

340. Fujita, T., et al., *Expansion of dysfunctional Tim-3-expressing effector memory CD8+ T cells during simian immunodeficiency virus infection in rhesus macaques.* J Immunol, 2014. **193**(11): p. 5576-83.
341. Barber, D.L., et al., *Restoring function in exhausted CD8 T cells during chronic viral infection.* Nature, 2006. **439**(7077): p. 682-7.
342. Gillespie, G.M., et al., *Functional heterogeneity and high frequencies of cytomegalovirus-specific CD8(+) T lymphocytes in healthy seropositive donors.* J Virol, 2000. **74**(17): p. 8140-50.
343. van de Berg, P.J., et al., *A fingerprint left by cytomegalovirus infection in the human T cell compartment.* J Clin Virol, 2008. **41**(3): p. 213-7.
344. Hansen, S.G., et al., *Prevention of tuberculosis in rhesus macaques by a cytomegalovirus-based vaccine.* Nat Med, 2018. **24**(2): p. 130-143.
345. Willer, D.O., et al. *Protection of Cynomolgus Macaques from Pathogenic SIV Following Vaccination with Varicella-Zoster Virus Based Vaccines*. in *AIDS Vaccine*. 2013. Barcelona, Spain.
346. Martins, M.A., et al., *Rare Control of SIVmac239 Infection in a Vaccinated Rhesus Macaque.* AIDS Res Hum Retroviruses, 2017. **33**(8): p. 843-858.
347. Debes, G.F., et al., *Chemokine receptor CCR7 required for T lymphocyte exit from peripheral tissues.* Nat Immunol, 2005. **6**(9): p. 889-94.
348. Bromley, S.K., S.Y. Thomas, and A.D. Luster, *Chemokine receptor CCR7 guides T cell exit from peripheral tissues and entry into afferent lymphatics.* Nat Immunol, 2005. **6**(9): p. 895-901.
349. Smith, C.J., H. Turula, and C.M. Snyder, *Systemic hematogenous maintenance of memory inflation by MCMV infection.* PLoS Pathog, 2014. **10**(7): p. e1004233.
350. Gerlach, C., et al., *The Chemokine Receptor CX3CR1 Defines Three Antigen-Experienced CD8 T Cell Subsets with Distinct Roles in Immune Surveillance and Homeostasis.* Immunity, 2016. **45**(6): p. 1270-1284.
351. Sathaliyawala, T., et al., *Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets.* Immunity, 2013. **38**(1): p. 187-97.
352. Osborn, J.F., et al., *Enzymatic synthesis of core 2 O-glycans governs the tissue-trafficking potential of memory CD8(+) T cells.* Sci Immunol, 2017. **2**(16): p. eaan6049.
353. Thome, J.J., et al., *Spatial map of human T cell compartmentalization and maintenance over decades of life.* Cell, 2014. **159**(4): p. 814-28.
354. Iijima, N. and A. Iwasaki, *T cell memory. A local macrophage chemokine network sustains protective tissue-resident memory CD4 T cells.* Science, 2014. **346**(6205): p. 93-8.
355. Masopust, D., et al., *Preferential localization of effector memory cells in nonlymphoid tissue.* Science, 2001. **291**(5512): p. 2413-7.
356. Smith, C.J., et al., *Murine CMV Infection Induces the Continuous Production of Mucosal Resident T Cells.* Cell Rep, 2015. **13**(6): p. 1137-1148.
357. Anderson, K.G., et al., *Cutting edge: intravascular staining redefines lung CD8 T cell responses.* J Immunol, 2012. **189**(6): p. 2702-6.
358. Knickelbein, J.E., et al., *Noncytotoxic lytic granule-mediated CD8+ T cell inhibition of HSV-1 reactivation from neuronal latency.* Science, 2008. **322**(5899): p. 268-71.

359. Gebhardt, T. and L.K. Mackay, *Skin-resident memory T cells keep herpes simplex virus at bay.* Immunol Cell Biol, 2013. **91**(7): p. 441-2.
360. Zhu, J., et al., *Immune surveillance by CD8alphaalpha+ skin-resident T cells in human herpes virus infection.* Nature, 2013. **497**(7450): p. 494-7.
361. Zens, K.D., J.K. Chen, and D.L. Farber, *Vaccine-generated lung tissue-resident memory T cells provide heterosubtypic protection to influenza infection.* JCI Insight, 2016. **1**(10).
362. Lee, Y.N., et al., *A Novel Vaccination Strategy Mediating the Induction of Lung-Resident Memory CD8 T Cells Confers Heterosubtypic Immunity against Future Pandemic Influenza Virus.* J Immunol, 2016. **196**(6): p. 2637-45.
363. Thom, J.T. and A. Oxenius, *Tissue-resident memory T cells in cytomegalovirus infection.* Curr Opin Virol, 2016. **16**: p. 63-69.
364. Suni, M.A., et al., *CD4+CD8dim T lymphocytes exhibit enhanced cytokine expression, proliferation and cytotoxic activity in response to HCMV and HIV-1 antigens.* European Journal of Immunology, 2001. **31**(8): p. 2512-2520.
365. Macagno, A., et al., *Isolation of human monoclonal antibodies that potently neutralize human cytomegalovirus infection by targeting different epitopes on the gH/gL/UL128-131A complex.* J Virol, 2010. **84**(2): p. 1005-13.
366. Gerna, G., et al., *Human cytomegalovirus serum neutralizing antibodies block virus infection of endothelial/epithelial cells, but not fibroblasts, early during primary infection.* J Gen Virol, 2008. **89**(Pt 4): p. 853-65.
367. Forthal, D.N., et al., *Antibody-dependent cellular cytotoxicity independently predicts survival in severely immunocompromised human immunodeficiency virus-infected patients.* J Infect Dis, 1999. **180**(4): p. 1338-41.
368. Spencer, J.V., et al., *Potent immunosuppressive activities of cytomegalovirus-encoded interleukin-10.* J Virol, 2002. **76**(3): p. 1285-92.
369. Ahn, K., et al., *The ER-Luminal Domain of the HCMV Glycoprotein US6 Inhibits Peptide Translocation by TAP.* Immunity, 1997. **6**(5): p. 613-621.
370. Jones, T.R., et al., *Multiple independent loci within the human cytomegalovirus unique short region down-regulate expression of major histocompatibility complex class I heavy chains.* J Virol, 1995. **69**(8): p. 4830-41.
371. Wiertz, E.J., et al., *Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction.* Nature, 1996. **384**(6608): p. 432-8.
372. Gewurz, B.E., et al., *Antigen presentation subverted: Structure of the human cytomegalovirus protein US2 bound to the class I molecule HLA-A2.* Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6794-9.
373. Pande, N.T., et al., *Rhesus cytomegalovirus contains functional homologues of US2, US3, US6, and US11.* J Virol, 2005. **79**(9): p. 5786-98.
374. Wiertz, E.J., et al., *The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol.* Cell, 1996. **84**(5): p. 769-79.
375. Lehner, P.J., et al., *The human cytomegalovirus US6 glycoprotein inhibits transporter associated with antigen processing-dependent peptide translocation.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(13): p. 6904-6909.

376. Ahn, K., et al., *Human cytomegalovirus inhibits antigen presentation by a sequential multistep process.* Proc Natl Acad Sci U S A, 1996. **93**(20): p. 10990-5.
377. Jones, T.R., et al., *Human cytomegalovirus US3 impairs transport and maturation of major histocompatibility complex class I heavy chains.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(21): p. 11327-11333.
378. Johnson, D.C. and N.R. Hegde, *Inhibition of the MHC class II antigen presentation pathway by human cytomegalovirus.* Curr Top Microbiol Immunol, 2002. **269**: p. 101-15.
379. Burwitz, B.J., et al., *Cross-Species Rhesus Cytomegalovirus Infection of Cynomolgus Macaques.* PLoS Pathog, 2016. **12**(11): p. e1006014.
380. Powers, C.J. and K. Früh, *Signal Peptide-Dependent Inhibition of MHC Class I Heavy Chain Translation by Rhesus Cytomegalovirus.* PLOS Pathogens, 2008. **4**(10): p. e1000150.
381. Krmpotic, A., et al., *MCMV glycoprotein gp40 confers virus resistance to CD8+ T cells and NK cells in vivo.* Nat Immunol, 2002. **3**(6): p. 529-35.
382. Sturgill, E.R., et al., *Natural Killer Cell Evasion Is Essential for Infection by Rhesus Cytomegalovirus.* PLoS Pathog, 2016. **12**(8): p. e1005868.
383. Skaletskaya, A., et al., *A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation.* Proc Natl Acad Sci U S A, 2001. **98**(14): p. 7829-34.
384. Goldmacher, V.S., et al., *A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2.* Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12536-41.
385. McMichael, A.J. and L.J. Picker, *Unusual antigen presentation offers new insight into HIV vaccine design.* Curr Opin Immunol, 2017. **46**: p. 75-81.
386. Murray, S.E., et al., *Fibroblast-adapted human CMV vaccines elicit predominantly conventional CD8 T cell responses in humans.* J Exp Med, 2017. **214**(7): p. 1889-1899.
387. Yi, H.A., et al., *Cellular signals involved in cyclooxygenase-2 expression induced by human cytomegalovirus.* Virus Res, 2009. **146**(1-2): p. 89-96.
388. Esser, M.T., et al., *Memory T cells and vaccines.* Vaccine, 2003. **21**(5-6): p. 419-30.
389. van Duikeren, S., et al., *Vaccine-induced effector-memory CD8+ T cell responses predict therapeutic efficacy against tumors.* J Immunol, 2012. **189**(7): p. 3397-403.
390. Beverley, P.C., et al., *A novel murine cytomegalovirus vaccine vector protects against Mycobacterium tuberculosis.* J Immunol, 2014. **193**(5): p. 2306-16.
391. Marzi, A., et al., *Cytomegalovirus-based vaccine expressing Ebola virus glycoprotein protects nonhuman primates from Ebola virus infection.* Sci Rep, 2016. **6**: p. 21674.
392. Tsuda, Y., et al., *A replicating cytomegalovirus-based vaccine encoding a single Ebola virus nucleoprotein CTL epitope confers protection against Ebola virus.* PLoS Negl Trop Dis, 2011. **5**(8): p. e1275.
393. Tsuda, Y., et al., *A cytomegalovirus-based vaccine provides long-lasting protection against lethal Ebola virus challenge after a single dose.* Vaccine, 2015. **33**(19): p. 2261-6.
394. Dekhtiarenko, I., et al., *The context of gene expression defines the immunodominance hierarchy of cytomegalovirus antigens.* J Immunol, 2013. **190**(7): p. 3399-409.

395. Trsan, T., et al., *Superior induction and maintenance of protective CD8 T cells in mice infected with mouse cytomegalovirus vector expressing RAE-1gamma.* Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16550-5.
396. Bray, M., et al., *A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever.* J Infect Dis, 1998. **178**(3): p. 651-61.
397. Bray, M., *The role of the Type I interferon response in the resistance of mice to filovirus infection.* J Gen Virol, 2001. **82**(Pt 6): p. 1365-73.
398. Jones, S.M., et al., *Assessment of a vesicular stomatitis virus-based vaccine by use of the mouse model of Ebola virus hemorrhagic fever.* J Infect Dis, 2007. **196 Suppl 2**: p. S404-12.
399. Bilello, J.P., et al., *Vaccine protection against simian immunodeficiency virus in monkeys using recombinant gamma-2 herpesvirus.* J Virol, 2011. **85**(23): p. 12708-20.
400. Asher, D.M., et al., *Persistent shedding of cytomegalovirus in the urine of healthy Rhesus monkeys.* Proc Soc Exp Biol Med, 1974. **145**(3): p. 794-801.
401. Chang, W.L. and P.A. Barry, *Cloning of the full-length rhesus cytomegalovirus genome as an infectious and self-excisable bacterial artificial chromosome for analysis of viral pathogenesis.* J Virol, 2003. **77**(9): p. 5073-83.
402. Gill, R.B., et al., *Coding potential of UL/b' from the initial source of rhesus cytomegalovirus Strain 68-1.* Virology, 2013. **447**(1-2): p. 208-12.
403. Carlson, J.R., et al., *Rhesus brain microvascular endothelial cells are permissive for rhesus cytomegalovirus infection.* J Gen Virol, 2005. **86**(Pt 3): p. 545-9.
404. Wu, H.L., et al., *The Role of MHC-E in T Cell Immunity Is Conserved among Humans, Rhesus Macaques, and Cynomolgus Macaques.* J Immunol, 2017.
405. Fooden, J., *Comparative review of Fascicularis-group species of macaques (Primates: Macaca).* Feildiana Zool, 2006. **107**: p. 1-43.
406. Delson, E., *Evolutionary history of the Cercopithecidae.* Contrib Primatol, 1975. **5**: p. 167-217.
407. Ambagala, A.P., et al., *Isolation and characterization of cynomolgus macaque (Macaca fascicularis) cytomegalovirus (CyCMV).* Virology, 2011. **412**(1): p. 125-35.
408. Delson, E., *Fossil Macaques, Phyletic Relationships and a Scenario of Deployment*, in *The Macaques: Studies in ecology, behaviour, and evolution*, D.G. Lindburg, Editor. 1980, Van Nostrand-Reinhold: New York. p. 10-30.
409. Tosi, A.J., J.C. Morales, and D.J. Melnick, *Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys.* Evolution, 2003. **57**(6): p. 1419-35.
410. Li, Q.Q. and Y.P. Zhang, *Phylogenetic relationships of the macaques (Cercopithecidae: Macaca), inferred from mitochondrial DNA sequences.* Biochem Genet, 2005. **43**(7-8): p. 375-86.
411. Arnason, U., et al., *The "Phoca standard": an external molecular reference for calibrating recent evolutionary divergences.* J Mol Evol, 1996. **43**(1): p. 41-5.
412. Campbell, K.J., et al., *Characterization of 47 MHC class I sequences in Filipino cynomolgus macaques.* Immunogenetics, 2009. **61**(3): p. 177-87.
413. Blancher, A., et al., *Study of cynomolgus monkey (Macaca fascicularis) Mhc DRB gene polymorphism in four populations.* Immunogenetics, 2012. **64**(8): p. 605-14.

414. Leuchte, N., et al., *MhcDRB-sequences from cynomolgus macaques (Macaca fascicularis) of different origin.* Tissue Antigens, 2004. **63**(6): p. 529-37.
415. Melnick, D.J. and K.K. Kidd, *Genetic and Evolutionary Relationships among Asian Macaques.* International Journal of Primatology, 1985. **6**(2): p. 123-160.
416. Trask, J.A.S., et al., *Single-Nucleotide Polymorphisms Reveal Patterns of Allele Sharing Across the Species Boundary Between Rhesus (Macaca mulatta) and Cynomolgus (M. fascicularis) Macaques.* American Journal of Primatology, 2013. **75**(2): p. 135-144.
417. Fooden, J., *Comparative review of Fascicularis-group species of macaques (Primates: Macaca) / Jack Fooden*. Vol. n.s. no.107(2006). 2006, Chicago, Ill. :: Field Museum of Natural History.
418. Kanthaswamy, S., et al., *Interspecies Hybridization and the Stratification of Nuclear Genetic Variation of Rhesus (Macaca Mulatta) and Long-Tailed Macaques (Macaca Fascicularis).* Int J Primatol, 2008. **29**(5): p. 1295-1311.
419. Tosi, A.J. and C.S. Coke, *Comparative phylogenetics offer new insight into the biogeographic history of Macaca fascicuaris and the origin of the Mauritian macaque.* Moleculat Phylogenetics and Evolution, 2007. **42**: p. 498-504.
420. Tosi, A.J., J.C. Morales, and D.J. Melnick, *Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history.* Mol Phylogenet Evol, 2000. **17**(2): p. 133-44.
421. Tosi, A.J., J.C. Morales, and D.J. Melnick, *Y-chromosome and mitochondrial markers in Macaca fascicularis indicate introgression with Indochinese M-mulatta and a biogeographic barrier in the Isthmus of Kra.* International Journal of Primatology, 2002. **23**(1): p. 161-178.
422. Kawamoto, Y., et al., *Genetic diversity of longtail macaques (Macaca fascicularis) on the island of Mauritius: an assessment of nuclear and mitochondrial DNA polymorphisms.* J Med Primatol, 2008. **37**(1): p. 45-54.
423. Stevison, L.S. and M.H. Kohn, *Divergence population genetic analysis of hybridization between rhesus and cynomolgus macaques.* Mol Ecol, 2009. **18**(11): p. 2457-75.
424. Lankau, E.W., et al., *Use of Nonhuman Primates in Research in North America.* Journal of the American Association for Laboratory Animal Science : JAALAS, 2014. **53**(3): p. 278-282.
425. Rogers, J. and R.A. Gibbs, *Applications of Next-Generation Sequencing Comparative Primate Genomics: Emerging Patterns of Genome Content and Dynamics.* Nature Reviews Genetics, 2014. **15**(5): p. 347-359.
426. Yan, G., et al., *Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques.* Nat Biotechnol, 2011. **29**(11): p. 1019-23.
427. Kumar, S. and S.B. Hedges, *A molecular timescale for vertebrate evolution.* Nature, 1998. **392**(6679): p. 917-20.
428. *Demands for rhesus monkeys in biomedical research: a workshop report.* ILAR J, 2003. **44**(3): p. 222-35.
429. Zhou, Y., et al., *SIV infection of rhesus macaques of Chinese origin: a suitable model for HIV infection in humans.* Retrovirology, 2013. **10**: p. 89.

430. Kanthaswamy, S., et al., *The genetic composition of populations of cynomolgus macaques (Macaca fascicularis) used in biomedical research.* J Med Primatol, 2013. **42**(3): p. 120-31.
431. Sussman, R.W. and I. Tattersall, *Behavior and ecology ofmacaca fascicularis in Mauritius: A preliminary study.* Primates, 1981. **22**(2): p. 192-205.
432. Matsubayashi, K., et al., *Clinical examinations on crab-eating macaques in mauritius.* Primates, 1992. **33**(2): p. 281-288.
433. Kondo, M., et al., *Population genetics of crab-eating macaques (Macaca fascicularis) on the island of mauritius.* American Journal of Primatology, 1993. **29**(3): p. 167-182.
434. Lawler, S.H., R.W. Sussman, and L.L. Taylor, *Mitochondrial DNA of the Mauritian macaques (Macaca fascicularis): an example of the founder effect.* Am J Phys Anthropol, 1995. **96**(2): p. 133-41.
435. Satkoski Trask, J., et al., *Population and landscape genetics of an introduced species (M. fascicularis) on the island of Mauritius.* PLoS One, 2013. **8**(1): p. e53001.
436. Tosi, A.J. and C.S. Coke, *Comparative phylogenetics offer new insights into the biogeographic history of Macaca fascicularis and the origin of the Mauritian macaques.* Mol Phylogenet Evol, 2007. **42**(2): p. 498-504.
437. Hatziioannou, T. and D.T. Evans, *Animal models for HIV/AIDS research.* Nat Rev Microbiol, 2012. **10**(12): p. 852-67.
438. Geraghty, D.E., et al., *Polymorphism at the HLA-E locus predates most HLA-A and -B polymorphism.* Human Immunology, 1992. **33**(3): p. 174-184.
439. Grimsley, C., et al., *Definitive high resolution typing of HLA-E allelic polymorphisms: Identifying potential errors in existing allele data.* Tissue Antigens, 2002. **60**(3): p. 206-12.
440. Grimsley, C. and C. Ober, *Population genetic studies of HLA-E: evidence for selection.* Hum Immunol, 1997. **52**(1): p. 33-40.
441. CITES, *CITES Trade Database*. 2017, United Nations Environment Programme World Conservation Monitoring Centre: Cambridge, UK.
442. Yap, M.W., S. Nisole, and J.P. Stoye, *A Single Amino Acid Change in the SPRY Domain of Human Trim5α Leads to HIV-1 Restriction.* Current Biology. **15**(1): p. 73-78.
443. Brennan, G., Y. Kozyrev, and S.-L. Hu, *TRIMCyp expression in Old World primates Macaca nemestrina and Macaca fascicularis.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(9): p. 3569-3574.
444. Morrow, W.J., et al., *Small animals are not susceptible to human immunodeficiency virus infection.* J Gen Virol, 1987. **68 (Pt 8)**: p. 2253-7.
445. Levy, J.A., et al., *AIDS retrovirus (ARV-2) clone replicates in transfected human and animal fibroblasts.* Science, 1986. **232**(4753): p. 998-1001.
446. Alter, H.J., et al., *Transmission of HTLV-III infection from human plasma to chimpanzees: an animal model for AIDS.* Science, 1984. **226**(4674): p. 549-52.
447. O'Neil, S.P., et al., *Progressive infection in a subset of HIV-1-positive chimpanzees.* J Infect Dis, 2000. **182**(4): p. 1051-62.
448. Fultz, P.N., et al., *Persistent infection of chimpanzees with human T-lymphotropic virus type III/lymphadenopathy-associated virus: a potential model for acquired immunodeficiency syndrome.* J Virol, 1986. **58**(1): p. 116-24.

449. Agy, M.B., et al., *Infection of Macaca nemestrina by human immunodeficiency virus type-1.* Science, 1992. **257**(5066): p. 103-6.
450. Greenwood, E.J.D., et al., *Simian Immunodeficiency Virus Infection of Chimpanzees (Pan troglodytes) Shares Features of Both Pathogenic and Non-pathogenic Lentiviral Infections.* PLOS Pathogens, 2015. **11**(9): p. e1005146.
451. Hatziioannou, T., et al., *A macaque model of HIV-1 infection.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(11): p. 4425-4429.
452. Liao, C.H., et al., *A novel fusion gene, TRIM5-Cyclophilin A in the pig-tailed macaque determines its susceptibility to HIV-1 infection.* Aids, 2007. **21 Suppl 8**: p. S19-26.
453. Hirsch, V.M., et al., *An African primate lentivirus (SIVsm) closely related to HIV-2.* Nature, 1989. **339**(6223): p. 389-92.
454. Murphey-Corb, M., et al., *Isolation of an HTLV-III-related retrovirus from macaques with simian AIDS and its possible origin in asymptomatic mangabeys.* Nature, 1986. **321**: p. 435.
455. Peeters, M., et al., *Isolation and characterization of a new chimpanzee lentivirus (simian immunodeficiency virus isolate cpz-ant) from a wild-captured chimpanzee.* Aids, 1992. **6**(5): p. 447-51.
456. Silvestri, G., et al., *Nonpathogenic SIV infection of sooty mangabeys is characterized by limited bystander immunopathology despite chronic high-level viremia.* Immunity, 2003. **18**(3): p. 441-52.
457. Daniel, M.D., et al., *Isolation of T-cell tropic HTLV-III-like retrovirus from macaques.* Science, 1985. **228**(4704): p. 1201-4.
458. Apetrei, C., et al., *Molecular Epidemiology of Simian Immunodeficiency Virus SIVsm in U.S. Primate Centers Unravels the Origin of SIVmac and SIVstm.* Journal of Virology, 2005. **79**(14): p. 8991-9005.
459. Voevodin, A. and P. Marx, *Lentivirus AIDS Models*, in *Simian Virology*. 2009, Wiley-Blackwell: Singapore. p. 119-154.
460. Morton, W.R., et al., *Transmission of the simian immunodeficiency virus SIVmne in macaques and baboons.* J Med Primatol, 1989. **18**(3-4): p. 237-45.
461. Hirsch, V., et al., *A molecularly cloned, pathogenic, neutralization-resistant simian immunodeficiency virus, SIVsmE543-3.* J Virol, 1997. **71**(2): p. 1608-20.
462. Kestler, H., et al., *Induction of AIDS in rhesus monkeys by molecularly cloned simian immunodeficiency virus.* Science, 1990. **248**(4959): p. 1109-12.
463. Klatt, N.R., et al., *Dynamics of simian immunodeficiency virus SIVmac239 infection in pigtail macaques.* J Virol, 2012. **86**(2): p. 1203-13.
464. Li, J., et al., *Infection of cynomolgus monkeys with a chimeric HIV-1/SIVmac virus that expresses the HIV-1 envelope glycoproteins.* Journal of acquired immune deficiency syndromes, 1992. **5**(7): p. 639.
465. Li, J.T., et al., *Persistent infection of macaques with simian-human immunodeficiency viruses.* Journal of Virology, 1995. **69**(11): p. 7061-7067.
466. Reimann, K.A., et al., *An env gene derived from a primary human immunodeficiency virus type 1 isolate confers high in vivo replicative capacity to a chimeric simian/human immunodeficiency virus in rhesus monkeys.* Journal of Virology, 1996. **70**(5): p. 3198-3206.

467. Igarashi, T., et al., *Emergence of a highly pathogenic simian/human immunodeficiency virus in a rhesus macaque treated with anti-CD8 mAb during a primary infection with a nonpathogenic virus.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(24): p. 14049-14054.
468. Joag, S.V., et al., *Chimeric simian/human immunodeficiency virus that causes progressive loss of CD4+ T cells and AIDS in pig-tailed macaques.* J Virol, 1996. **70**(5): p. 3189-97.
469. Nishimura, Y., et al., *Highly pathogenic SHIVs and SIVs target different CD4(+) T cell subsets in rhesus monkeys, explaining their divergent clinical courses.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(33): p. 12324-12329.
470. Harouse, J.M., et al., *Distinct pathogenic sequela in rhesus macaques infected with CCR5 or CXCR4 utilizing SHIVs.* Science, 1999. **284**(5415): p. 816-9.
471. Hsu, M., et al., *Increased mucosal transmission but not enhanced pathogenicity of the CCR5-tropic, simian AIDS-inducing simian/human immunodeficiency virus SHIV(SF162P3) maps to envelope gp120.* J Virol, 2003. **77**(2): p. 989-98.
472. Chang, H.-W., et al., *Generation and Evaluation of Clade C Simian-Human Immunodeficiency Virus Challenge Stocks.* Journal of Virology, 2015. **89**(4): p. 1965-1974.
473. Himathongkham, S., et al., *Simian-Human Immunodeficiency Virus Containing a Human Immunodeficiency Virus Type 1 Subtype-E Envelope Gene: Persistent Infection, CD4+ T-Cell Depletion, and Mucosal Membrane Transmission in Macaques.* Journal of Virology, 2000. **74**(17): p. 7851-7860.
474. Willer, D.O., et al., *Generation of recombinant parental Oka strain VZV-vectored SIV vaccine candidates.* 2012, University of Toronto.
475. Willer, D.O., et al., *Experimental infection of Cynomolgus Macaques (Macaca fascicularis) with human varicella-zoster virus.* J Virol, 2012. **86**(7): p. 3626-34.
476. Willer, D.O., et al., *Multi-low-dose mucosal simian immunodeficiency virus SIVmac239 challenge of cynomolgus macaques immunized with "hyperattenuated" SIV constructs.* J Virol, 2010. **84**(5): p. 2304-17.
477. Barry, P.A., et al., *Nonhuman Primate Models of Intrauterine Cytomegalovirus Infection.* ILAR Journal, 2006. **47**(1): p. 49-64.
478. Itell, H.L., et al., *Rhesus monkeys for a nonhuman primate model of cytomegalovirus infections.* Current Opinion in Virology, 2017. **25**(Complete): p. 126-133.
479. Barry, P.A. and L. Strelow, *Development of breeding populations of rhesus macaques (Macaca mulatta) that are specific pathogen-free for rhesus cytomegalovirus.* Comp Med, 2008. **58**(1): p. 43-6.
480. Rivailler, P., et al., *Genomic sequence of rhesus cytomegalovirus 180.92: insights into the coding potential of rhesus cytomegalovirus.* J Virol, 2006. **80**(8): p. 4179-82.
481. Yu, D., et al., *Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene.* J Virol, 2002. **76**(5): p. 2316-28.
482. Murphy, E., et al., *Coding potential of laboratory and clinical strains of human cytomegalovirus.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14976-81.
483. Dargan, D.J., et al., *Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture.* J Gen Virol, 2010. **91**(Pt 6): p. 1535-46.

484. Lilja, A.E. and T. Shenk, *Efficient replication of rhesus cytomegalovirus variants in multiple rhesus and human cell types.* Proc Natl Acad Sci U S A, 2008. **105**(50): p. 19950-5.
485. Bierle, C.J., et al., *Targeted Mutagenesis of Guinea Pig Cytomegalovirus Using CRISPR/Cas9-Mediated Gene Editing.* J Virol, 2016. **90**(15): p. 6989-98.
486. Smith, G.A. and L.W. Enquist, *Construction and transposon mutagenesis in Escherichia coli of a full-length infectious clone of pseudorabies virus, an alphaherpesvirus.* J Virol, 1999. **73**(8): p. 6405-14.
487. Shizuya, H., et al., *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, 1992. **89**(18): p. 8794-8797.
488. Cui, X., et al., *Cloning the complete guinea pig cytomegalovirus genome as an infectious bacterial artificial chromosome with excisable origin of replication.* J Virol Methods, 2008. **149**(2): p. 231-9.
489. Smith, G.A. and L.W. Enquist, *A self-recombining bacterial artificial chromosome and its application for analysis of herpesvirus pathogenesis.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4873-8.
490. Gahan, M.E., et al., *Bacterial Antigen Expression Is an Important Component in Inducing an Immune Response to Orally Administered Salmonella-Delivered DNA Vaccines.* PLOS ONE, 2009. **4**(6): p. e6062.
491. Goussard, S., C. Grillot-Courvalin, and P. Courvalin, *Eukaryotic promoters can direct protein synthesis in Gram-negative bacteria.* J Mol Microbiol Biotechnol, 2003. **6**(3-4): p. 211-8.
492. Paredes, A.M. and D. Yu, *Human cytomegalovirus: bacterial artificial chromosome (BAC) cloning and genetic manipulation.* Curr Protoc Microbiol, 2012. **Chapter 14**: p. Unit14E 4.
493. ten Haaft, P., et al., *Comparison of early plasma RNA loads in different macaque species and the impact of different routes of exposure on SIV/SHIV infection.* J Med Primatol, 2001. **30**(4): p. 207-14.
494. Antony, J.M. and K.S. MacDonald, *A critical analysis of the cynomolgus macaque, Macaca fascicularis, as a model to test HIV-1/SIV vaccine efficacy.* Vaccine, 2015. **33**(27): p. 3073-83.
495. Brunk, C.G., et al., *Follow-up Report on the Health Canada Non-Human Primate Colony: An Expert Panel Follow-up Report prepared at the request of the Royal Society of Canada for Health Canada*. 2003, The Royal Society of Canada/La Société royale du Canada: Ottawa, Ontario. p. 15.
496. Creager, H.M., et al., *Characterization of full-length MHC class II sequences in Indonesian and Vietnamese cynomolgus macaques.* Immunogenetics, 2011. **63**(9): p. 611-8.
497. Alcendor, D.J., et al., *Patterns of divergence in the vCXCL and vGPCR gene clusters in primate cytomegalovirus genomes.* Virology, 2009. **395**(1): p. 21-32.
498. Oxford, K.L., et al., *Protein coding content of the UL)b' region of wild-type rhesus cytomegalovirus.* Virology, 2008. **373**(1): p. 181-8.

499. Davison, A.J., *Comparative analysis of the genomes*, in *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis*, C.-F.G. Arvin A, Mocarski et al.,, Editor. 2007, Cambridge University Press: Cambridge.
500. Kolb, A.W., C. Ané, and C.R. Brandt, *Using HSV-1 Genome Phylogenetics to Track Past Human Migrations.* PLoS ONE, 2013. **8**(10): p. e76267.
501. Smith, A.L., D.H. Black, and R. Eberle, *Molecular evidence for distinct genotypes of monkey B virus (herpesvirus simiae) which are related to the macaque host species.* J Virol, 1998. **72**(11): p. 9224-32.
502. Chou, S., *Reactivation and Recombination of Multiple Cytomegalovirus Strains from Individual Organ Donors.* The Journal of Infectious Diseases, 1989. **160**(1): p. 11-15.
503. Lees, D.N., et al., *Herpesvirus simiae (B virus) antibody response and virus shedding in experimental primary infection of cynomolgus monkeys.* Lab Anim Sci, 1991. **41**(4): p. 360-4.
504. Jacobs, J.P., C.M. Jones, and J.P. Baille, *Characteristics of a human diploid cell designated MRC-5.* Nature, 1970. **227**(5254): p. 168-70.
505. Zerbino, D.R. and E. Birney, *Velvet: algorithms for de novo short read assembly using de Bruijn graphs.* Genome Res, 2008. **18**(5): p. 821-9.
506. Borst, E.M. and M. Messerle, *Analysis of human cytomegalovirus oriLyt sequence requirements in the context of the viral genome.* J Virol, 2005. **79**(6): p. 3615-26.
507. Darling, A.C., et al., *Mauve: multiple alignment of conserved genomic sequence with rearrangements.* Genome Res, 2004. **14**(7): p. 1394-403.
508. Katoh, K., et al., *MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.* Nucleic Acids Res, 2002. **30**(14): p. 3059-66.
509. Darriba, D., et al., *jModelTest 2: more models, new heuristics and parallel computing.* Nat Methods, 2012. **9**(8): p. 772.
510. Huelsenbeck, J.P. and F. Ronquist, *MRBAYES: Bayesian inference of phylogenetic trees.* Bioinformatics, 2001. **17**(8): p. 754-5.
511. Chris Loken, D.G., Leslie Groer, Richard Peltier, Neil Bunn, Michael Craig, Teresa Henriques, Jillian Dempsey, Ching-Hsing Yu, Joseph Chen, L Jonathan Dursi, Jason Chong, Scott Northrup, Jaime Pinto, Neil Knecht and Ramses Van Zon, *SciNet: Lessons Learned from Building a Power-efficient Top-20 System and Data Centre* J. Phys.: Conf. Ser., 2010. **256**(1): p. 012026
512. Lilja, A.E., et al., *Functional genetic analysis of rhesus cytomegalovirus: Rh01 is an epithelial cell tropism factor.* J Virol, 2008. **82**(5): p. 2170-81.
513. Omoto, S. and E.S. Mocarski, *Cytomegalovirus UL91 is essential for transcription of viral true late (gamma2) genes.* J Virol, 2013. **87**(15): p. 8651-64.
514. Wong, W.S.W. and R. Nielsen, *Detecting selection in noncoding regions of nucleotide sequences.* Genetics, 2004. **167**(2): p. 949-958.
515. Ohta, T., *The Nearly Neutral Theory of Molecular Evolution.* Annual Review of Ecology and Systematics 1992. **23**: p. 263-286
516. Assaf, B.T., et al., *Limited dissemination and shedding of the UL128 complex-intact, UL/b'-defective rhesus cytomegalovirus strain 180.92.* J Virol, 2014. **88**(16): p. 9310-20.
517. Ryckman, B.J., et al., *Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells.* J Virol, 2008. **82**(1): p. 60-70.

518. Nicholas, J., *Evolutionary aspects of oncogenic herpesviruses.* Mol Pathol, 2000. **53**(5): p. 222-37.
519. Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.
520. Arnolds, K.L., A.P. Lares, and J.V. Spencer, *The US27 gene product of human cytomegalovirus enhances signaling of host chemokine receptor CXCR4.* Virology, 2013. **439**(2): p. 122-31.
521. Golden, M.P., et al., *Cytomegalovirus vasculitis. Case reports and review of the literature.* Medicine (Baltimore), 1994. **73**(5): p. 246-55.
522. Arslan, F., et al., *Cytomegalovirus (CMV)-related cutaneous necrotizing vasculitis: case report and literature review.* Braz J Infect Dis, 2012. **16**(5): p. 482-5.
523. Cohen, P. and L. Guillevin, *Vascularites associées aux infections virales.* . La Presse Médicale, 2004. **33**: p. 1371-1384.
524. Sacre, K., et al., *A role for cytomegalovirus-specific CD4+CX3CR1+ T cells and cytomegalovirus-induced T-cell immunopathology in HIV-associated atherosclerosis.* AIDS, 2012. **26**(7): p. 805-14.
525. Grattan, M.T., et al., *Cytomegalovirus infection is associated with cardiac allograft rejection and atherosclerosis.* JAMA, 1989. **261**(24): p. 3561-3566.
526. Penfold, M.E., et al., *Characterization of the rhesus cytomegalovirus US28 locus.* J Virol, 2003. **77**(19): p. 10404-13.
527. Neubauer, S.R., *Characterization of the genome of baboon cytomegalovirus strain (OCOM4-37) isolated from the olive baboon, Papio cynocephalus anubis*. 2011, Oklahoma State University.
528. Rogers, J. and R.A. Gibbs, *Comparative primate genomics: emerging patterns of genome content and dynamics.* Nat Rev Genet, 2014. **15**(5): p. 347-59.
529. Gerna, G., et al., *The attenuated Towne strain of human cytomegalovirus may revert to both endothelial cell tropism and leuko- (neutrophil- and monocyte-) tropism in vitro.* J Gen Virol, 2002. **83**(Pt 8): p. 1993-2000.
530. Numazaki, K., M. Ikehata, and S. Chiba, *Subtyping of cytomegalovirus strains obtained from immunocompetent children.* In Vivo, 2000. **14**(6): p. 745-6.
531. Marchini, A., H. Liu, and H. Zhu, *Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes.* J Virol, 2001. **75**(4): p. 1870-8.
532. Mohr, C.A., et al., *Engineering of cytomegalovirus genomes for recombinant live herpesvirus vaccines.* Int J Med Microbiol, 2008. **298**(1-2): p. 115-25.
533. Kirchoff, V., et al., *Generation of a life-expanded rhesus monkey fibroblast cell line for the growth of rhesus rhadinovirus (RRV).* Arch Virol, 2002. **147**(2): p. 321-33.
534. Ambagala, A.P., et al., *Establishment of an immortal cynomolgus macaque fibroblast cell line for propagation of cynomolgus macaque cytomegalovirus (CyCMV).* Arch Virol, 2013. **158**(5): p. 955-65.
535. Chang, W.L., et al., *Replication of rhesus cytomegalovirus in life-expanded rhesus fibroblasts expressing human telomerase.* J Virol Methods, 2002. **104**(2): p. 135-46.
536. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform.* Bioinformatics, 2010. **26**(5): p. 589-95.
537. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7: improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.

538. Hartung, M. and B. Kisters-Woike, *Cre mutants with altered DNA binding properties.* J Biol Chem, 1998. **273**(36): p. 22884-91.
539. Missirlis, P.I., D.E. Smailus, and R.A. Holt, *A high-throughput screen identifying sequence and promiscuity characteristics of the loxP spacer region in Cre-mediated recombination.* BMC Genomics, 2006. **7**: p. 73.
540. Semprini, S., et al., *Cryptic loxP sites in mammalian genomes: genome-wide distribution and relevance for the efficiency of BAC/PAC recombineering techniques.* Nucleic Acids Res, 2007. **35**(5): p. 1402-10.
541. Hoess, R.H., A. Wierzbicki, and K. Abremski, *The role of the loxP spacer region in P1 site-specific recombination.* Nucleic Acids Res, 1986. **14**(5): p. 2287-300.
542. Abremski, K., et al., *Bacteriophage P1 Cre-loxP site-specific recombination. Site-specific DNA topoisomerase activity of the Cre recombination protein.* J Biol Chem, 1986. **261**(1): p. 391-6.
543. Santoro, S.W. and P.G. Schultz, *Directed evolution of the site specificity of Cre recombinase.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4185-90.
544. Sinzger, C., et al., *Modification of human cytomegalovirus tropism through propagation in vitro is associated with changes in the viral genome.* J Gen Virol, 1999. **80 (Pt 11)**: p. 2867-77.
545. McCormick, A.L., et al., *Differential function and expression of the viral inhibitor of caspase 8-induced apoptosis (vICA) and the viral mitochondria-localized inhibitor of apoptosis (vMIA) cell death suppressors conserved in primate and rodent cytomegaloviruses.* Virology, 2003. **316**(2): p. 221-33.
546. Sekulin, K., et al., *Analysis of the variability of CMV strains in the RL11D domain of the RL11 multigene family.* Virus Genes, 2007. **35**(3): p. 577-83.
547. Orme, I.M. and F.M. Collins, *Adoptive protection of the Mycobacterium tuberculosis-infected lung. Dissociation between cells that passively transfer protective immunity and those that transfer delayed-type hypersensitivity to tuberculin.* Cell Immunol, 1984. **84**(1): p. 113-20.
548. Lechner, F., et al., *Analysis of successful immune responses in persons infected with hepatitis C virus.* J Exp Med, 2000. **191**(9): p. 1499-512.
549. Traina-Dorge, V., et al., *Recombinant varicella vaccines induce neutralizing antibodies and cellular immune responses to SIV and reduce viral loads in immunized rhesus macaques.* Vaccine, 2010. **28**(39): p. 6483-90.
550. Asano, Y., et al., *Immunogenicity of wild and attenuated varicella-zoster virus strains in rhesus monkeys.* J Med Virol, 1984. **14**(4): p. 305-12.
551. Staprans, S.I., et al., *Enhanced SIV replication and accelerated progression to AIDS in macaques primed to mount a CD4 T cell response to the SIV envelope protein.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(35): p. 13026-13031.
552. Dekhtiarenko, I., et al., *Peptide Processing Is Critical for T-Cell Memory Inflation and May Be Optimized to Improve Immune Protection by CMV-Based Vaccine Vectors.* PLoS Pathog, 2016. **12**(12): p. e1006072.
553. Reynolds, M.R., et al., *A trivalent recombinant Ad5 gag/pol/nef vaccine fails to protect rhesus macaques from infection or control virus replication after a limiting-dose heterologous SIV challenge.* Vaccine, 2012. **30**(30): p. 4465-75.

554. Guo, J.-C., et al., *Highly Efficient CRISPR/Cas9-Mediated Homologous Recombination Promotes the Rapid Generation of Bacterial Artificial Chromosomes of Pseudorabies Virus.* Frontiers in Microbiology, 2016. **7**: p. 2110.
555. Thyagarajan, B., et al., *Mammalian genomes contain active recombinase recognition sites.* Gene, 2000. **244**(1-2): p. 47-54.
556. Gopaul, D.N., F. Guo, and G.D. Van Duyne, *Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination.* EMBO J, 1998. **17**(14): p. 4175-87.
557. Hel, Z., et al., *Improved vaccine protection from simian AIDS by the addition of nonstructural simian immunodeficiency virus genes.* J Immunol, 2006. **176**(1): p. 85-96.
558. Schachtel, G.A., et al., *Evidence for selective evolution in codon usage in conserved amino acid segments of human alphaherpesvirus proteins.* J Mol Evol, 1991. **33**(6): p. 483-94.
559. Vogt, C., et al., *ORF57 Overcomes the Detrimental Sequence Bias of Kaposi's Sarcoma-Associated Herpesvirus Lytic Genes.* J Virol, 2015. **89**(9): p. 5097-109.
560. Desrosiers, R. *Impressive Vaccine Protection Against SIVmac239 by Recombinant Gamma-2 Herpesvirus*. in *AIDS Vaccine*. 2013. Barcelona, Spain.
561. Wilkinson, G.W.G., et al., *Human cytomegalovirus: taking the strain.* Medical Microbiology and Immunology, 2015. **204**(3): p. 273-284.
562. Neff, B.J., et al., *Clinical and laboratory studies of live cytomegalovirus vaccine Ad-169.* Proc Soc Exp Biol Med, 1979. **160**(1): p. 32-7.
563. Plotkin, S.A., J. Farquhar, and E. Horberger, *Clinical trials of immunization with the Towne 125 strain of human cytomegalovirus.* J Infect Dis, 1976. **134**(5): p. 470-5.
564. Chen, S., et al., *Cytomegalovirus seropositivity is associated with glucose regulation in the oldest old. Results from the Leiden 85-plus Study.* Immunity & Ageing, 2012. **9**(1): p. 18.
565. Mahnke, Y.D., et al., *The who's who of T-cell differentiation: human memory T-cell subsets.* Eur J Immunol, 2013. **43**(11): p. 2797-809.
566. Faint, J.M., et al., *Memory T cells constitute a subset of the human CD8+CD45RA+ pool with distinct phenotypic and migratory characteristics.* J Immunol, 2001. **167**(1): p. 212-20.
567. Montefiori, D.C., *Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays.* Curr Protoc Immunol, 2005. **Chapter 12**: p. Unit 12.11.
568. Rosner, B., *Fundamentals of Biostatistics*. 7 ed. 2011, Boston, MA: Brooks/Cole.
569. Del Prete, G.Q., et al., *Molecularly tagged simian immunodeficiency virus SIVmac239 synthetic swarm for tracking independent infection events.* J Virol, 2014. **88**(14): p. 8077-90.
570. Quinnan, G.V., Jr., et al., *Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus.* Ann Intern Med, 1984. **101**(4): p. 478-83.