Cytomegalovirus diversity among patients of all ages in Johannesburg, South Africa, 2022-2023

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

By

Pholenyana Petronela Letlonkana

A Dissertation submitted to the Faculty of Health Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science. Johannesburg, 2025.

2025

Supervisors: Dr Florette Treurnicht
Dr Allison Glass
Dr Bonolo Mashishi

DECLARATION

I, Pholenyana Petronela Letlonkana declare that this Dissertation is my own, unaided work. It is being submitted for Master of Science in medicine specialising in Virology at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____ (Signature of candidate) _____23rd

_____day of ____October_____2025____in__Auckland Park_____

CONFERENCE PRESENTATIONS FROM THIS STUDY

Petronela Letlonkana, Allison Glass, Bonolo Mashishi, Florette Treurnicht. Cytomegalovirus diversity among patients of all ages in Johannesburg, South Africa, 2022-2023. PathRed 2025 conference, Johannesburg, South Africa. 02-05 October 2025. Sprint Oral Poster. Winner, Virology Poster category.

ABSTRACT

Background: Cytomegalovirus (CMV), a common deoxyribonucleic acid (DNA) viral DNA pathogen with diverse genotypes, can possibly and significantly affect the severity of the disease and treatment outcomes. Accurate and reliable glycoprotein B (gB) genotype detection is important in understanding the epidemiology of CMV infection and disease.

Methods: A total of 111 specimens were collected from patients who accessed care in public and private healthcare settings in Johannesburg, South Africa for CMV related illness. Study samples were tested for CMV using the Argene CMV-R gene real-time PCR detection kit. Thereafter CMV positive samples underwent genotyping using an in-house real-time multiplex PCR, targeting gB1, gB2, gB3, and gB4. Samples that failed genotyping were further analysed using conventional PCR followed by sequencing. In addition, some CMV positive samples were selected for amplification and sequencing of the unique long (UL) 54 and UL97 genes to identify mutations associated with drug resistance.

Results: Among the total study population 25/111 (23%) were excluded due to sample haemolysis or insufficient volume. From the remaining 86/111 (77%) tested for CMV, 86% (74/86) tested positive. The gB3 genotype was the most prevalent at 77% (57/74). No single gB2, gB1 and gB4 genotypes were detected. Mixed genotype infections were detected in 6.7% (5/74) of patient samples. Unfortunately, we were unable to generate UL54 and UL97 sequence data to analyse for drug resistance.

Discussion and Conclusions: Our analysis revealed genotype gB3 as the predominant strain within the study population. The presence of mixed infections in 6.7% of samples highlights the complex nature of CMV epidemiology. Overall, this research enhances the understanding of CMV genetic diversity, which is important for informing targeted public health and treatment strategies.

ACKNOWLEDGEMENTS

I am deeply grateful to God for His faithfulness, grace, and mercy, which have sustained me throughout my journey. His promises are indeed true, and I have experienced His prosperity and timely intervention.

I extend my heartfelt thanks to the Poliomyelitis Research Foundation (PRF) and the Postgraduate Research Office for funding my degree.

To my supervisors, Dr Florette Treurnicht, Dr Allison Glass and Dr Bonolo Mashishi, I appreciate your guidance, support, and encouragement. Your expertise and mentorship made a significant difference in my academic journey.

TABLE OF CONTENTS

List of tables

List of figures

	CM76, F: CM25). The Y-axis indicates the sequence read coverage or number of reads aligning to each region and the X-axis (in base pairs) indicates the positions across the CMV genome. The genome regions where there is significant coverage are highlighted in red.
Figure 3.5	Phylogenetic analysis of UL55 gene fragments among CMV isolates. The tree was constructed using the maximum likelihood method with 500 bootstrap replicates under the Hasegawa-Kishino-Yano model. Branch lengths represent evolutionary distances. The multiple sequence alignment was trimmed to 280 bp in order to include partial coding sequences for UL55 genotype (gB1–gB4) references.
Figure 3.6	Maximum-likelihood phylogenetic analysis of the UL55 gene fragment among CMV isolates, incorporating the reported gB5 strain (Human betaherpesvirus 5 isolate BCR-VU). The tree was constructed using 500 bootstrap replicates, with confidence values indicated at branch nodes. The multiple sequence alignment was trimmed to 279 bp in order to include partial coding sequences for UL55 genotype (gB1–gB4) references as well as the gB5 strain. Branch lengths represent evolutionary distances, and the Hasegawa-Kishino-Yano model was applied to account for nucleotide substitutions.
Figure 3.7	CMV genotype distribution in single and mixed infections among persons of all ages, 2022-2023. The X-axis indicates the CMV genotypes, and the Y-axis represent the percentage of each genotype detected.
Figure 3.8	Gel electrophoresis image of 10 samples that were successfully amplified as seen by the presence of bands. A molecular marker (M) is used to identify and verify the size of the specific gene (UL54)
Figure 3.9	Screenshot taken from GenBank showing the position of the UL54 gene on the CMV virus graphics.

Figure 3.10	Genome coverage maps. These visual maps provide a detailed overview of the location of genetic data for UL54 gene sequences. The Y-axis indicates the sequence read coverage or number of reads aligning to each region and the X-axis indicates the positions across the CMV genome. The genome regions where there is significant coverage are highlighted in red.
Figure 3.11	The Maximum Likelihood technique based on the Hasegawa-Kishino-Yano model was used to infer the evolutionary history of the Human Cytomegalovirus UL54 gene. Next to the branches is the proportion of trees in which the related taxa clustered together. The tree is depicted to scale, and branch lengths are measured in substitutions per site. The study included 13 nucleotide sequences. The 6 sequences in the study were represented by a green circle and the green diamond presents other South African strains that were previously published. All positions with gaps and missing data were removed. MEGA11 was used to perform the evolutionary analysis.
Figure 3.12	Coverage maps of consensus contigs for UL97 obtained with Genome detective. The Y-axis represents the depth of coverage, and the X-axis represents the position of the sequence in the context of the complete genome.
Figure 3.13	A maximum likelihood phylogenetic tree was constructed for the UL97 gene of Human HCMV, with the Merlin strain (MH036940.1) serving as the root. Bootstrap values above 50 are shown. Genes identified in this study are denoted by red dots, while the root is marked by a black square.

List of abbreviations

°C	Degrees Celsius
µL	Microliters
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Treatment
BLAST	Basic Local Alignment Search Tool
bp	base pairs
CAD	Coronary Artery Disease
CDC	Centres for Disease Control and Prevention
CI	Confidence Interval
CMV	Cytomegalovirus
CSF	Cerebrospinal Fluid
Ct	Cycle Threshold
CV	Coefficient of Variation
CY5	Cyanine 5
DNA	Deoxyribonucleic Acid
dNTP	deoxyribonucleotide triphosphate
E	Early
EDTA	Ethylenediaminetetraacetic acid
FAM	6-carboxyfluorescein
FDA	Food and Drug Administration
gB	Glycoprotein B
GenBank	Genetic Bank
gH	Glycoprotein H
gL	Glycoprotein L
gM	Glycoprotein M

gN	Glycoprotein N
gO	Glycoprotein O
HEX	Hexachlorofluorescein
HIV	Human Immunodeficiency Virus
HKY	Hasegawa-Kishino-Yano
HLA	Human Leukocyte Antigen
HSCT	Haematopoietic Stem Cell Transplantation
IE	Immediate Early
IgG	Immunoglobulin G
IRL	Internal Repeat Long
IRS	Internal Repeat Short
IU	International Units
kbp	kilobase pairs
L	Late
MCL	Maximum Composite Likelihood
MHC	Major Histocompatibility Complex
Min	Minutes
mL	Millilitres
N/A	Not Available
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NHLS	National Health Laboratory Services
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
pp65	Phosphoprotein 65

ROX	Rhodamine X
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SD	Standard Deviation
Sec	Seconds
SOT	Solid Organ Transplantation
T1DM	Type 1 Diabetes Mellitus
TAE	Tris-Acetate-EDTA
Taq	Thermus aquaticus
TRL	Terminal Repeat Long
TRS	Terminal Repeat Short
UI	Uncertainty Interval
UL	Unique Long
UL51	Unique Long 51
UL54	Unique Long 54
UL55	Unique Long 55
UL56	Unique Long 56
UL89	Unique Long 89
UL97	Unique Long 97
USA	United States of America
UV	Ultraviolet
VIC	6-carboxyrhodamine
VL	Viral Load

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1. INTRODUCTION

Cytomegalovirus (CMV) is a double stranded DNA virus that affects individuals of all ages and leads to opportunistic infections of global concern. CMV is primarily transmitted through direct contact with or exchange of infected bodily fluids, such as genital secretions, breast milk, blood, and saliva (1). The virus demonstrates tropism for the mucous membrane that lines the surface of various bodily tissues (2). CMV seroprevalence is markedly higher in developing countries, nearing 100%, compared to developed countries, where seroprevalence rates are estimated at 45% (3). Specific populations with varying degrees of immunosuppression are at risk for severe CMV disease. These populations include neonates with congenitally acquired infection, people living with advanced HIV/AIDS, transplant recipients, the elderly, individuals with chronic conditions and those using immunomodulatory agents. CMV disease, if not treated timeously, can result in severe outcomes, including retinitis, leading to blindness, and life-threatening pneumonia and gastrointestinal disease (4).

CMV can be classified into distinct genotypes based on the glycoprotein B (gB). Five main genotypes of CMV have been identified, designated as gB1-gB5. While other glycoproteins like gH, gL, gM, gN, and gO can also be used for characterisation, the UL55 encoded gB is the most widely accepted and preferred method for distinguishing between CMV genotypes (5, 6) The FDA has several-approved treatment options for CMV disease including ganciclovir and its prodrug valganciclovir, foscarnet and cidofovir. Furthermore, FDA has approved letemovir for prophylaxis against CMV disease in specific transplant recipients. Ganciclovir and valganciclovir target the UL97 gene and are considered first-line treatment options, whilst foscarnet and cidofovir which target UL54, are second line drugs (1, 7, 8). There have been reports of mutations in these drug target genes resulting in resistance to treatment (9, 10). Currently there is a paucity of data on this subject, and up to date data on circulating genotypes and their associations with drug resistance is lacking (11).

1.1.1 STUDY RATIONALE

1.1.1.1 Study problem

CMV has great adaptability to its human hosts, estimated to infect 60%-70% of the population in developed countries, and close to 100% in developing countries. In otherwise healthy and asymptomatic individuals, infection with CMV is usually unremarkable (3,12). Infection by CMV is frequently detected in salivary glands and bodily fluid of healthy individuals showing no symptoms. In contrast, CMV infection in immunocompromised individuals can be life threatening (13). Three groups are disproportionately affected by CMV disease: infants born with congenital CMV, individuals living with HIV/AIDS, and transplant recipients, who are more susceptible to CMV due to their compromised immune status (14). It is the second most common congenital viral infection after HIV in developing countries (15).

Currently, the drugs licensed for the treatment of CMV disease and approved by the FDA are ganciclovir and its prodrug valganciclovir, as first-line treatment options, together with cidofovir and foscarnet as second-line treatment options (16, 17). In South Africa, CMV disease is treated with ganciclovir and its prodrug valganciclovir, with limited availability to second-line options (1). The clinical emergence of antiviral resistance poses a threat amongst different high-risk groups with CMV disease. Antiviral resistance, results from the selection of mutations in the UL54 and UL97 genes. The genetic diversity of CMV genotypes, including resistance mutations among South African patients, is unclear (14).

The current standard practice is that the possibility of drug resistance among patients is based on clinical suspicion together with persistently raised viral loads (VL) greater than 1250 international units per millilitre (IU/mL) in the setting of an appropriate treatment schedule and duration (17). This approach is flawed because an increase in viral load may be attributed to non-compliant drug administration and the general immune status of the patient (18). In some cases, viral load may not always accurately reflect the disease severity or resistance. Failure to correctly identify drug resistance may result in delayed or inappropriate treatment, and poor clinical outcomes. This creates a need for a laboratory method that is accurate in diagnosing CMV drug resistance to inform appropriate treatment. This is particularly relevant in South Africa, a country with a high burden of immunocompromised persons and congenital CMV

disease. There is paucity of data on the use of genotypic assays to detect and diagnose CMV drug resistance in South Africa (11).

1.1.1.2 Aim and objectives

1.1.1.2.1 Aim

The aim of the study is to characterise CMV strains to describe virus diversity and drug resistance mutation profiles associated with suspected clinical resistance among patients with CMV disease, South Africa, 2022-2023.

1.1.1.2.2 Objectives

- To amplify and sequence CMV UL55 genes for characterisation of CMV strains and determine the diversity of gB genotypes among South African patients, 2022 to 2023.
- To amplify and sequence drug resistance associated CMV genes UL54 and UL97 to describe mutation profiles associated with CMV genotypes among South African patients, 2022 to 2023.

1.2. LITERATURE REVIEW

1.2.1 History and epidemiological review

For over a century, CMV has been recognised as a significant viral pathogen associated with human infection (19). CMV was first observed in 1881 by Ribbert and the histological appearance of CMV infection was initially described in 1904, and it was officially identified in 1953 by Weller, following advancements in cell culture techniques (20).

CMV is a ubiquitous viral pathogen that infects individuals of all ages and has no notable seasonal fluctuations (13, 21). The transmission of CMV is primarily through direct contact with or exchange of infected bodily fluids such as respiratory droplets, saliva, genital secretions, blood and breast milk (22). Other modes of transmission include organ transplantation, and perinatal transmission (3, 23, 24). The natural course of CMV infection begins with primary infection followed by latent infection with intermittent viral reactivation. In some cases, re-infection with a different strain of CMV may occur, resulting in a new non-primary infection (22). In immunocompetent individuals, including children, primary CMV infection is usually asymptomatic, but may present with illness characterised by an acute flu-like syndrome with persistent fever, night sweats, profound fatigue, muscle pain, joint pain and transaminitis (12). Among immunologically compromised patients, CMV may cause opportunistic disease. Individuals who have undergone organ transplantation, long-term steroid or immunomodulator use, or have immunosuppressive conditions are especially vulnerable to opportunistic disease leading to systemic involvement, organ-specific diseases, and death (22, 25). CMV can also result in congenital infection which is due to in utero vertical transmission (13).

A systematic and meta-analysis review by Zuhair *et al.* reported global CMV seroprevalence in the general population to be at 83% (95% uncertainty interval (UI): 78-88), 86% (83-89) in women of childbearing age, and 86% (82-89) in blood donors (21). In developed countries, IgG antibodies, which indicate exposure, were reported to be around 40% in developed countries and nearing 100% in developing countries (25). Age-related increases in CMV seroprevalence were observed by Cannon *et al.* (3). Based on these studies, seroprevalence was 20–30% higher in non-white groups compared to white groups, with certain non-white

populations approaching a seroprevalence of almost 100% (3). Further studies have found infections to be more prevalent in women of child-bearing age ranging from 45% to 100% with Asia, Africa, and South America having the greatest seroprevalence of CMV (3). The lowest seroprevalence rates were observed in Western Europe and the United States of America (USA) (26). The World Health Organization (WHO) and the Centres for Disease Control and Prevention (CDC) stated that more than 50% of adults in the USA are CMV infected by the time they reach 40 years old, and one in three children contract CMV by the time they are five years old (27). An estimated 1 in 200 neonates world-wide and 1 in 30 newborns in USA contract CMV infection from their mothers. The severity of foetal CMV infection is dependent on the CMV sero-status of the mother and the timing of infection in the gestational period. Unborn babies of seronegative pregnant women who develop a primary infection during pregnancy are at highest risk for severe disease, but transmission from seropositive mothers can also have serious repercussions (26, 28, 29).

The clinical presentation of CMV disease post-transplant differs based on multiple factors including the type of transplantation, serological compatibility between donor and recipient, the type of immunosuppressive drugs used and the presence of additional risk factors for illness (30, 31). In solid organ transplant patients, the highest risk is associated with a serological mismatch resulting in transmission of CMV from the donated organ due to reactivation of latent virus or after primary infection of seronegative patients (32–34). The highest incidence of CMV disease was found to be in lung or heart-lung transplant patients, with the incidence between 50% and 75%, followed by approximately 50% of pancreas or kidney-pancreas transplant patients. The incidence of CMV disease was found to be between 9% and 23% after heart transplant, between 22% and 29% after liver transplant and 8% to 32% after kidney transplant. In patients undergoing allogeneic haematopoietic stem cell transplantation (HSCT), the incidence of disease was found to be 30% and approximately 5% in autologous HSCT transplant patients (31, 34).

Further risk factors include intense immunosuppression, administration of lymphocyte depleting antibodies, acute rejection, the presence of other viral infections, and genetic polymorphism (33, 34).

The prevalence of CMV disease in HIV patients varies depending on factors such as CD4 cell count, HIV viral load, and geographic location. The widespread availability of antiretroviral therapy (ART) has resulted in a decreased frequency of CMV disease in HIV positive individuals (35). During the pre-ART era, CMV disease was a common complication in patients with AIDS particularly those with CD4 cell counts <50 cells per microlitre (μL), affecting approximately 40% of individuals during their lifetime (36).

1.2.2 Taxonomic classification

CMV belongs to a family of viruses known as *Herpesviridae,* subfamily *Betaherpesviridae*, genus *Cytomegalovirus.* It is one of eight herpesviruses that infect humans and is reported to be the largest human herpes virus (37–39).

1.2.3 Morphological and physiological characteristics

CMV is the largest in the family of *Herpesviridae*, having a diameter of approximately 220 nanometres (nm). Similar to other family members, CMV is an enveloped virus that comprises a double stranded DNA genome (40).

A CMV viral particle that has reached its full maturation measures between 150-200nm in diameter (41, 42). The virus particle is made up of three distinct parts namely, the icosahedral nucleocapsid with a diameter between 100-110nm, the tegument layer and the viral envelope, which is a host-derived lipid bilayer (**Figure 1.1**). The nucleocapsid houses the virus DNA and is made up of one hundred and sixty-two capsomeres. Twelve of the capsomeres are pentameric, with the remaining 150 capsomeres being hollow along half of the axis at a hexagonal cross section (43). When visualised by electron microscopy, the nucleocapsid, appears to be highly electron-dense (44).

The protein-rich tegument houses the virus particle, encompassing a few of the dominant targets for T-lymphocyte response including the phosphoprotein-65 (pp65) (45–47). The tegument is enclosed by the envelope layer and consists of numerous virally encoded glycoproteins (g) together with complexes known as the gB complex, the glycoprotein M (gM)

and N (gN) complex, and finally the glycoprotein H (gH), glycoprotein L (gL) and the glycoprotein O (gO) complex (47, 48).

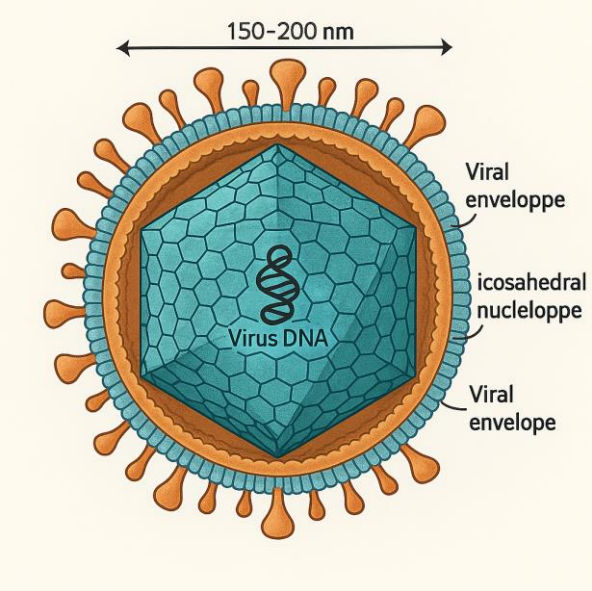

Figure 1.1 Structural diagram of a mature CMV virion showing the viral envelope, tegument layer, and icosahedral nucleocapsid containing viral DNA [*image created with BioRender* (49)]

As with other herpesviruses, CMV is vulnerable to degradation under acidic conditions, exposure to lipid solvents, and high temperatures. At 37°C, the virus has a half-life of around 60 minutes, indicating moderate instability. Furthermore, CMV is relatively unstable when stored at -20°C, requiring storage at -70°C or lower to preserve its infectious properties (42, 50).

1.2.4 Genome structure, organisation, and replication

The CMV genome consists of 235 kilo base pairs (kbp) of DNA, encoding 200 genes within 732 open reading frames (ORFs) (**Figure 1.2**). This intricate genetic structure allows CMV to interact with host cells, ensuring its persistence and transmission. The potential proteins

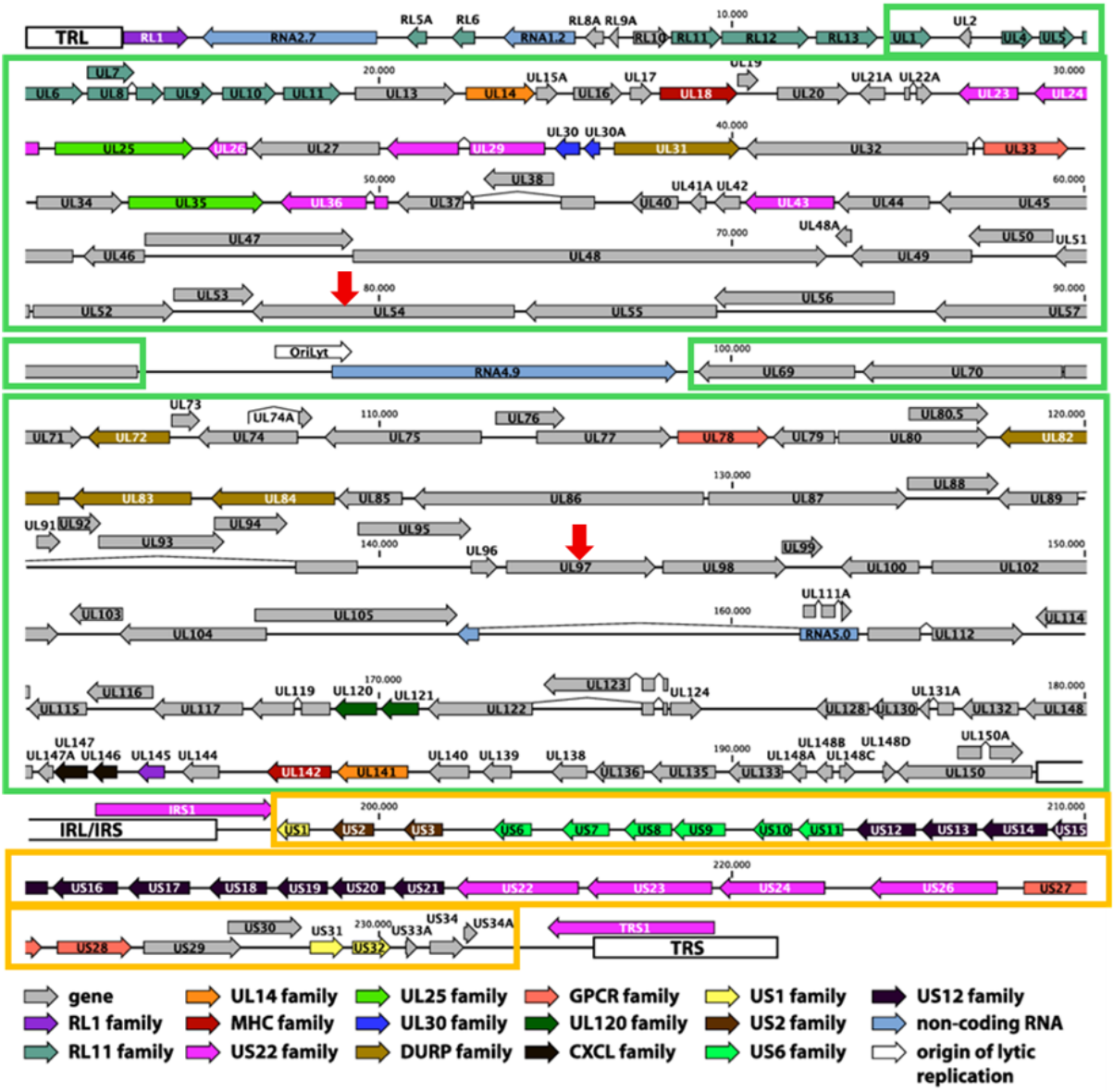

Figure 1.2 Genomic map of Human Cytomegalovirus (HCMV). The genome consists of unique long (UL, green boxes) and short (US, orange boxes) regions flanked by repeats and major gene families are color-coded. Study target genes are indicated by down arrows: red arrows: UL54 and UL97, blue arrow: UL55 and goldish arrow: UL83 (adapted from Sijmons *et al.*, 2014) (63).

encoded by these ORFs have been investigated, but the functions of most remain unknown, with only a handful having been defined (51). However, notable features have been identified, including nine distinct gene families and a substantial number of genes encoding glycoproteins. Furthermore, CMV's genome exhibits significant homology with key immune regulators, such as Human Leukocyte Antigen (HLA) class I genes and G protein-coupled receptor genes (26, 52). The genome exhibits an E-type structural organisation (**Figure 1.2**), characterised by two

large reverse-complement domains, designated as long (L) and short (S). The L and S domains are flanked by terminal inverted repeats at their extremities, while internal inverted repeats are strategically positioned within each domain (53, 54). Each domain is further subdivided into three distinct regions, namely, central unique, unique long (UL), and unique short (US) (55). During replication, the UL and the US segments can invert relative to one another through recombination at inverted repeats, generating four possible genomic isomers (56).

A study by Hega *et al.* reported that various clinical isolates of CMV demonstrate genetic variations in several key genes, particularly those encoding envelope glycoproteins, including UL55 (gB), UL73 (g N), and UL75 (g H) (**Figure 1.2**) (57). Glycoprotein B (gB) is an abundant and highly conserved glycoprotein, that plays a crucial role in the viral life cycle. Earlier research has established that gB facilitates both the initial entry of the virus into host cells and its subsequent spread from cell to cell. Glycoprotein H exhibits genetic variations leading to two main genotypes, gH1 and gH2, which may impact viral entry and immune interactions. Furthermore, the highly polymorphic UL73 gene encodes glycoprotein N (gN), which is classified into seven distinct genotypes (gN1, gN2, gN3a, gN3b, gN4a, gN4b, and gN4c), highlighting the extensive genetic diversity of CMV (51).

1.2.5 Replication, pathogenesis and virulence factors

The long history of CMV together with other *Herpesviruses* coevolving with their respective hosts has resulted in the virus' development of various complex strategies to evade the immune system of the host, enabling them to survive and maintain persistent infection (58). CMV infections are host-specific with the ability to infect a wide range of cells. The broad cellular range allows for efficient spread within the host and transmission to other persons (59, 60).

The establishment of CMV infection may differ at a cellular level based on receptor interactions between the host cell and the virus (61). Once CMV infection has been established, gB, which is a crucial target for virus-neutralising antibodies, plays a significant role in the initial phases of the infection (62). It facilitates the attachment and the binding of CMV to the host cells by forming a linkage with heparin sulphate glycosaminoglycans (HS-GAGs) on the cell surface (**Figure 1.3**).

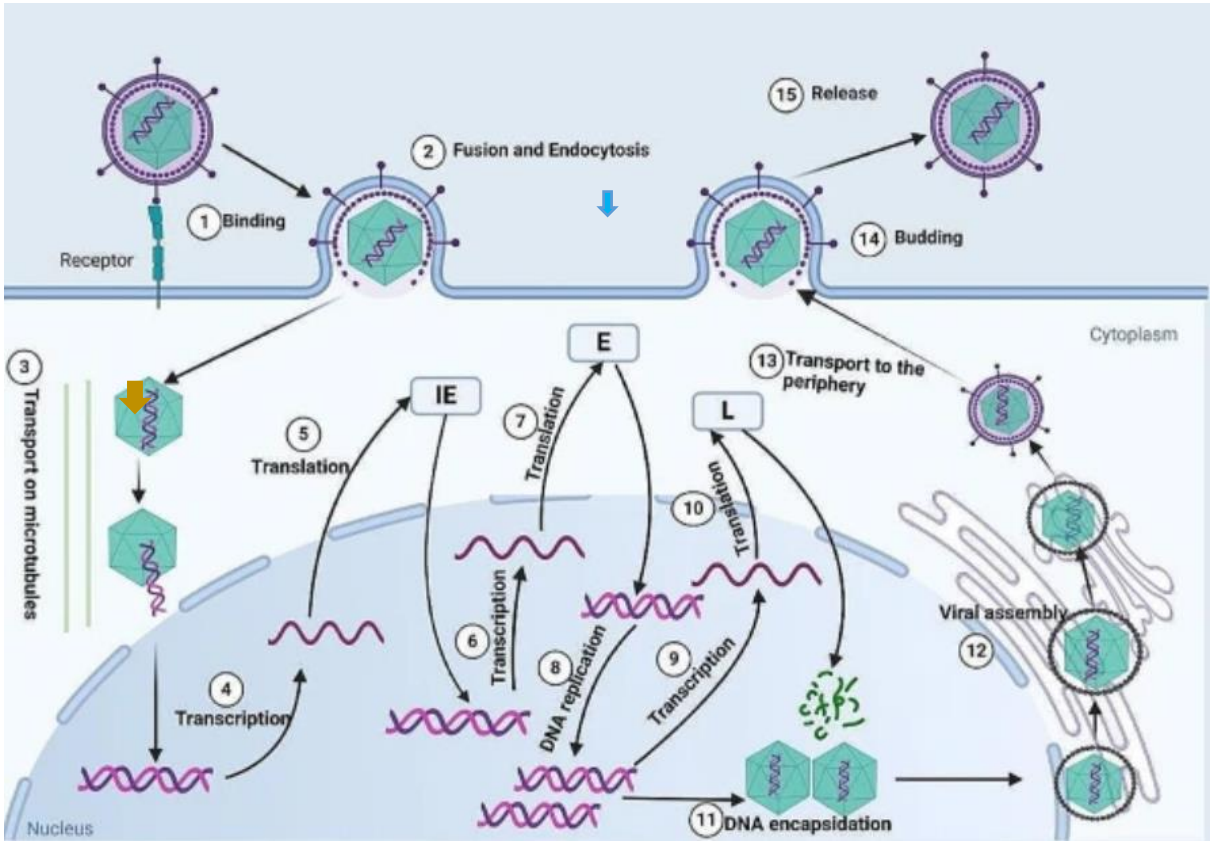

Figure 1.3 Cytomegalovirus (CMV) Replication Cycle. The diagram illustrates the replication cycle of CMV. The process begins with the virus entering the host cell, where the viral DNA is released and transported to the nucleus. Within the nucleus, the viral genome is replicated, and viral proteins are synthesized. These components then assemble to form new viral particles. Finally, these mature virions are released from the cell to infect new host cells, continuing the cycle (64). *Permission to use the figure was granted, see appendix A and B.*

The interaction between CMV and host cells facilitates the virus's initial attachment and subsequent penetration through the cell membrane, launching the infection process (60, 65). The virus relies on distinct glycoprotein complexes to enter different cell types: the gH/gL complex for fusion with fibroblast plasma membranes, and the gH/gL/UL128-131 complex for entry into epithelial and endothelial cells via micropinocytosis and endosomal fusion (65).

Subsequent to glycoprotein complexes attaching to the host cell, the alpha-helical coiled domain in the gB and gH fuses and integrates into the cell membrane (**Figure 1.3**), resulting in the release of the capsid into the cytoplasm (66). The capsid is then transported to the nucleus via microtubules, where the viral genome is released, and transcription begins (42, 67). The cycle of replication comprises of three phases, namely, immediate early (IE), early (E), and late (L) phases (Figure 1.3), which sequentially coordinate viral replication and growth (58). The predominant function of IE proteins is to facilitate the replication of the viral genome and alter the immune response of the host, allowing for smooth viral replication and evasion of the host's immune system (68, 69). Other functions of the IE genes may be suppressed upon infection, leading to a latent infection where the virus enters a state of dormancy, characterised by viral genome maintenance without active replication or viral production (70, 71). The host cell RNA polymerase II facilitates the transcription of immediate-early (IE) proteins in the nucleus. Notably, tegument proteins from the infecting virus particle serve as transactivators, enhancing the expression of IE genes (42, 72, 73). The IE proteins significantly play a role in regulating the subsequent stages of replication and modulate cellular processes, including the expression, transport, and presentation of class I Major Histocompatibility Complex (MHC) proteins, also known as HLA, to the proteasome for processing (72). The production of late (L) proteins is initiated by early (E) proteins. Primarily, the viral L genes result in the production of structural proteins that play an essential role in the assembly and release of new viral particles (58, 61, 71). The capsid assembly takes place in the nucleus of the cell, while the viruses are exported and enveloped at the inner nuclear membrane (and possibly other cellular membranes). Finally, mature virions are released from the cell, around 72-96 hours post infection, completing one lytic CMV infection (42, 58, 73).

Similar to other family members, CMV infection results in subsequent periods of latency characterised by restricted lytic gene expression with the absent production of viral particles (70). During the latent period the viral genome exists as an extrachromosomal plasmid with minimal replication, likely to remain undetected by the immune system of the host. Under specific stimuli, latent infections can intermittently reactivate leading to productive, lytic infections. The outcome of reactivation may facilitate ongoing viral transmission and further be associated with either asymptomatic or symptomatic disease of varying severity (74).

In addition to the crucial proteins for replication and structure, CMV produces specialised proteins that help it evade detection and attack by the immune system, disguising itself as a host protein (molecular mimicry) and evading immune responses (56, 75). The virulence factors (**Table 1.1**) play a very important role in viral reactivation and persistence, particularly in hosts who are immunocompromised (76–78).

Table 1.1 Key CMV Virulence Factors and Their Immune Evasion Mechanisms.

Gene/Protein	Function/ mechanism	Immune targeted
UL18	Mimics MHC class I to inhibit NK cell killing	MK recognition
UL40	Stabilizes HLA-E to prevent NK activation	NK toxicity
US2, US3, US6, US11	Degrade or retain MHC class I, blocking antigen presentation	CD8+ T-cell response
UL111A	Viral IL10 homolog suppresses inflammatory cytokines	Innate and adaptive immunity
UL141	Prevents expression of death receptors	MK cell and apoptosis

1.2.6 CMV and comorbidities

CMV disease is associated with significant morbidity and mortality (79). It poses a substantial threat to vulnerable populations, such as organ transplant recipients, individuals living with advanced HIV/AIDS, and congenitally infected newborns. CMV infection has been tentatively linked to various multi-systemic comorbidities, such as cardiovascular disease, autoimmune disorders, neurological conditions, atherosclerosis, and inflammatory diseases like diabetes (80–84).

Although considerable progress has been made in the field, CMV disease continues to be an obstacle in transplant medicine, increasing the likelihood of complications, graft rejection and mortality, thereby undermining the success of the transplant procedure (85). The risk of developing CMV disease in this group without prophylaxis is influenced by a complex interplay of factors, including the transplant type, donor-recipient serological matching and choice of immunosuppressive drugs (31). The adverse consequences of CMV disease in

transplant recipients is due to the direct cytopathic effects of CMV on multiple organs, primarily leading to debilitating conditions such as pneumonia, gastrointestinal disease, hepatitis, encephalitis, and retinitis, which can severely compromise the patient's health and well-being (86, 87).

Prior to the introduction of ART, CMV disease was a major cause of morbidity and mortality in people living with advanced HIV/AIDS (88). It can result in a range of serious and potentially life-threatening conditions in this group, namely retinitis, oesophagitis, colitis, and gastrointestinal ulceration (89, 90). HIV and CMV have a bidirectional relationship, with each infection capable of intensifying the effects of the other. The immunosuppressive nature of HIV increases the likelihood of CMV reactivation, while concurrent CMV infection can further weaken the immune system, creating a self-reinforcing cycle of deteriorating health outcomes (91). Individuals with CMV disease experience marked acceleration of progression to AIDS and mortality, with a 2.5 times higher risk compared to those without CMV (88, 92). A study by Gianella *et al.* (2014) suggested that CMV could be facilitating immune activation and replication of HIV. In this study which was focusing on men living with HIV and undergoing ART, it was reported that the presence of CMV DNA in semen correlated with increased T-cell activation and proliferation in the blood, as well as significantly higher levels of HIV DNA in mononuclear cells. These findings suggested that CMV reactivation may contribute to the persistence of HIV DNA in the body, even when ART is effective in suppressing the virus (93). Furthermore, another study by Gianella *et al.* (2016) revealed that intermittent presence of CMV DNA in peripheral blood mononuclear cells during ART was correlated significantly with a slower decline in HIV DNA levels (94).

In neonates, CMV infection remains one of the leading congenital viral infections, with birth prevalence ranging from 0.2%-6% in developed countries and between 0.6%-6.1% in developing countries (95, 96). Congenital CMV infection poses a significant threat due to it being closely tied to serious health implications, including permanent neurodevelopmental disabilities and sensorineural hearing loss (97). A wide range of clinical manifestations are exhibited with congenital CMV both in the immediate short-term and long-term including, petechiae, jaundice, hepatosplenomegaly, and neurological abnormalities like microcephaly

(98, 99). The occurrence of foetal infection can follow primary maternal infection, maternal reinfection with a different strain, or reactivation of latent maternal infection (100).

Further research has revealed a notable relationship between infection with CMV and the development of type 1 diabetes mellitus (T1DM). Specifically, Eltayib *et al.* (2014) demonstrated a significant correlation between the presence of CMV seropositivity and the onset of T1DM in paediatric populations, implying that CMV infection may play a role in triggering the disease (101). Similarly, researchers have noted that the detection of CMV genetic material in pancreatic beta cells is associated with the presence of autoantibodies targeting islet cells in individuals with newly diagnosed diabetes, suggesting that CMV infection may initiate an autoimmune response that contributes to the development of T1DM (102). Furthermore, a study by Batchy *et al.* (2022) proved that CMV infection is also linked to disrupted metabolic profiles and severe insulin deficiency in individuals with T1DM (103).

In the setting of cardiovascular disease, CMV infection is reported to be associated with an increased risk of developing atherosclerosis, ischaemic heart disease, and myocardial infarction, ultimately leading to death (80). In the early 1940s, research by Paterson and Cottral hinted at a correlation between herpes viruses and atherosclerosis (104). Years later, an analysis of plaque and tissue samples revealed a significant presence of CMV antigen in atherosclerotic lesions (52%) compared to non-atherosclerotic tissues (29%), implying that CMV may trigger the development of atherosclerosis. Also, arterial walls were identified as potential sites for latent CMV infection (105). Arnon and colleagues investigated the relationship between anti-CMV antibody titres and coronary artery disease (CAD). They discovered that patients with high anti-CMV antibody titres had a higher rate of re-stenosis compared to seropositive patients with low titres. The researchers further concluded that high anti-CMV antibody titres may serve as a strong predictor for CAD and could potentially forecast re-stenosis after coronary angioplasty (106). This may be due to the ability of CMV to suppress the p53 protein, which normally helps with the regulation of cell growth. Without the p53 protein, vascular cells may proliferate uncontrollably, resulting in the formation of neointima, scar tissue in the vessels, potentially exacerbating CAD (107). Previous research has consistently shown that infection by CMV is associated with cardiovascular disease. For instance, studies found that patients with stable coronary artery disease and acute coronary syndrome had significantly higher levels of CMV DNA in their blood compared to healthy individuals (108, 109). Similarly, Beyaz *et*

al. discovered that patients with bilateral carotid artery stenosis had significantly higher levels of CMV DNA compared to those without bilateral carotid artery stenosis. Overall, the evidence suggests that CMV infection may play a role in the pathogenesis of cardiovascular disease (110).

1.2.7 CMV diagnosis

The critical role that CMV disease plays in morbidity and mortality rates in immunosuppressed individuals and congenitally infected neonates has led to the development of diagnostic techniques that enable prompt diagnosis and management (111). The most widely used method for the detection and diagnosis of CMV is real-time PCR, which detects CMV DNA in clinical samples like blood, urine, CSF, and amniotic fluid. Due to its high specificity and sensitivity, PCR has become the gold standard for the diagnosis of CMV, providing fast and reliable results (112). An alternative amplification-based method for nucleic acid detection is loop-mediated isothermal amplification or LAMP (114). In addition to PCR methods, serological tests, which detect CMV-specific antibodies IgM and IgG, are utilized. Although they can detect recent or past infections, serological tests may not help to identify clinically significant viral activity, particularly in immunocompromised patients or those with weak antibody responses, resulting in false negative results (113, 114). The CMV antigenemia assay is also a widely used diagnostic tool that detects the pp65 protein in blood leukocytes, enabling early identification and monitoring of CMV infection. This assay is particularly useful for high-risk individuals, as it predicts the onset of CMV disease, allowing for timely treatment, thus improving patient outcomes (115, 116). Furthermore, the antigenemia assay serves as a supplemental tool for diagnosing tissue-invasive CMV diseases, despite its varying sensitivity (116). Quantitative nucleic acid amplification tests (QNATs) have introduced viral load measurement as a critical tool in CMV management, these assays not only detect CMV nucleic acid but also quantify it (117). Although not a primary diagnostic test, viral load measurement is crucial for predicting disease risk, guiding pre-emptive therapy, monitoring treatment response, determining therapy duration, and detecting relapse or resistance (118).

1.2.8 Treatment, prevention and control

Prevention of CMV disease is crucial for high-risk populations, such as HSCT and solid organ transplant (SOT) recipients, as it significantly reduces the incidence of CMV-related complications, thereby decreasing morbidity and mortality rates (119). For the prevention of CMV disease and its associated complications, two main strategies are employed, post-transplantation antiviral prophylaxis, which involves proactive antiviral treatment, and pre-emptive therapy, which involves close monitoring for CMV reactivation and swift treatment if necessary (120). In certain cases, a combination of these approaches is employed, which involves the initial use of antiviral prophylaxis during the period of highest risk, followed by a switch to a surveillance strategy, which involves monitoring for CMV reactivation and initiating pre-emptive treatment if necessary (121). There are investigational findings suggesting that the effectiveness of CMV prevention strategies can be enhanced by incorporating cell-mediated immunity monitoring. This approach may allow for more personalised and optimised prevention measures, tailoring interventions to an individual's specific immune response (120, 122).

The development of CMV-specific antiviral agents was initially due to the need to treat CMV end-organ disease in patients with HIV/AIDS (9). The discovery of virus-encoded enzymes, essential for viral replication, led to the development of antiviral drugs. Importantly, CMV-specific enzymes have very distinct structural variations from enzymes of the cell, allowing the antiviral drugs to selectively target and inhibit them, thus preventing harm to the host cells (9, 123).

Drugs that are widely used for the treatment and prevention of CMV are ganciclovir and its prodrug valganciclovir, cidofovir, foscarnet, and letermovir (16). Ganciclovir and its prodrug valganciclovir as first-line treatments have broader approvals, including serving as prophylaxis in transplant recipients, and treatment of CMV retinitis and CMV-related gastrointestinal disease. The typical treatment regimen, depending on the disease severity, involves intravenous administration of ganciclovir for 2-3 weeks (16, 124, 125). As second line treatment foscarnet and cidofovir have more limited approvals, specifically for treating CMV retinitis in AIDS patients, while letermovir is approved as prophylaxis for CMV-positive recipients of HSCT (9, 125). Furthermore, ganciclovir and its prodrug valganciclovir, foscarnet, and cidofovir usage

is complicated by significant toxicity and the development of drug resistance, prompting the urgent need for research and development of less toxic and effective antiviral drugs (14, 125).

Ganciclovir is a synthetic compound that mimics the structure of 2'-deoxyguanosine, a natural building block of DNA (9). Characteristic of nucleoside analogues targeting *Herpesviruses*, ganciclovir needs to be phosphorylated within cells to become active and exert its antiviral effects (126). A UL97 (**Figure 1.2**) encoded viral protein kinase monophosphorylates ganciclovir, which is further phosphorylated by the protein kinases of the cell to create an active form. In its active form, ganciclovir inhibits viral DNA replication by blocking the insertion of deoxyguanine triphosphate. In addition to this, the drug inserts into the viral DNA resulting in slowed and eventually terminated elongation of the viral DNA (127). Valganciclovir, the L-vanyl ester of ganciclovir, upon ingestion gets converted to ganciclovir, and thus possesses similar properties to ganciclovir (9).

Foscarnet is a pyrophosphate analogue and has a different mechanism of action compared to ganciclovir and valganciclovir. Foscarnet comparatively, does not rely on phosphorylation by viral kinases (128). Foscarnet achieves the inhibition of the viral DNA polymerase activity encoded by UL54 by targeting the pyrophosphate binding site, disrupting the nucleotide addition cycle and preventing the 3' elongation of viral DNA (7). Although foscarnet is the drug of choice in patients who have developed ganciclovir resistance or are intolerant to ganciclovir due to severe neutropenia or leukopenia (129), the biggest limiting factor for use of this drug is renal toxicity (7).

Synthetic antiviral agent cidofovir is an acyclic nucleoside phosphonate, chemically designated as 1-[(S)-3-hydroxy-2-(phosphonomethoxy)propyl] cytosine dihydrate. It has a wide range of antiviral activity, demonstrating potent efficacy against a broad spectrum of DNA viruses (130). It is indicated for the clinical management of CMV retinitis in AIDS patients. The drug's extended intracellular half-life allows for infrequent administration, making it a convenient treatment option (9). Upon entry into the host cell, cidofovir becomes phosphorylated by the cellular kinases to its active disphosphoryl form. In this form it competitively inhibits the viral DNA polymerase, resulting in premature termination of viral DNA replication (131). The greatest hindrance with the use of this drug is its nephrotoxicity (7).

Letermovir is a chemical compound belonging to the 3,4-dihydro-quinazoline class, specifically a derivative of quinazoline-4-yl-acetic acid (120). It functions by inhibiting the terminase complex of the CMV DNA. It achieves this by disrupting the packaging of viral DNA into new virions by a complex comprised of UL56, UL89 and UL51 (**Figure 1.2**), which works to allow precise cleavage and packaging of the viral DNA (132, 133). This drug is rapidly absorbed upon administration, unlike other anti-CMV drugs that require activation (120). It also possesses minimal risk of drug-to-drug interactions, making it a suitable drug of choice in patients on multiple medications (132). Different from other antiviral agents, letermovir does not result in myelotoxicity or nephrotoxicity, providing a safer alternative for patients requiring CMV prophylaxis (120). However, a major concern with this drug is the risk of breakthrough CMV disease, especially after stopping prophylaxis. Research has shown that patients who discontinue letermovir may be more likely to experience significant CMV disease, which may necessitate subsequent antiviral therapy (134).

Other treatment options for prevention and treatment of CMV that are being explored include maribavir, brincidofovir, cidofovir derivatives, cyclopropavir, adoptive T-cell therapy, and a combination of therapies strategy (135–137). The innovation of these drugs was as a response to the growing issue of antiviral resistance and drug toxicity, particularly amongst patients undergoing prolonged suppressive therapy (132, 137).

1.2.9 CMV drug resistance

The emergence of CMV strains that are resistant to treatment was first noted as a major concern in patients who were suffering from HIV/AIDS (138, 139). Research consistently shows that the development of drug resistant CMV, identified by phenotypic and genotypic testing, is closely linked to the advancement or recurrence of CMV related disease, especially retinitis in patients undergoing treatment for prolonged periods (10, 140). The first research into the prevalence of ganciclovir resistance was in HIV/AIDS patients, which involved the analysis of the urine from patients (n=31) receiving intravenous ganciclovir treatment for CMV retinitis. The research found that the patients who received treatment for more than three months still had virus present in their urine, and 38% had developed resistance, which accounted for 8% of the total patient cohort (141). Since then, there has been further research conducted through larger studies to examine the emergence of resistant strains over time.

Research has continually reported that genetic mutations in the genome of CMV, particularly in the UL97 and UL54 genes (**Figure 1.2**), confer resistance to various antiviral drugs. UL97 gene mutations confer resistance to ganciclovir, while UL54 gene mutations confer resistance to multiple drugs, including ganciclovir, foscarnet, and cidofovir (**Table 1.2**) (14, 142). In the UL97 gene, mutations at codons 460, 594, and 595 have been reported as contributors to resistance to ganciclovir. Notably, the L595S mutation is particularly impactful, conferring a remarkable 9.2-fold decrease in ganciclovir susceptibility, making it a critical mutation to monitor in resistant CMV strains (143, 144). Meanwhile, in UL54, K513T mutations have been reported to be responsible for resistance against both ganciclovir and cidofovir while remaining sensitive (145). Further notable mutations include D413A and V823A, which are associated with resistance to both ganciclovir and cidofovir. Certain mutations in the exonuclease domain of UL54, such as K513N and K513R, have been found to enhance the enzyme's proofreading ability, leading to more efficient excision of nucleotide analogues and consequently contributing to drug resistance (146–148).

Table 1.2 CMV genetic mutations and antiviral drug resistance

Gene	**Key mutations conferring resistance**	**Antiviral drugs affected**
UL97	M460V/I, A594V, L595S	Ganciclovir
UL54	K513T, K513N, K513R, D413A, V823A	Ganciclovir, Foscarnet, Cidofovir

1.2.9 Vaccine development

CMV vaccine research began in the 1970s, with the development of live attenuated virus-based preparations (149). Live attenuated virus-based vaccines were the most considered approach for the development of CMV vaccine, as they presented a broad of viral antigens. However, the experiments and human studies proved otherwise, revealing that immune responses to be weaker than anticipated, and the ability of the vaccine to reduce the incidence of CMV infection only partial (150). Vaccine development of CMV is complicated by the ability of CMV virus to evade the immune system, and its ability to persist in the body long-term (151).

Presently, research in CMV vaccine development has advanced with various strategies, which included subunit, live attenuated, plasmid DNA, and vectored vaccines. The CMVPepVax trial highlights a breakthrough in creating a safer, non-infectious subunit vaccine using specific T-cell targets, which could protect high-risk groups, like hematopoietic stem cell transplant patients, from the risks associated with live vaccines (152). Additionally, there have been promising developments in plasmid DNA vaccines, which targets specific CMV proteins (pp65 and gB) and uses a poloxamer adjuvant. Significant reduction in CMV viremia was demonstrated in patients receiving hematopoietic cell transplants (153). Research has demonstrated that vaccines that are gB based can stimulate strong antibody response, which is crucial the vaccine development (154). While these vaccines are promising, their effectiveness remains a concern, as evidenced by the need for further optimisation to enhance T-cell responses and neutralising antibodies, which are important effective protection against CMV infection (153, 154).

The primary reasons for the development of CMV vaccine is for the prevention of congenital CMV disease. A key strategy to achieve this goal is to vaccinate both adolescent girls and boys before they become sexually active, thereby protecting them from CMV infection, as sexual transmission Is a major risk factor for CMV infection beyond childhood (153, 155).

CHAPTER 2

METHODOLOGY

2.1 ETHICAL CONSIDERATIONS

Ethical clearance was obtained from the University of Witwatersrand Human Research Ethics Committee (Medical) with project number: M210822. Permission was granted by the Gauteng Department of Health and Lancet Laboratories for retrieval of specimens sent for CMV qualitative and quantitative testing from their storage bank.

2.2 STUDY DESIGN

This was a laboratory-based study that employed a descriptive quantitative study design to determine the molecular characteristics of CMV strains and the resistance profiles among South African patients.

2.3 STUDY POPULATION

The study population included patients who seek care in the private and public sector from whom clinical samples were submitted for CMV qualitative and viral load testing. Samples from private sector patients were retrieved from Lancet Laboratories and for the public sector from the National Health Laboratory Service (NHLS) Virology Laboratory at Charlotte Maxeke Johannesburg Academic Hospital (CMJAH). The study included high risk patients of all ages who were either experienced in CMV antiviral treatment or naïve. Patient identifiers were anonymized and replaced by a laboratory number with only the study team having access to the study data. HIV status, CD4 count, CMV treatment and duration of treatment, and other laboratory results were not available. The study was conducted from 01 February 2022 to 31 November 2023.

2.4 SAMPLE SIZE AND STUDY SAMPLES

2.4.1 Sample size

Epi Info version 7.2.3.1 (Centers for Disease Control and Prevention, USA) was used to calculate the minimum required sample size. Assuming an expected CMV prevalence of 50% to give the largest sample size, at 80% power and 95% confidence and a population size of 999 999, a minimum sample size of 194 was required.

2.4.2 Study samples

Study samples included serum, plasma, whole blood, CSF, breast milk, nasopharyngeal aspirates, and urines on which CMV quantitative and qualitative testing were requested from Lancet Laboratories and NHLS Virology Laboratory at Charlotte Maxeke Johannesburg Academic Hospital. Samples that were haemolysed, indicated by dark brown color due to the release of hemoglobin, or had volumes less than 150µl were excluded for analysis in this study.

2.5 LABORATORY TESTING

Strict adherence to Good Clinical Laboratory Practice guidelines was ensured throughout the study, with the inclusion of appropriate controls to validate assay performance and prevent contamination (156). Quality control measures were implemented to guarantee the accuracy, precision, and reliability of test results, thereby ensuring the integrity of the data generated. Furthermore, robust standard operating procedures were followed in the laboratory of the Division of Virology, University of the Witwatersrand and NHLS to minimise the risk of sample mix-up, contamination, or errors in data interpretation.

2.5.1 DNA extraction

DNA was extracted manually using the DaAngene nucleic acid isolation and purification reagents from Sun Yat-Sen university (Guangzhou, China). Two of the 48 reaction test kits were prepared by adding 10mL absolute ethyl alcohol into each of the inhibitor remover bottles, and 44mL absolute ethyl alcohol into each of the deionised solution bottles. 50µL of proteinase K was added into a sterile 1.5mL centrifuge tube, followed by 200µL of specimen and lysis

solution. The mixture was vortexed and centrifuged at 140000g and then incubated for 10 minutes at 72ºC, simultaneously preheating the elution buffer at 72ºC. Afterwards, 250µL of ethanol was added to the lysate mixture, which was vortexed and transferred into a spin column, centrifuged at 12000g for a minute at room temperature before fitting the spin column into a new collection tube. Then 500µL of deionised solution was added as wash buffer to the spin column, which was centrifuged at 12000g for one minute and fitted into a new collection tube. This wash process was repeated three times. The spin column was then placed into a new collection tube and centrifuged at 14000g for three minutes to remove residual alcohol. The spin column was then fitted into a new 1.5mL centrifuge tube. The tubes were incubated for two minutes at the temperature of 72ºC with their caps open to evaporate residual ethanol whereafter 50µL warm elution buffer was added to each spin column and centrifuged at 14000g for one minute to collect the nucleic acid in the 1.5mL centrifuge tubes. Nucleic acid samples were stored at a temperature of -20ºC.

A summary of all laboratory investigations to follow (PCR and sequencing) is summarised in **Table 2.1** below.

Table 2.1 Summary of CMV genes targeted for PCR and sequencing in this study with the expected amplicon size indicated.

Gene	Protein	Assay	Purpose	Amplicon Size (bp)
UL83	pp65 protein	Argene CMV-R kit	qualitative PCR for CMV detection	not indicated
UL55 (gB)	glycoprotein B	In-house Real-Time multiplex PCR	CMV genotyping (gB1–gB4)	~100–200 bp per target
UL55 (gB)	glycoprotein B	Nested Conventional PCR + NGS	Sequence full gB genes for phylogenetic analysis	2709 bp
UL54	DNA polymerase	Nested PCR + NGS	To detect drug resistance mutations	1654 bp
UL97	phosphotransferase	Nested PCR + NGS	To detect drug resistance mutations	1298 bp

PCR: polymerase chain reaction; NGS: next generation sequencing.

2.5.2 Qualitative real-time PCR

Argene CMV-R gene detection kit (BioMérieux, France), targeting UL83 gene (**Figure 1.2**) coding for ppUL83 protein was used to confirm the presence of CMV in each of the DNA extracts prepared from samples obtained from both NHLS and Lancet laboratories, making use of the Taq Polymerase enzyme. The kit performs amplification using the 5' nuclease TaqMan® technology (patent nº: 5210015, 5487972) also called hydrolysis probes. The ready to use premix contained the primers, dNTPs, amplification buffer, *Taq* Polymerase enzyme, CMV specific primers and probes and internal control specific primers and probes. Ten microlitres of the test sample or positive control were added to 15µl of the amplification premix, then transferred to the tubes. The mixture was homogenised using a pipette, centrifuged, and transferred to the thermocycler. The CMV R-gene was detected in FAM channel and the internal control in VIC channel. The results were generated in the form of amplification curves and cycle threshold (Ct) values. Positive results were indicated by an exponential increase in fluorescence signal. Ct values ≤38 was accepted as positive. The conditions employed for PCR cycling are shown in **Table 2.2**.

Table 2.2 PCR cycling conditions for CMV qualitative test.

Steps	**Temperature**	**Time**	
Taq polymerase activation	95ºC	15 minutes	1 cycle
Denaturation	95ºC	10 seconds	45 cycles
Annealing/elongation	60ºC	40 seconds	
Hold	4ºC	∞	

2.5.3 CMV UL55 genotyping

2.5.3.1 UL55 genotyping by real-time PCR

Samples positive for CMV DNA by qualitative PCR were genotyped using an in-house real-time PCR developed with primers and probes (**Table 2.3**) for identification of genotypes gB1 to gB4 as described by Pang *et al.* (2008) (157).

Table 2.3 Summary of genotype-specific primers and probes used for CMV UL55 genotyping

Genotype	Primer sequence (5'-3')	Probe sequence (5'-3')	Channels
gB1	CATACGACGTCTGCTGCTCACT(F) GCTGACCGTTTGGGAAGAAG (R)	TCGATCCGGTTCAGTCTCT	FAM+
gB2	TCTTTGGTGGAATTGGAACGT (F) TGTCACTCGTACTTCTTCTGGTCCTA(R)	ATCCAGTCTGAATATCA	HEX
gB3	TGTTGGAACTGGAACGTTTGG (F) TGCCCGTACTTCTCTTGGTTCT(R)	CGGTGTGAACTCCA	Cy5
gB4	AAACGTGTCCGTCTTCGAAACT(F) TCCACCAGAGATTTTTGCTTGA(R)	CCGGCGGACTAGTAGT	Rox

The real-time PCR was performed using the MOLBIOL lyophilised 1-step RT-PCR polymerase mix (TIB MOLBIOL, Germany). The enzyme master mix was reconstituted with 990µL of ice-cold reconstitution buffer. Ten microlitres of the premixed master mix was added to the tube with 4µL of 10mM primer mix and 1µL of 5mM probe mix. Five microlitres of the template DNA was then added. The mixture was homogenised with a pipette, centrifuged, and transferred to the QuantStudio 5 real-time PCR system (ThermoFisher Scientific, USA). The PCR cycling conditions used are provided in **Table 2.4**.

Table 2.4 Real-Time PCR Conditions and Parameters for CMV UL55 Genotyping (157).

Steps	Temperature	Time	
Taq polymerase activation	95ºC	10 minutes	1 cycle
Denaturation	94ºC	20 seconds	45 cycles
Annealing/elongation	60ºC	1 minute	
Hold	4ºC	∞	

2.5.3.2 UL55 genotyping by conventional PCR and sequencing

Nested PCR of the UL55 gene (gB protein gene)

The samples in which CMV was detected in the initial qualitative screening but did not yield genotype results by real-time PCR were subsequently amplified using a conventional nested PCR assay. Briefly, for both the 1st round and 2nd round PCR reaction a 50µL reaction mixture was prepared by combining 2µL of the premixed Long Amp Hot Start Taq Polymerase mix with 10 mM concentrations of both forward and reverse primers (**Table 2.5**) and water to make up to total volume based on the DNA template volume for each PCR. For the 1st round PCR 5µL of DNA/TNA extract was used and for the 2nd round PCR, 2µL of the first-round reaction product was used as the template. The primers listed in Table 2.4 were previously described by Nelson *et al.* (2019) and are detailed in the supplementary material of their publication (158). The conditions for both 1st round and 2nd round PCR cycling are outlined in **Table 2.6**.

Table 2.5 Primer sets employed for the conventional PCR of UL55 (164).

Primer name	**Primer sequence (5'-3')**	**PCR**	**Amplicon size**
FULLGB-R1	TTGAAAAACATAGCGGACCG	1st round reverse primer	~2784 bp
FULLGB-F1	ACACGCAAGAGACCACGACG	1st round forward primer	
FULLGB-R2	TCAGACGTTCTCTTCTTCGT	2nd round reverse primer	~2709 bp
FULLGB-F2	ATGGAATCCAGGATCTGGTG	2nd round forward primer	

Table 2.6 Conditions utilised for the conventional PCR of UL55 gene

Steps	Temperature	Time	
Initial denaturation	94ºC	30 seconds	
Denaturation	94ºC	10 seconds	30 cycles
Annealing/elongation	60ºC	2,30 minutes	
Hold	4ºC	∞	

Gel electrophoresis

All PCR amplicons were subsequently analysed by gel electrophoresis to detect successful amplification of the target DNA sequence. A 1% agarose gel was prepared in 1X Tris-Acetate-EDTA (TAE) buffer, stained with 3µL of gelred (Biotium, San Francisco), resulting in a final dye concentration of 0.03µg/mL. Gel electrophoresis was performed at a constant voltage of 120V for a duration of 45 minutes. The anticipated amplicon size was approximately 2709 bp. Following gel electrophoresis, the DNA fragments were visualised under ultraviolet (UV) illumination using the Syngene Vacutec gel doc system (Cambridge, United Kingdom).

Sequencing and phylogenetic analysis of UL55 gene fragments

Samples that demonstrated amplification of the UL55 locus via conventional PCR were deemed suitable for further genetic characterisation. The resulting amplicons were isolated and subjected to DNA next generation sequencing (NGS) at the Sequencing Core Facility of the National Institute for Communicable Diseases (NICD), Johannesburg, South Africa to elucidate the precise nucleotide sequence of the UL55 gene for the purpose of characterisation. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter) following manufacturer's instructions, followed by tagmentation and addition of index primers using the Illumina Nextera XT library preparation kit (Illumina Inc, USA). DNA libraries were prepared for sequencing on the Illumina NextSeq 500 system using indexed paired-end libraries (Illumina Inc, USA).

Assembly of sequence reads.

The sequence reads were extracted and assembled using Genome detective virus tool (https://www.genomedetective.com/) and aligned using MAFFT (version 7) multiple alignment program for amino acid and nucleotide sequences (159). BioEdit software was then used for sequence curation (160), which enabled the visualisation and editing of the DNA sequences. The aligned sequences were used to generate a phylogenetic tree using MEGA11, a molecular evolutionary genetics analysis tool (161). The phylogenetic tree was constructed using the maximum likelihood method, which allowed for the inference of evolutionary relationships between sequences. The tree was used to identify clusters and relationships between sequences, providing insights into the genetic diversity and evolution of the CMV strains.

Analysis and retrieval of sequences from GenBank

To identify closely related sequences, patient consensus sequences were compared using NCBI's BLAST tool (162). Relevant sequences representing diverse strains were retrieved from GenBank for comparative analysis with accession numbers: OQ511294.1, OQ511292.1, OQ511300.1, OQ127294.1, GU365820.1, GU365821.1, OQ511297.1, OQ511296.1, KT987993.1, KT987992.1, 4 MK157427.1, OQ511295.1, OQ511302.1, KT987991, KT987990.1, OQ511304.1, OQ511303.1, LC496860.1, LC425070.1, LC496859, GU365824.1, KT987995.1, GU365825.1, MK157450.1, GU365822.1, KT987994.1, GU365823.1. Following phylogenetic tree generation, two samples remained unassigned genotypes. To further resolve their genetic relationships, an additional phylogenetic analysis was conducted incorporating two reference sequences from GenBank: gB5 (accession numbers KY436014.1 and KY436009.1). This expanded analysis aimed to clarify the classification of the unassigned samples.

2.5.3.3 Phylogenetic tree

A maximum likelihood phylogenetic tree was constructed through the MEGA11 software using 500 bootstrap replicates to assess branch support Branch lengths represent evolutionary

distances between sequences (161). The Hasegawa-Kishino-Yano (HKY) model was employed to account for nucleotide substitutions.

2.5.4 Amplification and sequencing of UL54 and UL97

2.5.4.1 Amplification of UL54 and UL97

The Q5 high fidelity 2X master mix kit (New England Biolabs, Ipswich, MA) was used with the manufacturer's protocol as guide. To achieve a final volume of 50µL, 400nM of forward and reverse primers, as outlined in **Table 2.7** (UL54) and **Table 2.8** (UL97), were added together with 25µL of the master mix. Then 5µL of the DNA extract was used as template for the first round PCR and for the second round of the nested PCR the 1st round PCR products was used as templates. Simultaneous multiplex PCR assays were performed using specific primer pairs for UL97 and UL54 (163). The cycling conditions (**Table 2.9**) were consistent across both rounds of nested PCR.

Table 2.7 Primer sets for the conventional nested PCR of UL54 as previously described (163).

Primer name	**Primer sequence (5'-3')**	**PCR**	**Amplicon size**
UL54(long)-F1	CGCGTCGCCGTTGCACGTAG	1st round forward primer	2410 bp
UL54(long)-R1	AGTGTCGGTAGCCGTTTTTGCGA	1st round reverse primer	
UL54(long)-F2	AGTCCACGCCGCCTCATCTC	2nd round forward primer	1654 bp
UL54(long)-R2	GTGAGAGGCTGACAGCTTACGA	2nd round reverse primer	

Table 2.8 Primers for the conventional nested PCR of UL97 as previously described (163).

Primer name	Primer sequence (5'-3')	PCR	Amplicon size
UL97(long)-F1	GCCACCTTGGTGGACTCGGT	1st round forward primer	1844 bp
UL97(long)-R1	CTTTCGGCGCACCACCAGCA	1st round reverse primer	
UL97(long)-F2	GCAATCCCCGTCACGCCTCTG	2nd round forward primer	1298 bp
UL97(long)-R2	GGACGCGGAACGTGACGGTT	2nd round reverse primer	

Table 2.9 Conditions employed for the nested amplification of UL54 and UL97.

Steps	**Temperature**	**Time**
Initial denaturation	98ºC	30 seconds
35 cycles	98ºC 55ºC 72ºC	10 seconds 30 seconds 2.30 seconds
Final extension	72ºC	2 minutes
Hold	4ºC	∞

2.5.4.2 Gel electrophoresis

The amplicons for both the UL54 and UL97 were visualised by gel electrophoresis for the detection of successful amplification of the target genes. A 1% agarose gel was prepared in 1X Tris-Acetate-EDTA (TAE) buffer, stained with 3µL of gelred (Biotium, San Francisco), resulting in a final dye concentration of 0.03µg/mL. Gel electrophoresis was performed at a constant voltage of 120V for a duration of 45 minutes. The expected band size was 1844 bp for UL54 and 1298 bp for UL97 (14). Subsequent to gel electrophoresis, the resulting DNA fragments were visualised under UV illumination using the Syngene Vacutec gel doc system (Cambridge, United Kingdom). The target genes (UL54 and UL97) were pooled per patient for

the purposes of library preparation for NGS at the Sequencing Core Facility of the NICD, Johannesburg, South Africa.

2.5.4.3 Sequencing and phylogenetic analysis of UL54 gene fragments

Samples that demonstrated amplification products via conventional PCR were deemed suitable for further genetic characterisation. The resulting amplicons were isolated and subjected to DNA next generation sequencing (NGS) to elucidate the precise nucleotide sequence of the UL54 gene. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter) following manufacturer's instructions, followed by tagmentation and addition of index primers using the Illumina Nextera XT library preparation kit (Illumina Inc, USA). DNA libraries were prepared for sequencing on the Illumina NextSeq 500 system using indexed paired-end libraries (Illumina Inc, USA). The input DNA met the recommended specifications of 10ng/μL concentration in 10mM Tris-Cl buffer at pH 8.5.

Assembly of UL54 sequence reads.

The sequence reads were extracted and assembled using Genome detective virus tool v2.27 (164), and aligned using BioEdit software (160), which enabled the visualisation and editing of the DNA sequences. The aligned sequences, including those retrieved from GenBank representing other strains, were used to generate a phylogenetic tree using MEGA11, a molecular evolutionary genetics analysis tool (161). The phylogenetic tree was constructed using the maximum likelihood method, which allowed for the inference of evolutionary relationships between sequences. The tree was used to identify clusters and relationships between sequences.

Analysis and retrieval of sequences from GenBank

For the identification of closely related sequences, the patient consensus sequences were compared using the NCBI's BLAST tool (162). Sequences which were relevant were retrieved from GenBank (accession numbers: OU342916.1, OU342917.1, MK42217.1, AH013698.2 [Toledo strain], OK000912.1 [SOMA strain] and FJ527563.1) (165). The most suitable

sequences were chosen based on the sequence identity percentage. The closest sequences were analysed for nucleotide differences, strain type and geographic origin.

2.5.4.4 Sequencing and phylogenetic analysis of UL97 gene fragments

Twenty-nine out of 40 samples that demonstrated amplification products via conventional PCR were deemed suitable for further genetic characterisation underwent a similar process of sequencing as the UL54 gene. The resulting amplicons were isolated and subjected to DNA NGS at the Sequencing Core Facility of the NICD, Johannesburg, South Africa to elucidate the precise nucleotide sequence of the UL54 gene. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter) following manufacturer's instructions, followed by tagmentation and addition of index primers using the Illumina Nextera XT library preparation kit (Illumina Inc, USA). DNA libraries were prepared for sequencing on the Illumina NextSeq 500 system using indexed paired-end libraries (Illumina Inc, USA).

Assembly of UL97 sequence reads.

The sequence reads were extracted and assembled using Genome detective virus tool v2.27 (164), and aligned using BioEdit software (160), which enabled the visualisation and editing of the DNA sequences. The aligned sequences, including those retrieved from GenBank representing other strains, were used to generate a phylogenetic tree using MEGA11, a molecular evolutionary genetics analysis tool (161). The phylogenetic tree was constructed using the maximum likelihood method, which allowed for the inference of evolutionary relationships between sequences. The tree was used to identify clusters and relationships between sequences (166).

Analysis and retrieval of sequences from GenBank

The sequencing reads were in FASTQ format. For the identification of sequences which were related closely, the patient consensus sequences were compared using the NCBI's BLAST tool (162). Relevant sequences were retrieved with accession numbers MT044482.1, MT044478.1, KY490063.1, OV100760.1, MT649470.1, MW980585.1, MW439039.1, MT070138.1,

MW197154.1, OU342916.1, OU342916.1, KP745728.1, KU550088.1, OR546128.1, KT726950.2, KY490069.1, KY490075.1, MK290742.1, MK290743.1, MK290744.1, KY490072.1, MK422176.1, KP745713.1, KP745674.1, KP745663.1, KJ872539.1, KJ361971.1, JX512205.1, KP745721.1, KP745644.1 and MH036940.1 (Merlin strain) from GenBank. The most suitable sequences were chosen based on the sequence identity percentage. The closest sequences were analysed for nucleotide differences, strain type and geographic origin.

2.5.4.5 Phylogenetic tree

The phylogenetic trees for the UL54 and UL97 genes were constructed using the Hasegawa-Kishino-Yano (HKY) model to infer evolutionary history, as described by Hasegawa *et al.* (1985). For UL54, the Maximum Likelihood (ML) method was employed to estimate the most likely tree topology. In contrast, UL97 utilized the Maximum Composite Likelihood (MCL) tree drawing method, combining neighbour-join and BioNJ algorithms to identify the optimal tree topology based on pairwise distances and log likelihood scores. To account for rate variation among sites, a discrete gamma distribution with five categories (+G) was applied. Phylogenetic analyses were conducted using MEGA11, leveraging the ML algorithm for UL54 and MCL algorithm for UL97 (161).

2.6 DATA CAPTURE

All data was captured on Microsoft Excel (Microsoft office 365).

2.7 DATA ANALYSIS

Real-time PCR data for both CMV qualitative detection and genotyping was analysed using Applied Biosystems Quantstudio 5 design and analysis software (167). The conventional PCR results were analysed using 1% agarose gel viewed under Syngene Vacutec gel doc system (Cambridge, United Kingdom). Sequence data was assembled and extracted using the online Genome Detective Virus Tool (https://www.genomedetective.com/). Extracted consensus sequences for each sample UL55, UL54 and UL97 genes fragments were aligned by multiple

sequence alignment together with reference sequences obtained from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) using Clustal W application in the BioEdit software program (160). MEGA11 was used to generate phylogenetic tree for evolutionary analysis (161).

Descriptive statistics, including mean, standard deviation, and coefficient of variation, were calculated for the Ct values of positive samples. The proportions of positive samples (prevalence rates) were calculated from the total number of samples, and the 95% confidence interval was determined to estimate the precision of the observed prevalence.

2.8 RELIABILITY AND VALIDITY

Repeated capturing of data on Microsoft Excel, and inclusion of relevant controls were practiced ensuring reliability and validity results. Standard operating procedures, packages inserts, and manufacturer's instructions were also adhered to.

CHAPTER 3

RESULTS

3.1 STUDY POPULATION, DEMOGRAPHICS AND CLINICAL DETAILS

A total of 111 clinical samples were gathered of which 25 (22.5%) were excluded due to either haemolysis or insufficient volume (**Figure 3.1**). The 86 samples retained in the study were from patients seeking care in the private healthcare sector, accessed from Lancet laboratories (87%; 75/86), and the public sector, accessed from the NHLS Virology Laboratory at Charlotte Maxeke Johannesburg Academic Hospital (12.8%, 11/86) (**Table 3.1**).

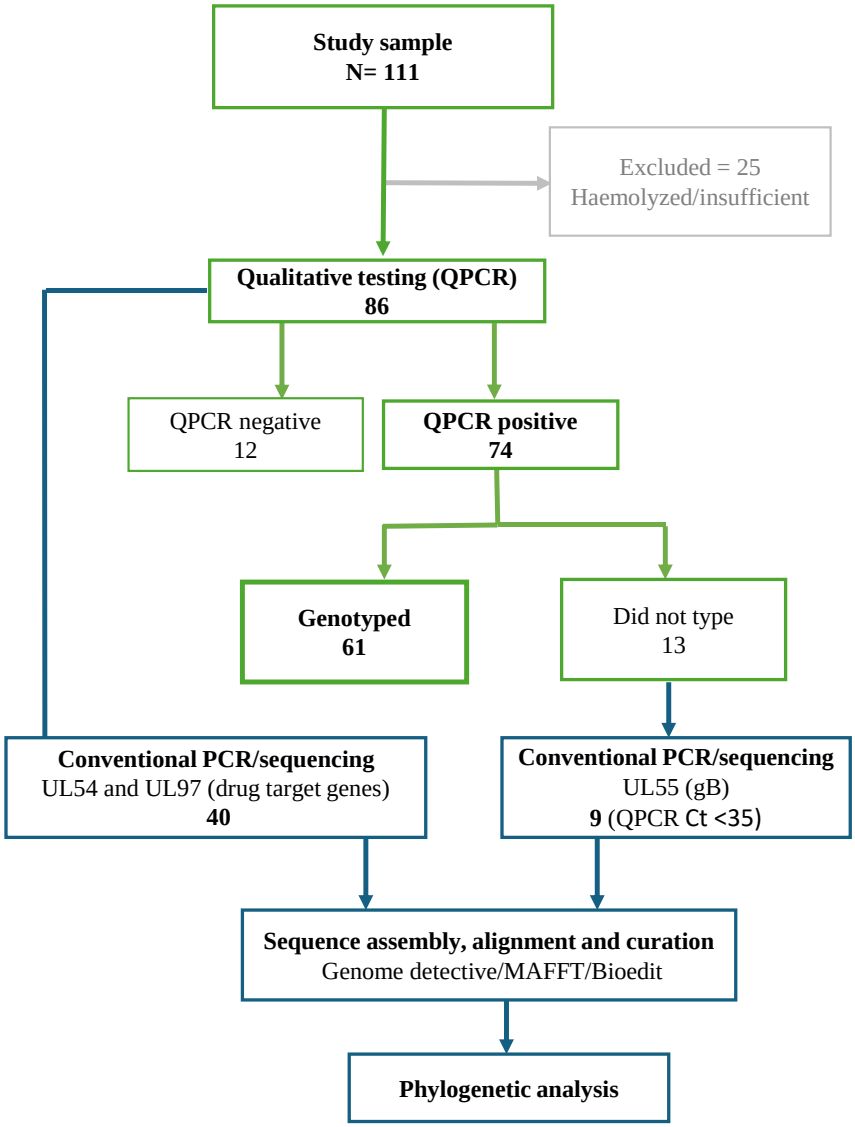

Figure 3.1 Flowchart summarising sample flow for CMV QPCR, genotyping, conventional PCR and sequencing.

Table 3.1. Demographic information for patients with CMV from Lancet Laboratories and NHLS.

No.	Study ID.	Specimen type	Test site	Sex	Age	No.	Study ID.	Specimen type	Test site	Sex	Age
1	CM1	Plasma	NHLS	M	37	**44**	CM68	Blood	Lancet	F	43
2	CM6	Plasma	NHLS	F	44	**45**	CM69	Blood	Lancet	M	36
3	CM7	Plasma	NHLS	M	40	**46**	CM70	Blood	Lancet	M	66
4	CM9	Plasma	NHLS	F	28	**47**	CM71	Blood	Lancet	N/A	N/A
5	CM13	Plasma	NHLS	F	39	**48**	CM72	Blood	Lancet	F	51
6	CM15	Plasma	NHLS	M	36	**49**	CM73	Blood	Lancet	N/A	N/A
7	CM18	Plasma	NHLS	M	6 mo.	**50**	CM74	Blood	Lancet	F	9
8	CM23	Plasma	NHLS	M	48	**51**	CM75	Blood	Lancet	F	66
9	CM25	Plasma	NHLS	F	67	**52**	CM76	Blood	Lancet	M	43
10	CM27	Plasma	NHLS	M	33	**53**	CM77	Blood	Lancet	F	44
11	CM33	Serum	Lancet	N/A	N/A	**54**	CM78	Blood	Lancet	N/A	N/A
12	CM34	Plasma	NHLS	M	48	**55**	CM79	Blood	Lancet	F	49
13	CM35	Serum	Lancet	N/A	N/A	**56**	CM80	Blood	Lancet	N/A	N/A
14	CM36	Serum	Lancet	N/A	N/A	**57**	CM82	Blood	Lancet	F	49
15	CM37	Serum	Lancet	N/A	N/A	**58**	CM83	Blood	Lancet	M	1M,28D
16	CM39	Serum	Lancet	N/A	N/A	**59**	CM84	Blood	Lancet	N/A	N/A
17	CM40	Serum	Lancet	N/A	N/A	**60**	CM86	Blood	Lancet	M	50
18	CM41	Serum	Lancet	N/A	N/A	**61**	CM88	Urine	Lancet	N/A	N/A
19	CM42	Serum	Lancet	N/A	N/A	**62**	CM89	Sputum	Lancet	M	63
20	CM43	Serum	Lancet	N/A	N/A	**63**	CM98	Blood	Lancet	M	59
21	CM44	Serum	Lancet	N/A	N/A	**64**	CM99	Sputum	Lancet	N/A	N/A
22	CM45	Serum	Lancet	N/A	N/A	**65**	CM96	CSF	Lancet	N/A	N/A
23	CM46	Serum	Lancet	N/A	N/A	**66**	CM95	CSF	Lancet	F	28
24	CM47	Blood	Lancet	F	38	**67**	CM90	Sputum	Lancet	N/A	N/A
25	CM48	Blood	Lancet	F	63	**68**	CM94	CSF	Lancet	M	48
26	CM49	Blood	Lancet	M	57	**69**	CM91	Sputum	Lancet	N/A	N/A
27	CM50	Blood	Lancet	F	74	**70**	CM92	Sputum	Lancet	N/A	3M,18D
28	CM51	Blood	Lancet	F	21	**71**	CM97	CSF	Lancet	N/A	1M,26D
29	CM52	Blood	Lancet	F	38	**72**	CM93	CSF	Lancet	N/A	1M 1D
30	CM53	Blood	Lancet	F	3	**73**	CM106	Urine	Lancet	M	51
31	CM54	Blood	Lancet	F	8	**74**	CM100	Sputum	Lancet	M	73
32	CM55	Blood	Lancet	M	49	**75**	CM107	Urine	Lancet	N/A	0M 18D
33	CM56	Blood	Lancet	F	30	**76**	CM102	Urine	Lancet	M	5
34	CM57	Blood	Lancet	F	50	**77**	CM108	Urine	Lancet	N/A	0M 11D
35	CM58	Blood	Lancet	F	38	**78**	CM111	Urine	Lancet	M	42
36	CM59	Blood	Lancet	M	80	**79**	CM101	Urine	Lancet	F	0M 11D
37	CM60	Blood	Lancet	F	25	**80**	CM87	Urine	Lancet	N/A	N/A
38	CM61	Blood	Lancet	M	51	**81**	CM109	Urine	Lancet	N/A	N/A
39	CM63	Blood	Lancet	N/A	N/A	**82**	CM105	Breast milk	Lancet	N/A	N/A
40	CM65	Blood	Lancet	N/A	N/A	**83**	CM120	Blood	Lancet	M	60
41	CM66	Blood	Lancet	N/A	N/A	**84**	CM121	Blood	Lancet	M	53
42	CM67	Blood	Lancet	M	66	**85**	CM122	Blood	Lancet	F	49
43	CM123	Blood	Lancet	F	49	**86**	CM110	Sputum	Lancet	N/A	N/A

Blood= whole blood; N/A: not applicable; NHLS: National Health Laboratory Service

The predominant specimen types were whole blood samples (48%; 41/86), followed by serum samples (14%, 12/86), plasma samples (13%, 11/86), urine (10%, 9/86), sputum/nasopharyngeal aspirates (8%, 7/86), cerebrospinal fluid (6%, 5/86), and breast milk (1%, 1/86). For most of the specimens limited demographic data could be retrieved for 57/86 (66%). Among the 57 patients age data on sex was available for 52 (91%) with equal distribution between males (26/52; 50%) and females (26; 50%) (**Table 3.1**). Age was indicated for all 57 patients with the youngest according to the data available an 8-day old male from the public sector and the oldest an 80-year-old male from the private sector. The available data indicated that samples included patients of all ages and races. Detailed clinical data was unavailable.

3.2 QUALITATIVE SCREENING OF CMV BY REAL-TIME PCR

3.2.1 Qualitative real-time PCR results of CMV

Among the 86/111 (77.5%) specimens evaluated using the qualitative PCR (QPCR) assay, 86% (74/86) yielded positive results, whilst 14% (12/86) tested negative (**Figure 3.1**). The positive samples exhibited a range of Ct values, from 26 to 38, with a median value of 34 (**Table 3.2**). Notably, all 11 plasma samples tested positive for CMV. In contrast, various other specimen types showed lower positivity rates: 2 out of 12 (16.7%) serum samples, 4 out of 41 (9.8%) whole blood samples, 1 out of 9 (11%) urine samples, 3 out of 7 (42.9%) sputum/nasopharyngeal aspirate samples, and 2 out of 5 (40%) cerebrospinal fluid samples tested negative for CMV DNA.

Further analysis of the Ct values revealed a mean of 33.55, a standard deviation of 4.33, and a coefficient of variation of 12.95%. Additionally, the 95% confidence interval for the proportion of positive samples was estimated to be between 32.32 and 34.78, with a margin of error of approximately 1.20.

Table 3.2. CMV qualitative PCR results

No.	Sample ID.	QPCR	Ct value	No.	Sample ID.	QPCR	Ct value	No.	Sample ID.	QPCR	Ct value
1	CM1	+	38	**30**	CM53	+	36	**59**	CM84	+	29
2	CM6	+	33	**31**	CM54	+	32	**60**	CM86	+	37
3	CM7	+	28	**32**	CM55	-	-	**61**	CM88	+	34
4	CM9	+	35	**33**	CM56	+	38	**62**	CM89	+	28
5	CM13	+	33	**34**	CM57	+	38	**63**	CM98	+	32
6	CM15	+	37	**35**	CM58	+	36	**64**	CM99	+	27
7	CM18	+	35	**36**	CM59	+	28	**65**	CM96	+	26
8	CM23	+	37	**37**	CM60	+	37	**66**	CM95	-	-
9	CM25	+	35	**38**	CM61	+	31	**67**	CM90	+	26
10	CM27	+	32	**39**	CM63	+	35	**68**	CM94	+	34
11	CM33	+	35	**40**	CM65	+	37	**69**	CM91	-	-
12	CM34	+	37	**41**	CM66	+	34	**70**	CM92	-	-
13	CM35	+	28	**42**	CM67	+	33	**71**	CM97	+	28
14	CM36	+	38	**43**	CM123	+	27	**72**	CM93	-	-
15	CM37	+	28	**44**	CM68	+	38	**73**	CM106	+	34
16	CM39	+	38	**45**	CM69	+	35	**74**	CM100	+	38
17	CM40	+	38	**46**	CM70	+	37	**75**	CM107	-	-
18	CM41	-	-	**47**	CM71	+	32	**76**	CM102	+	37
19	CM42	+	37	**48**	CM72	+	36	**77**	CM108	+	29
20	CM43	+	34	**49**	CM73	+	38	**78**	CM111	+	31
21	CM44	-	-	**50**	CM74	+	34	**79**	CM101	+	34
22	CM45	+	26	**51**	CM75	+	32	**80**	CM87	+	34
23	CM46	+	28	**52**	CM76	+	31	**81**	CM109	+	34
24	CM47	+	34	**53**	CM77	+	36	**82**	CM105	+	35
25	CM48	+	37	**54**	CM78	+	36	**83**	CM120	+	33
26	CM49	+	38	**55**	CM79	-	-	**84**	CM121	+	34
27	CM50	+	34	**56**	CM80	-	-	**85**	CM122	+	30
28	CM51	-	-	**57**	CM82	+	26	**86**	CM110	-	-
29	CM52	+	30	**58**	CM83	+	35				

3.3 CMV GENOTYPING

3.3.1 CMV genotyping by real-time PCR

The 74/86 (86%) QPCR positive samples were processed for genotyping by real-time multiplex PCR and genotypes could be identified for 61/74 (82%) samples. Amplification curve examples are shown in **Figure 3.2**. The plots display the fluorescence intensity (ΔRn) against the cycle number, with each curve representing a specific CMV genotype (gB1, gB2, gB3, and gB4).

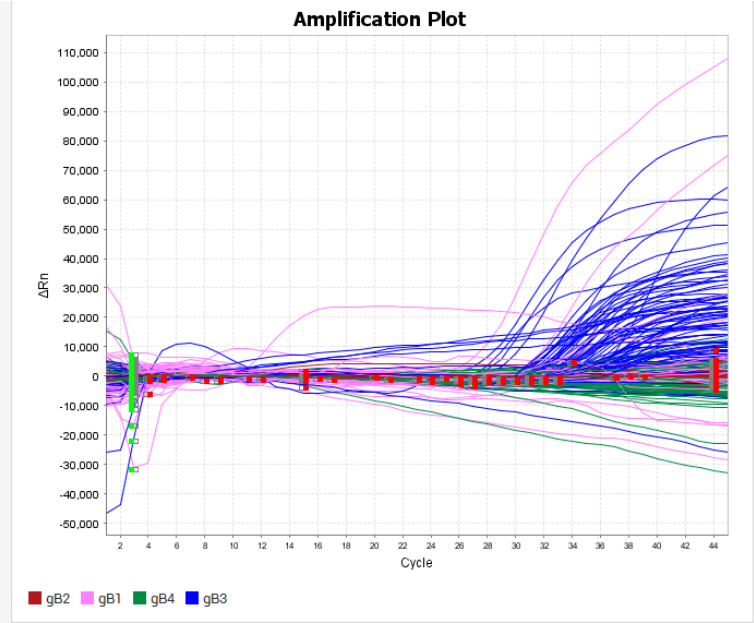

Figure 3.2. Real-time PCR amplification plots for CMV genotype determination. Amplification plots represent the amplification curves for genotypes gB1, gB2, gB3, and gB4. The X- axis shows the cycle numbers.

Genotype gB3 was detected in 77% (57/74) of the samples. Three samples (3/74; 4%) had dual detection of gB3 and gB4 strains and two samples (2/74; 2.7%) had dual detection of gB1 and gB3 strains (**Figure 3.3**). Genotype gB2 was not detected in this study population. Genotyping was unsuccessful for 13 samples (**Table 3.2**). Mixed infections were also detected with the QPCR genotyping assay at a prevalence of 6.7% (5/74) representing gB1/gB3 in 2 patients (2.7%) and gB3/gB4 in 3 patients (4%) (**Figure 3.3**).

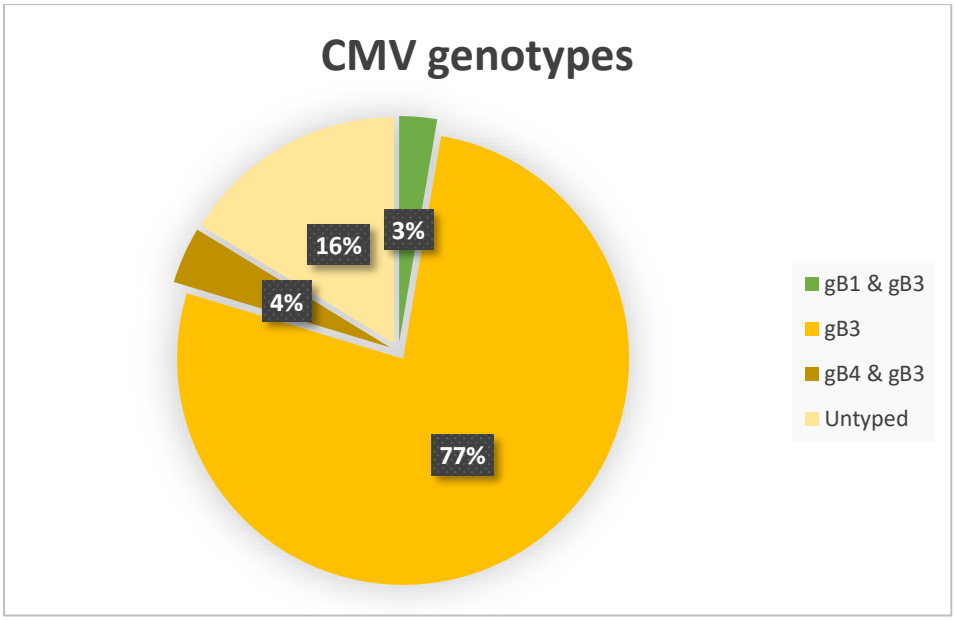

Figure 3.3. Pie chart illustrating the frequency of CMV genotypes identified in the study population (n = 74).

Table 3.3. List of 13 specimens that were positive for CMV but could not be classified into any of the four expected genotypes using the real-time PCR assay.

No.	Sample ID.	Qualitative PCR results	Ct value
1	CM13	+	33
2	CM25	+	35
3	CM47	+	34
4	CM56	+	38
5	CM68	+	38
6	CM71	+	32
7	CM75	+	32
8	CM76	+	31
9	CM78	+	36
10	CM83	+	35
11	CM98	+	32
12	CM97	+	28
13	CM100	+	38

These samples showed clear amplification curves for more than one genotype, indicating co-infection rather than cross-reactivity. The Ct values for these samples were within the same

range as single-genotype samples, suggesting adequate sample quality. None of the untyped samples overlapped with the mixed-genotype group.

3.3.2 Conventional PCR for CMV UL55 gene

Among the 13 samples that tested positive for CMV but failed the real-time PCR genotyping assay, a subset of 9 samples exhibiting Ct values of 35 or lower were selected for further analysis using conventional nested PCR. This approach was intended to enhance the sensitivity of detection and facilitate genotyping. The gene of interest was UL55, spanning 2709 base pairs in length. Successful amplification of the UL55 gene was achieved in 6 out of 9 samples, which were subsequently subjected to DNA sequencing to determine their genotypic profile. **Table 3.4** presents the results of the conventional nested PCR amplification of the UL55 gene, highlighting the samples that yielded successful amplification. The amplified products were subsequently resolved on a 1% agarose gel.

Table 3.4. Summary of nested conventional PCR amplification of the UL55 gene from CMV positive untyped samples.

No.	Sample ID.	QPCR Ct value	Conventional UL55 PCR results
6	CM13	33	-
10	CM25	35	+
25	CM47	34	-
47	CM71	32	+
51	CM75	32	+
52	CM76	31	+
58	CM83	35	-
63	CM97	28	+
71	CM98	32	+

3.3.4 DNA sequencing and phylogenetic analysis of UL55 genes

Of the seven biological samples subjected to sequencing analysis, a single sample failed to yield amplifiable DNA, resulting in a success rate of 85.7% (6/7) for generating usable sequencing data (**Table 3.5**). The sequencing data for sample CM25 revealed a total of 602 high-quality reads, achieving an average depth of coverage of 158.7-fold. The genomic coverage was calculated to be 0.9%. Furthermore, the nucleotide identity (NT identity) of the sample's sequence was determined to be 81.9% similar to the CMV reference strain (**Table 3.5**). In **Figure 3.4** examples of the sequence identity output are provided in the context of the complete CMV genome. The sequence segments obtained from CM25 were outside of the UL55 gene and were located at the position of approximately 10 000 bp which is outside of the targeted gene (**Figure 3.4F**). This indicates that the assembled UL55 gene sequences obtained located to the expected genomic position of approximately 80 000 bp for CM71, CM75, CM97, CM98 and CM76 (**Figure 3.4A** to **3.5E**, respectively).

Table 3.5 Genome coverage and identity metrics for each sample, enabling an assessment of the sequencing data's quality and accuracy.

Sample	#Reads	Depth of coverage	Genome coverage	NT identity	Assembled sequence length in base pairs (bp)
CM25	602	158.7	0.9%	81.9%	908 bp
CM71	1592122	68717.0	1.2%	97.7%	2935 bp
CM75	2269975	97942.5	1.2%	99.5%	2938 bp
CM76	1862544	79466.2	1.3%	93.4%	2943 bp
CM97	1867753	80067.8	1.2%	97.8%	2927 bp
CM98	3097874	126776.0	1.2%	95.0%	2 938bp

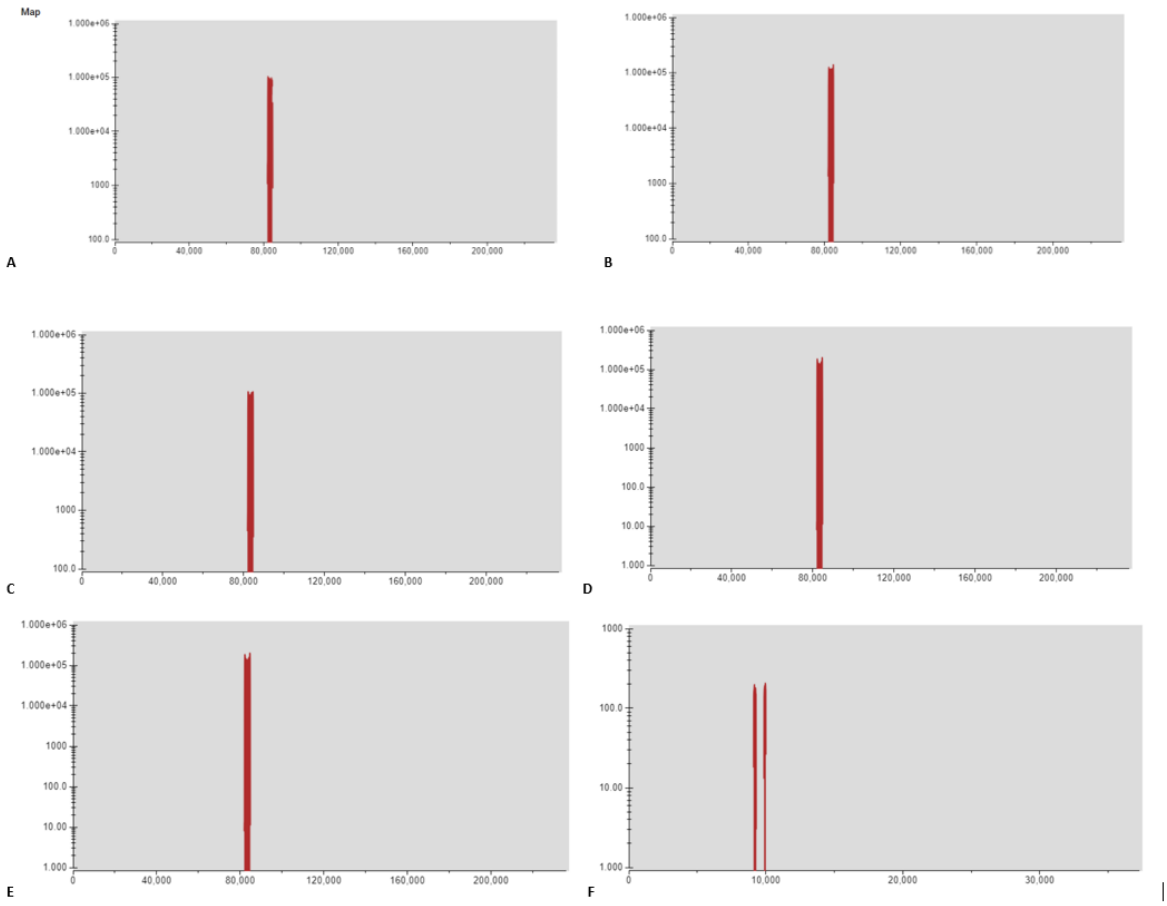

Figure 3.4 Genome coverage distribution for UL55 sequencing. This image offers a visualisation of the distribution of genome coverage for the sequenced samples: A: CM71, B: CM75, C: CM97, D: CM98, E: CM76, F: CM25). The Y-axis indicates the sequence read coverage or number of reads aligning to each region and the X-axis (in base pairs) indicates the positions across the CMV genome. The genome regions where there is significant coverage are highlighted in red.

Phylogenetic analysis of the UL55 gene segment showed that CM98 clusters with gB2 strains with 99% bootstrap support (**Figure 3.5**). CM71 and CM97 cluster together with 87% bootstrap support outside of the gB2 group. The gB1 group is supported by 75% bootstrap value and only CM75 clusters in this group. CM76 clustered in the gB3 group supported by bootstrap value of 100%. No gB4 strains were identified.

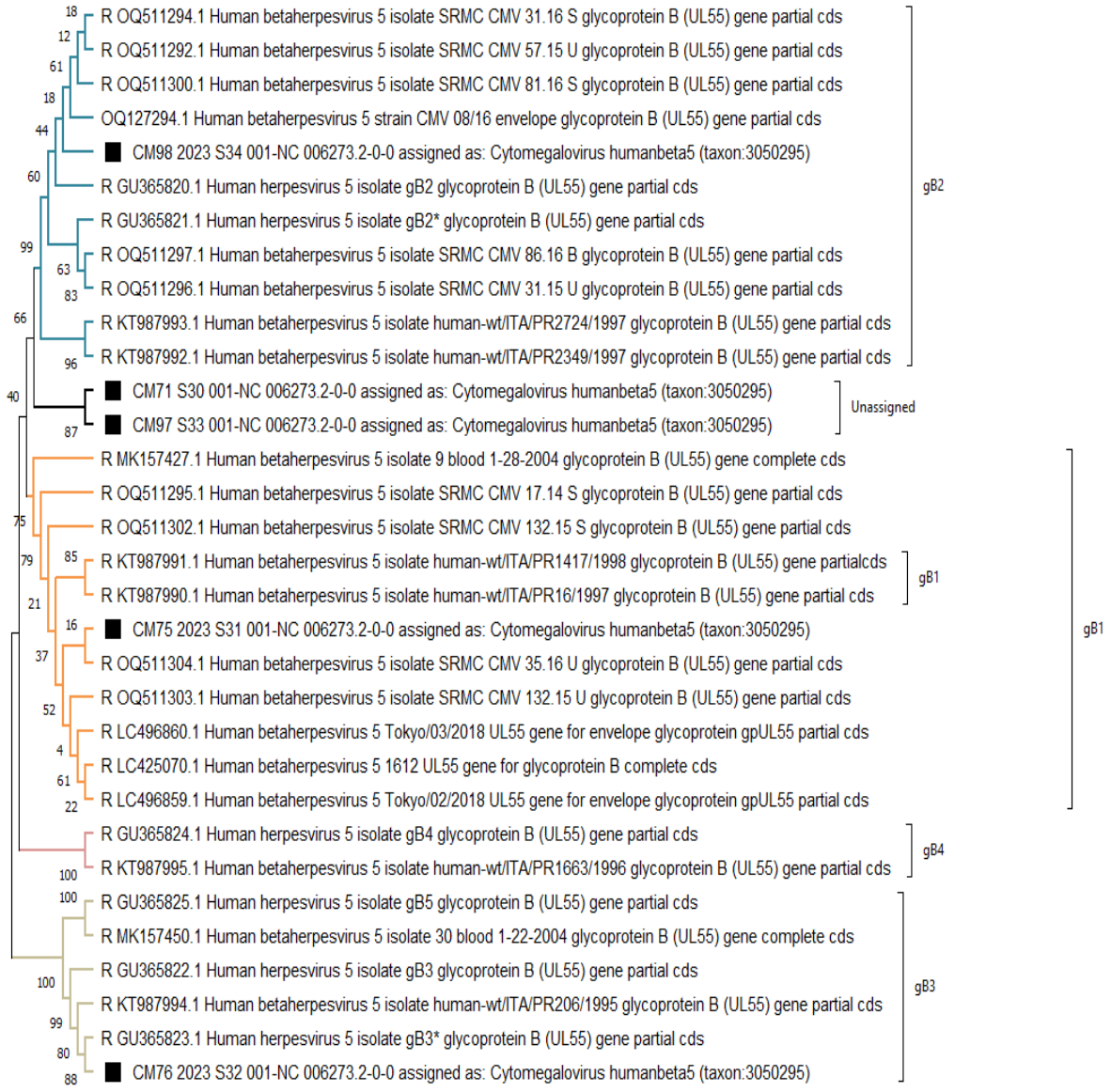

Figure 3.5. Phylogenetic analysis of UL55 gene fragments among CMV isolates. The tree was constructed using the maximum likelihood method with 500 bootstrap replicates under the Hasegawa-Kishino-Yano model. Branch lengths represent evolutionary distances. The multiple sequence alignment was trimmed to 280 bp in order to include partial coding sequences for UL55 genotype (gB1–gB4) references.

Subsequently a gB5 reference strain were identified, and the phylogenetic analysis (**Figure 3.6**) was repeated. However, despite the inclusion of the gB5 reference strain, the two samples CM71 and CM97 remain genotype unassigned as in **Figure 3.6**, but with ancestral links to gB1 and gB4 forming a cluster with them with 72% bootstrap support.

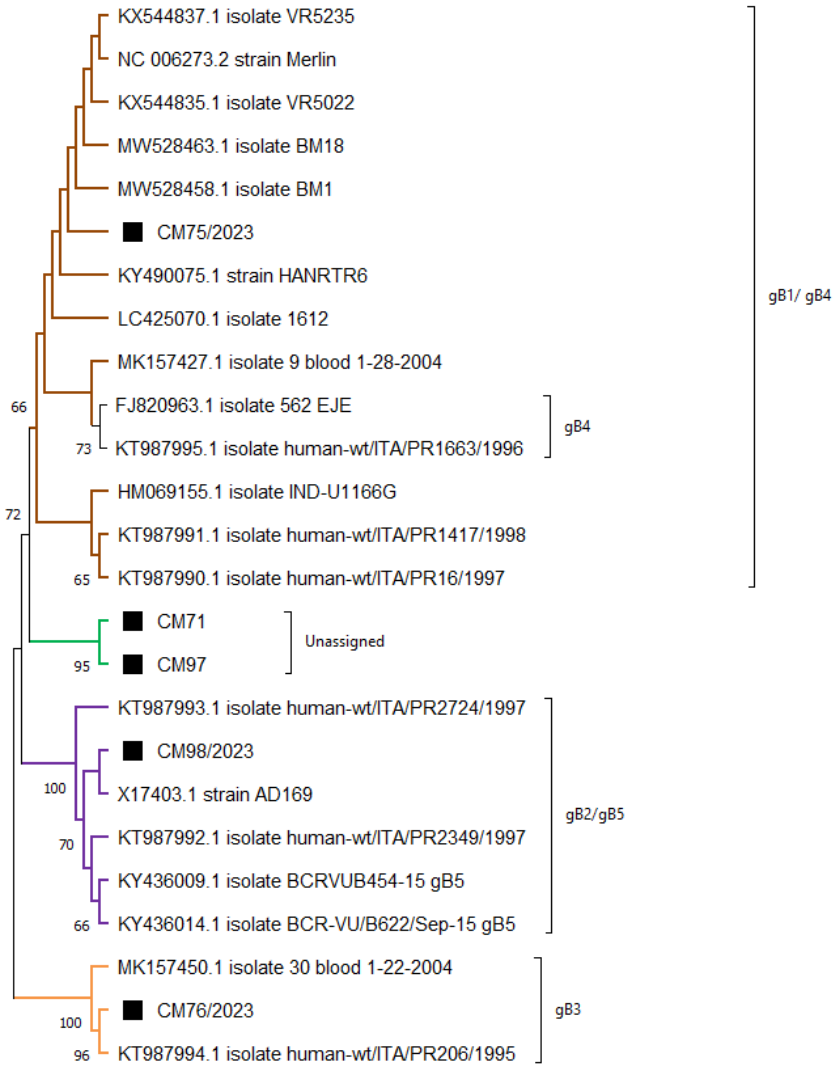

Figure 3.6 Maximum-likelihood phylogenetic analysis of the UL55 gene fragment among CMV isolates, incorporating the reported gB5 strain (Human betaherpesvirus 5 isolate BCR-VU). The tree was constructed using 500 bootstrap replicates, with confidence values indicated at branch nodes. The multiple sequence alignment was trimmed to 279 bp in order to include partial coding sequences for UL55 genotype (gB1–gB4) references as well as the gB5 strain. Branch lengths represent evolutionary distances, and the Hasegawa-Kishino-Yano model was applied to account for nucleotide substitutions.

3.3.5 Summary of gB genotypes assigned by real-time PCR and sequencing

In summary, a total of 74/86 (86%) CMV-positive samples were subjected to genotyping by real-time PCR targeting the UL55 gene. Genotypes were successfully assigned to 61/74 (82.4%) samples. The 9/13 (Ct <35) samples that could not be genotyped by real-time PCR were evaluated by conventional PCR and of these a genotype was successfully assigned to 3/9 (33.3%) by sequencing and phylogenetic analysis increasing the overall genotyping success rate to 86% (64/74) (**Figure 3.7**). However, 2/9 (22.2%) remained "unassigned" to a specific genotype, but clustering with gB1 and gB4 with a bootstrap value of 72%. Collectively, the results reveal (**Figure 3.7**) a predominantly gB3 population (58/74, 78.4%), with minor representation of gB2 (1/74, 1.4%) and gB1 (1/74, 1.4%). Notably, 3 samples (3/74, 4.1%) exhibited dual detection of gB3 and gB4 strains, and two samples (2/74, 2.7%) showed dual detection of gB1 and gB3 strains, highlighting the presence of mixed infections.

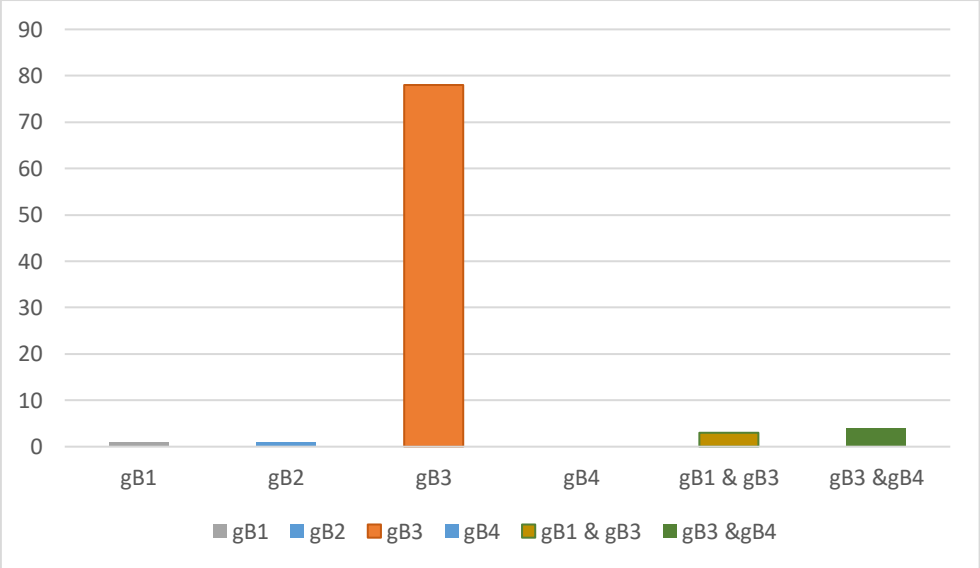

Figure 3.7 CMV genotype distribution in single and mixed infections among persons of all ages, 2022-2023. The X-axis indicates the CMV genotypes, and the Y-axis represent the percentage of each genotype detected.

3.4 CMV DRUG RESISTANCE TESTING

3.4.1 Amplification and sequencing of UL54

A subset of 40 samples was subjected to PCR amplification for further analysis, and only 10 of them were successfully amplified. Despite visible bands on the gel, their sizes do not correlate with the expected amplicon size of 1654bp. The observed band size was 1000bp which are visible together with multiple smaller band sizes (**Figure 3.8**). Optimization of the PCR was attempted by the adjustment of the PCR conditions and the use of different PCR kits. Due to time and financial limitations, further optimization could not be possible. The 10 successful samples were taken for further analysis by sequencing.

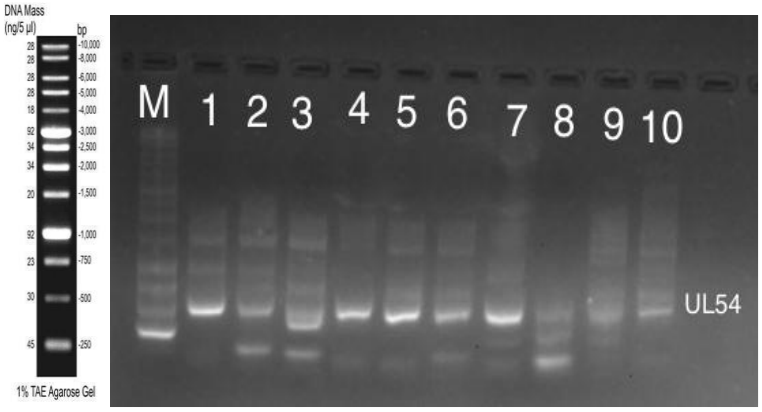

Figure 3.8 Gel electrophoresis image of 10 samples that were successfully amplified as seen by the presence of bands. A molecular marker (M) is used to identify and verify the size of the specific gene (UL54).

Sequence assembly and coverage for UL54

Genome Detective was used for the assembly of the 10 samples. Six samples showed the presence of CMV after assembly. Notably, the assembly report also revealed the presence of other viruses, with HIV being the most prevalent co-infection. However, further analysis and discussion of these findings were limited by the absence of confirmatory ELISA/PCR results. The four samples excluded from further analysis showed no CMV detection, had significant size discrepancies, and poor alignment quality compared to the other sequenced samples.

According to the genome graphics in **Figure 3.9**, the UL54 gene is located between nucleotide positions 78,000 and 82,000. The sequences from this study samples all had significant coverage at before 80,000. For NCV1936, NCV1939, NCV1942 and NCV1943 significant coverage were also observed outside the region of interest.

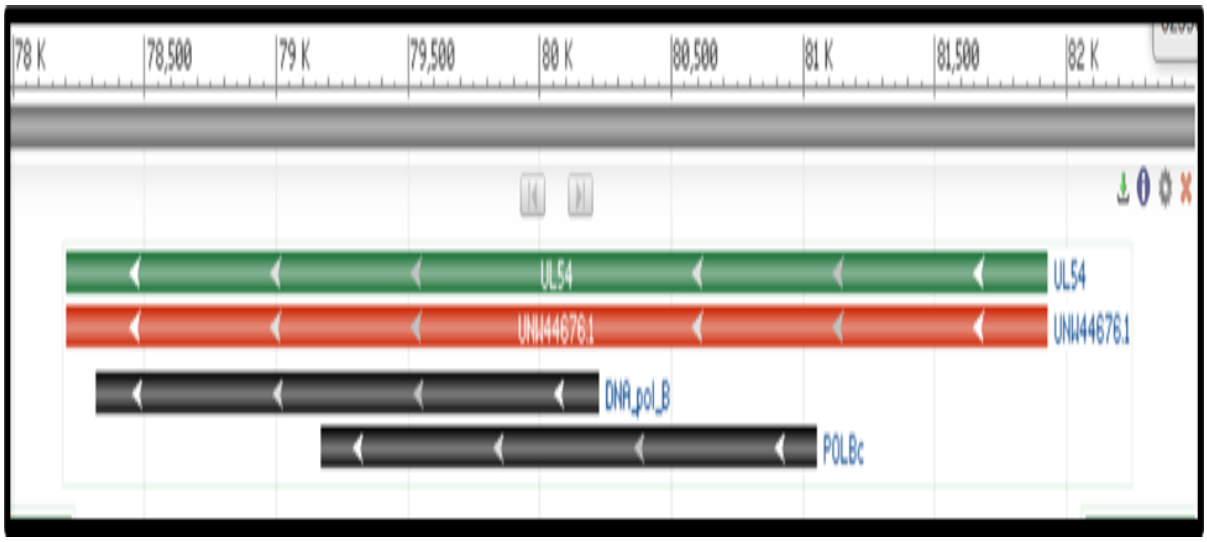

Figure 3.9 Screenshot taken from GenBank showing the position of the UL54 gene on the CMV virus graphics.

Notably, all sequences showed low percentage coverage, with NCV1938 and NCV1940 showing the lowest coverage of approximately 0.3% and a coverage length of around 700 base pairs (1 contig) (**Table 3.6** and **Figure 3.10**). In contrast, NCV1943 stood out with the highest percentage of coverage 0.8% and a longest length coverage of 1856 base pairs, spread across 7 contigs (**Table 3.6** and **Figure 3.10**). Notably, the sequence concordance was remarkably high, ranging from 94.2% to 99.1% (**Table 3.6**), indicating a strong similarity between the patient sequences and the Merlin reference strain used by the software during assembly.

In summary, this analysis yielded partial CMV sequences that were very similar to the reference CMV strain, but unfortunately, not relating to the gene of interest. To better understand the results, visual map was created that shows where the detected nucleotides would fit on the reference CMV Merlin strain.

Table 3.6: A comparative analysis of the quality of the UL54 sequence assemblies to the reference Merlin strain generated with Genome Detective.

Sequence ID	Length coverage	% Coverage	% Concordance	Matches (%)	Identities (%)
NCV1936	1068 (2 contigs)	0.5	98.6	1068 (99.9%)	1063 (99.4%)
NCV1938	781 (1 contig)	0.3	94.2	789 (100%)	766 (97.1%)
NCV1939	1082 (2 contigs)	0.5	98.2	1082 (100%)	1072 (99.1%)
NCV1940	739 (1 contig)	0.3	98.2	739 (99.9%)	735 (99.3%)
NCV1942	1072 (2 contigs)	0.5	99.1	1072 (100%)	1067 (99.5%)
NCV1943	1856 (7 contigs)	0.8	98.7	1856 (100%)	1844 (99.4%)

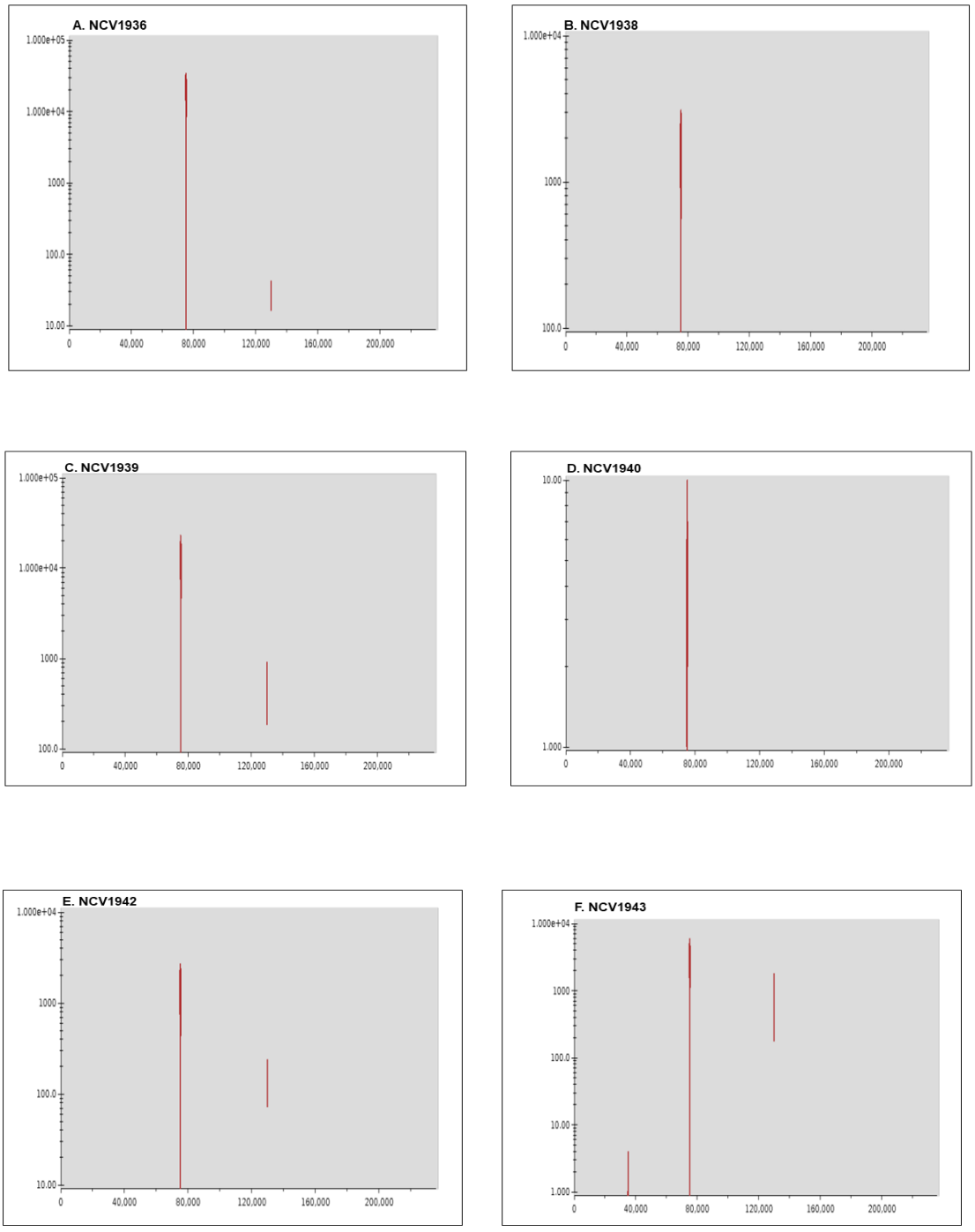

Figure 3.10 Genome coverage maps. These visual maps provide a detailed overview of the location of genetic data for UL54 gene sequences. The Y-axis indicates the sequence read coverage or number of reads aligning to each region and the X-axis indicates the positions across the CMV genome. The genome regions where there is significant coverage are highlighted in red.

Phylogenetic analysis UL54

The tree in **Figure 3.11** was generated to visualise the evolutionary relationship among the strains of CMV based on their UL54 gene sequences. The analysis was achieved through the use of the Maximum Likelihood technique with the Hasegawa-Kishino-Yano model. The branch lengths represent the evolutionary distance between the strains, the longer the branch the greater the genetic divergence. Due to its representation of a standard consensus sequence, the AD169 strain was selected as the reference point, as it is commonly used for comparative analysis of drug-resistant HCMV strains.

Two significant branches were identified. The Zambian strain (MK4222176.1) served as the anchor for the first branch, with the prototype strains NCV1943 and NCV1942 clustering closely around it (**Figure 3.11**). The Toledo strain (AH013698.2) served as the hub for three subgroups, comprising NCV1940, NCV1939, and NCV1936, which clustered around it (**Figure 3.11**). A clear distance was seen between the Zambian and Toledo strains and the other patient sequences, which formed two distinct clusters. This divergence was reinforced by a bootstrap value of 63%, providing moderate support for this grouping.

The other primary branch comprised the other South African strains OU342916.1 and OU342917.1 (**Figure 3.11**), which formed a tight cluster with the Merlin and Soma reference strains, supported by a bootstrap value of 66%. This grouping suggests a close evolutionary relationship between these strains.

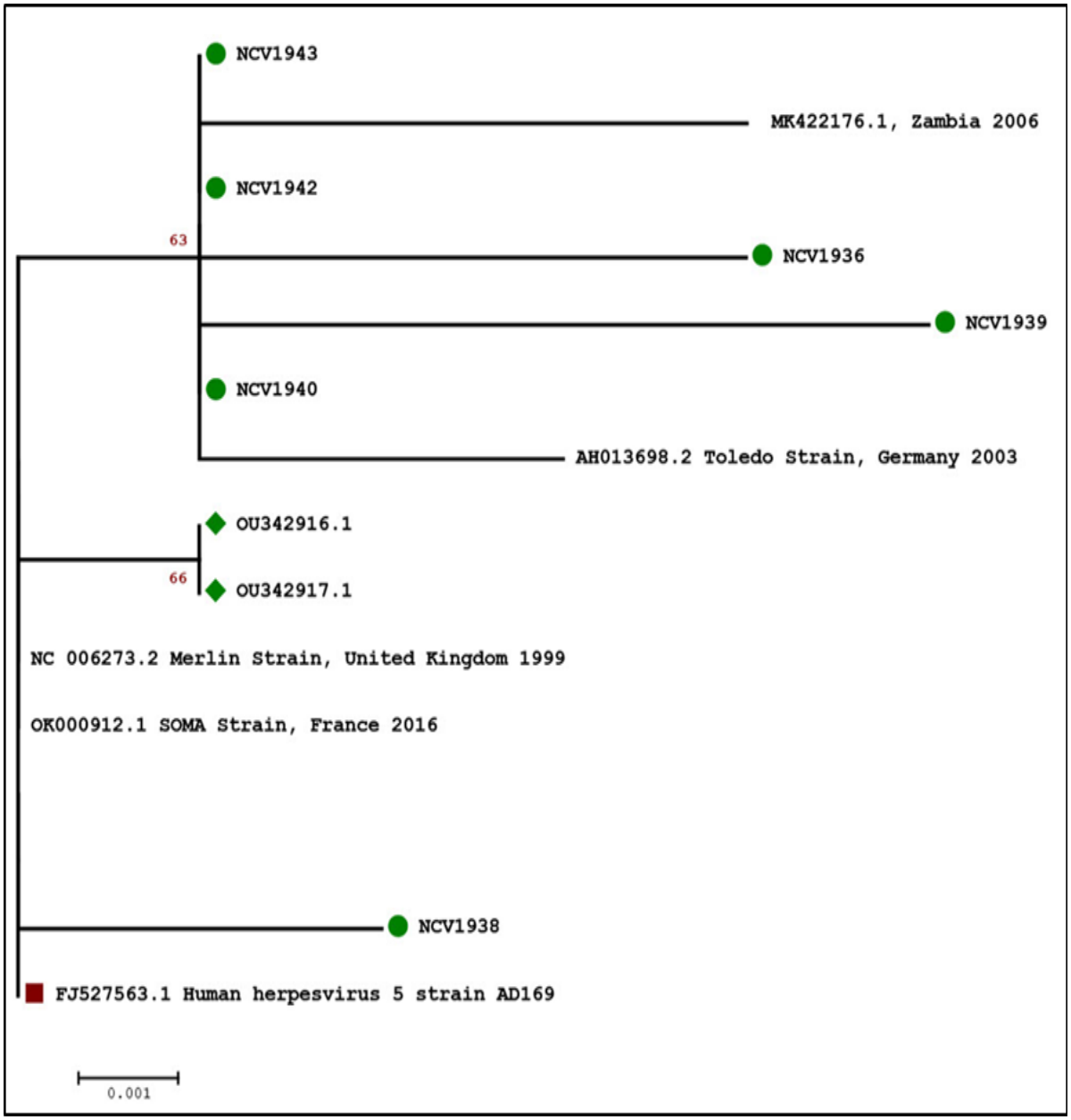

Figure 3.11 The Maximum Likelihood technique based on the Hasegawa-Kishino-Yano model was used to infer the evolutionary history of the Human Cytomegalovirus UL54 gene. Next to the branches is the proportion of trees in which the related taxa clustered together. The tree is depicted to scale, and branch lengths are measured in substitutions per site. The study included 13 nucleotide sequences. The 6 sequences in the study were represented by a green circle, and the green diamond presents other South African strains that were previously published. All positions with gaps and missing data were removed. MEGA11 was used to perform the evolutionary analysis.

3.4.2 Amplification and sequencing of UL97

Sequencing was performed on 29/40 samples, yielding assembled sequence reads. However, only 6/29 samples yielded identifiable CMV contigs (NCV1906, NCV1908, NCV1912, NCV1916, and NCV1919) the read length was ranging from 81bp to 825bp (**Table 3.7**). **Figure 3.12** provides an illustration of the genomic location of the identified CMV gene fragments withing the complete CMV genome. BLASTN analysis revealed most fragments (NC1906, NCV1907, and NCV1912) to be mapped to the UL98 gene, upstream of UL97. On the contrary, the NCV1919 gene segment was located in a genomic region encompassing the UL54 and the UL56 genes.

Table 3.7 Summary of Genome Detective assembly statistics for UL97 sequence reads.

Study ID	Read length	Genome Coverage	Matches (%)	Identities (%)
NCV1906	81	0.1%	81 (100%)	80 (98.8%)
NCV1907	825	0.3%	665 (100%)	656 (98.6%)
NCV1908	790	0.1%	285 (100%)	282 (98.9%)
NCV1912	135	0.1%	135 (100%)	132 (97.8%)
NCV1915	693	0.2%	494 (100%)	490 (99.2%)
NCV1919	150	0.1%	150 (100%)	135 (90.0%)

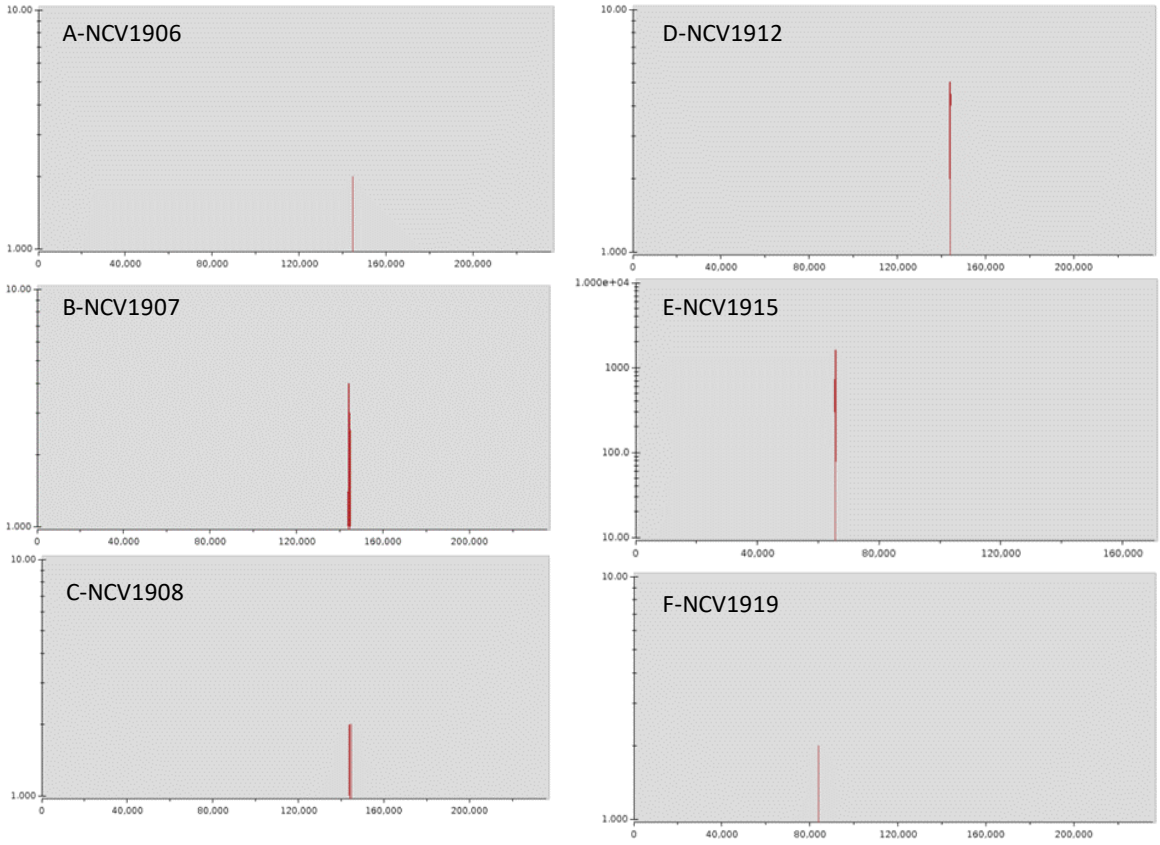

Figure 3.12 Coverage maps of consensus contigs for UL97 obtained with Genome Detective. The Y-axis represents the depth of coverage, and the X-axis represents the position of the sequence in the context of the complete genome.

Phylogenetic analysis UL97

A Maximum Likelihood phylogenetic tree was constructed, incorporating a rate variation model (5 categories + G, parameter = 0.0500), which revealed that nearly half of the sites (49.27%) were evolutionarily invariable. The resulting tree, drawn to scale with branch lengths representing substitutions per site, showed that two study samples (NCV1915 and NCV1907) clustered with globally circulating cytomegalovirus (CMV) strains. Notably, NCV1907 formed a well-supported subcluster (54% bootstrap) with CMV strains identified in Belgium in 2010 supported by a 54% bootstrap value (**Figure 3.13**). NCV1915 clustered with a diverse group of CMV strains from various geographic locations, including Durban (South Africa), Zambia

(2002-2003), Belgium (2010-2011), and London (United Kingdom, 2011), indicating widespread genetic relatedness.

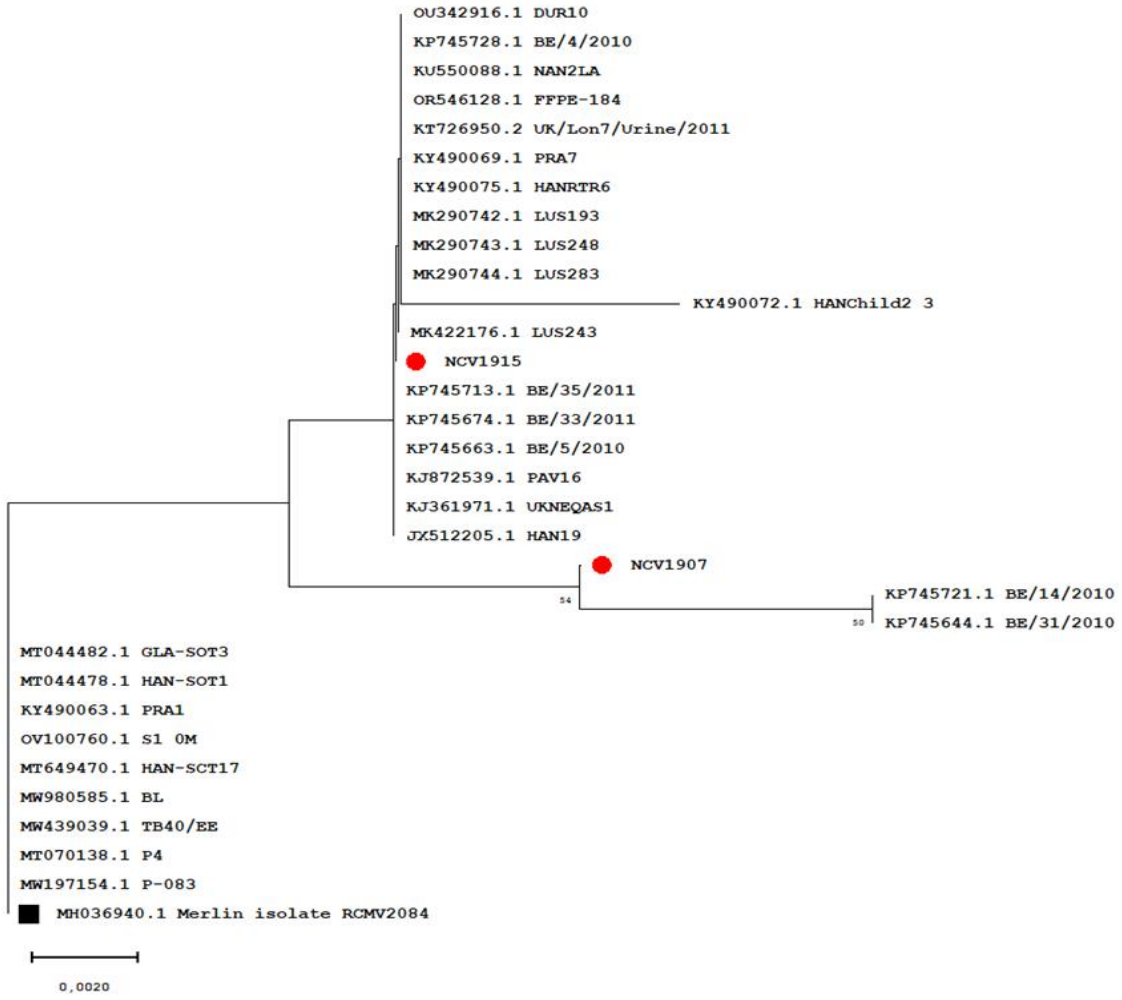

Figure 3.13 A maximum likelihood phylogenetic tree was constructed for the UL97 gene of Human HCMV, with the Merlin strain (MH036940.1) serving as the root. Bootstrap values above 50 are shown. Genes identified in this study are denoted by red dots, while the root is marked by a black square.

CHAPTER 4

DISCUSSION AND CONCLUSIONS

4.1 DISCUSSION

CMV is a ubiquitous virus that is clinically significant, posing a substantial global health concern with seroprevalence nearing 100% in developing countries, and approximately 40% in developed countries, vulnerable individuals to CMV disease being the immunocompromised and neonates (21, 25). Despite its significance, there is paucity of data on CMV epidemiology and resistance patterns in Africa. In this study among 74/86 (86%) CMV positive samples the gB3 genotype was predominantly prevalent at 77% (57/74). This dominance, coupled with the absence of mono infections with gB4 and single cases of gB1 and gB2 identified by sequencing, suggests distinct local epidemiological patterns possibly influenced by specific transmission dynamics in South Africa. Findings on genotype prevalence in this study is in contrast to a study conducted in Pretoria in 2014 where among immunocompromised patients gB2 was detected in 35% (7/20) followed by gB4 in 30% (6/20) of patients as the most prevalent genotypes (168). This study also identified genotypes gB1 and gB3, despite the small population size. Another South African study with a limited sample of 12 HIV exposed infants with CMV pneumonia from Durban KwaZulu Natal Province, also in contrast to this study identified gB2 (8/12; 67%) as the most prevalent with GB5 as the only other genotype detected (169). Unfortunately, in this study GB5 was not included in the genotyping assay and sequencing could not differentiate the genotypes appropriately due to the short sequence length (<500 bp) analysed during phylogenetic analysis. These observations indicate that differences geographic distribution patterns of CMV genotypes exist in South Africa. Overall, genotyping by real-time multiplex PCR revealed a high success rate of 82.4% (61/74).

Furthermore, there was dual detection of gB1 and gB3 in 2.7% (2/74) and gB3 and gB4 in 4% (3/74) by QPCR-based genotyping. The dual detection observed indicates that infections by multiple strains of CMV in our study population. This finding is similar to a study by Külekci *et al.* where 5 out of 10 donor lungs were CMV-positive, with three of those cases involving multiple CMV strains (170). Previous studies have with consistency demonstrated that mixed infections with multiple strains are prevalent, specifically among HSCT patients, and the presence of multiple strains is influenced by factors such as viral load and the specimen type

analysed (171, 172). The detection of mixed genotype infections at an overall prevalence of 6.7% (5/74) in this study is lower than reported in the literature. For example, Tarragó *et al.* reported a 27.5% rate of mixed gB1/gB3 infections in serum samples (173), while Handous *et al.* found mixed gB genotypes in 31% of non-HSCT haematological malignancy patients, often associated with higher viral loads and symptomatic disease (6). Similarly, in solid organ transplant recipients, mixed gB infections were detected in ~17% of samples, frequently involving gB3, whereas Vasiljević *et al.* observed an even higher rate of 55.5% in HSCT patients (157, 172). In congenital CMV, Sarkar *et al.* identified mixed infections in 23.5% of symptomatic neonates, most often involving gB1 with gB2 or gB3 (174). Notably, the gB1+gB3 and gB3+gB4 pairs we observed are consistent with genotype combinations reported in other settings (173, 174). These findings emphasise that although mixed infections appear relatively uncommon in this population, they remain clinically relevant and may contribute to within-host diversity, viral persistence, and treatment challenges. Interestingly, no mono-infections with gB1, gB2, or gB4 were detected in this study. It remains possible that these genotypes are circulating at low prevalence but were not captured. Alternatively, the predominance of gB3 may suggest regional clustering or selective advantages for this genotype in the South African context, which warrants further investigation.

The negative sample types included urine, sputum/NPA, CSF, serum and whole blood samples. DNA degradation resulting from delayed processing and temperature fluctuations during transportation from collaborating laboratories to our research facilities is the most likely contributing factor to negative results. The statistical analysis of the CMV qualitative PCR Ct values revealed a Ct mean of 33.55 which was suggestive of a moderate level of CMV DNA in the samples. The standard deviation of 4.33 and coefficient of variation of 12.95% was suggestive of a significant amount of variability in viral load amongst the samples. This could be as a result of factors such as differences in the immune responses of different individuals, the stage of the disease, sample type and possibly site of sample collection. Overall, this analysis is suggestive of the fact that there is a moderate level of CMV infection in the study population, with significant variability in viral load.

Samples that could not be genotyped (18%; 13/74) underwent further analysis using conventional PCR and sequencing, to ensure comprehensive characterisation of CMV diversity in this cohort. Among the 13 samples that were CMV-positive but were not genotyped, a subset

of 9 samples which had Ct values less than 35 were selected to undergo further investigation, and a successful amplification of UL55 gene sequences was achieved in 5 out of 9 samples. Sequencing of the 5 samples yielded a successful genotyping for only 3, of which the phylogenetic analysis revealed them to be clustered in three genotypes: gB1, gB2, and gB3. Despite further investigation, including the incorporation of gB5 in the phylogenetic analysis, the remaining 2 samples could not be clearly assigned a genotype. The inability to assign a genotype for these 2 samples is due to the limited sequence length used during phylogenetic analysis. Unfortunately, we had to trim our study sequences to the limited sequence lengths of all the relevant genotype specific reference sequences included in the multiple sequence alignments. Of the remaining samples there was a sample that did not yield the expected sequencing output, most likely due to degraded DNA that may have hindered accurate sequencing. The successful genotyping of samples by conventional PCR that initially failed real-time PCR suggest primer-probe mismatches as the greatest limitation. The real-time PCR primers and probes may not have been suitably matched for the genetic diversity present in the study samples. This emphasises the significance of primer sequences optimisation for accuracy in CMV genotyping. Additionally, due to the sensitivity of real-time PCR, even small mismatches in the primers and probes can make a difference, whereas in conventional PCR bands may be visible even though incorrect in size, or not clear enough to warrant sequencing.

While our study successfully characterized gB genotypes using both QPCR-based genotyping and sequencing and phylogenetic analysis we were unable to generate UL54 and UL97 sequence data for drug resistance analysis. Despite the challenges encountered, this study provides insights into the circulating CMV strains in Johannesburg, South Africa, and their diversity. We tried to sequence single fragments more than 1200bp in size for both UL54 and UL97 whereas in a study by Kleiboeker *et al.* they amplified and sequenced a single fragment of approximately 600bp for the UL97 gene whereas they generated four fragments across the UL54 gene ranging in size from 337bp to 798bp (175). The dominance of gB3 strains suggests they are primarily responsible for CMV disease in the region. Detailed clinical information would have offered a deeper understanding of clinical outcomes related to specific strains, as previous studies have hypothesised that different CMV strains may result in different unique clinical manifestations (176).

This study had several limitations and challenges, including a smaller sample size, limited geographic scope, and lack of comprehensive clinical data. The sample size calculated was 194

but even after extension of the study period the sample only reached 111 and of these 25 had to be excluded whereafter we remained with 86 samples for analysis. Unfortunately, the fact that CMV viral load testing is not conducted in the NHLS Virology laboratory at CMJAH residual specimens of cases tested at the referral laboratory (based at Chris Hani Baragwanath Hospital, Soweto, South Africa) resulted in logistical constraints as these samples were not regularly returned as requested. Technical issues, time constraints, and financial limitations also hindered the amplification of UL54 and UL97 genes. While PCR primers utilized for UL54 and UL97 in this study did not yield satisfactory results despite exploring multiple PCR conditions it should be noted that Sahoo *et al.* (2013) used CMV virus isolates while clinical samples were used here. Subsequent, mapping of the primers showed that they were located outside of the genes of interest (UL54 and UL97). Therefore, future studies explore the use of primers described by Kleiboeker *et al.* as they used clinical samples. Strengthening the surveillance of UL97 and UL54 gene mutations will also inform antiviral drug development, optimize treatment options, and improve disease management (175). Future studies should prioritize characterization of gB protein variations across diverse geographic locations to inform vaccine development. The study also suffers from information bias, selection and sampling bias, as some samples had incomplete information and were conveniently collected.

4.2 CONCLUSION

In conclusion, this study demonstrated that the gB3 strain is the most dominant CMV strain in Johannesburg, South Africa during 2021 to 2023 compared to earlier periods when gB2 dominated. The findings of this study contribute to the understanding of CMV genetic diversity and have implications for public health strategies and treatment approaches. This research underscores the significance of continued research into CMV genetic composition, particularly in relation to disease progression, treatment outcomes, and CMV vaccine development.

REFERENCES

1. Andrei G, De Clercq E, Snoeck R. Drug Targets in Cytomegalovirus Infection. Infect Disord Drug Targets. 2012;9(2):201–22.

2. Sun L, Chen J min, Yang K, Zhang L, Ma Z yuan, Chen X mei, et al. Cytomegalovirus cell tropism and clinicopathological characteristics in gastrointestinal tract of patients with HIV/AIDS. Diagn Pathol. 2022 Dec 1;17(1):90.

3. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. Rev Med Virol. 2010 Jul;20(4):202–13.

4. Bates M, Brantsaeter AB. Human cytomegalovirus (CMV) in Africa: a neglected but important pathogen. J Virus Erad. 2016;2(3):136–42.

5. Huang CY, Cheng YC, Hwang YS, Kang EYC, Hsiao CH. Cytomegalovirus Glycoprotein B Genotype in Patients with Anterior Segment Infection. Int J Mol Sci. 2023 Apr 1;24(7):6304.

6. Handous I, Hannachi N, Achour B, Marzouk M, Hazgui O, Khelif A, et al. Cytomegalovirus Glycoprotein Polymorphisms and Increasing Viral Load in Non-Transplant Patients with Hematological Malignancies Undergoing Chemotherapy: A Prospective Observational Study. Infect Dis Ther. 2021 Sep 1;10(3):1549–66.

7. Biron KK. Antiviral drugs for cytomegalovirus diseases. Antiviral Res. 2006;71(2-3 SPEC. ISS.):154–63.

8. Walti CS, Khanna N, Avery RK, Helanterä I. New Treatment Options for Refractory/Resistant CMV Infection. Transplant International. 2023 Oct 12 ;36(10):11785.

9. Ahmed A. Antiviral Treatment of Cytomegalovirus Infection. Infect Disord Drug Targets. 2012;11(5):475–503.

10. Gilbert C, Boivin G. Human cytomegalovirus resistance to antiviral drugs. Antimicrob Agents Chemother. 2005 Mar;49(3):873–83.

11. Nkosi NP. Analysis of cytomegalovirus UL97 drug resistance mutation [Internet]. [Cape Town]: University of Cape Town; 2017 [cited 2024 Aug 29]. Available from: https://open.uct.ac.za/items/f1f11009-a1c0-4d18-986a-93277f4 Dec20

12. Nolan N, Halai UA, Regunath H, Smith LP, Rojas-Moreno C, Salzer W. Primary cytomegalovirus infection in immunocompetent adults in the United States – A case series. IDCases. 2017; 10:123–6.

13. Manicklal S, Emery VC, Lazzarotto T, Boppana SB, Gupta RK. The “Silent” global burden of congenital cytomegalovirus. Clin Microbiol Rev. 2013;26(1):86–102.

14. Lurain NS, Chou S. Antiviral drug resistance of human cytomegalovirus. Clin Microbiol Rev. 2010;23(4):689–712.

15. Mhandire D, Rowland-Jones S, Mhandire K, Kaba M, Dandara C. Epidemiology of Cytomegalovirus among pregnant women in Africa. J Infect Dev Ctries. 2019 Oct 31;13(10):865–76.

16. U.S Food and Drug Administration. FDA briefing document [Internet]. Silver Spring, MD; 2021 Sep [cited 2024 Jun 1]. Available from: https://www.fda.gov/media/152176/download

17. Mcintosh M, Hauschild B, Miller V. Human cytomegalovirus and transplantation: drug development and regulatory issues. J Virus Erad. 2016 Jul 1;2(3):143–8.

18. Drew LawrenceW. Cytomegalovirus resistance testing: Pitfalls and problems for the clinician. Clinical Infectious Diseases. 2010;50(5):733–6.

19. Weller T, Macauley J, Craig J, Wirth P. Isolation of intranuclear inclusion-producing agents from infants with illnesses resembling cytomegalic inclusion disease. Journal of Clinical Investigation. 1957;36(1):1–12.

20. Ho M. The history of cytomegalovirus and its diseases. Med Microbiol Immunol. 2008 Jun;197(2):65–73.

21. Zuhair M, Smit GSA, Wallis G, Jabbar F, Smith C, Devleesschauwer B, et al. Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. Rev Med Virol. 2019 May 1;29(3):e2034.

22. Flanders D, Lally C, Dilley A, Diaz-Decaro J. Estimated cytomegalovirus seroprevalence in the general population of the United States and Canada. J Med Virol. 2024 Mar 1;96(3):e29525.

23. Bartlett JG. Cytomegalovirus infection in immunocompetent patients. Infectious Diseases in Clinical Practice. 2004;12(3):212–3.

24. Bartlett AW, Hall BM, Palasanthiran P, McMullan B, Shand AW, Rawlinson WD. Recognition, treatment, and sequelae of congenital cytomegalovirus in Australia: An observational study. Journal of Clinical Virology. 2018 Nov 1;108:121–5.

25. Griffiths P, Baraniak I, Reeves M. The pathogenesis of human cytomegalovirus. Journal of Pathology. 2015 Jan 1;235(2):288–97.

26. Grgic I, Gorenec L. Human cytomegalovirus (HCMV) genetic diversity, drug resistance testing and prevalence of the resistance mutations: A literature review. Trop Med Infect Dis. 2024 Feb 1;9(2):49.

27. Mana H Al, Yassine HM, Younes NN, Al-Mohannadi A, Al-Sadeq DW, Alhababi D, et al. The current status of cytomegalovirus (CMV) prevalence in the MENA region: A systematic review. Pathogens. 2019 Oct 31;8(4):213.

28. Revello MG, Gerna G. Human cytomegalovirus tropism for endothelial/epithelial cells: Scientific background and clinical implications. Rev Med Virol. 2010 May;20(3):136–55.

29. Davis NL, King CC, Kourtis AP. Cytomegalovirus infection in pregnancy. Birth Defects Res. 2017 Mar 15;109(5):336–46.

30. Simon DM, Levin S. Infectious complications of solid organ transplantation. Infect Dis Clin North Am. 2001;15(3):521–49.

31. Azevedo LS, Pierrotti LC, Abdala E, Costa SF, Strabelli TMV, Campos SV, et al. Cytomegalovirus infection in transplant recipients. Clinics. 2015 Jul 28;70(7):515–23.

32. Humar A, Snydman D. Cytomegalovirus in solid organ transplant recipients. American Journal of Transplantation. 2009;9(Suppl 4): S78–86.

33. Fishman JA, Emery V, Freeman R, Pascual M, Rostaing L, Schlitt HJ, et al. Cytomegalovirus in transplantation - Challenging the status quo. Clin Transplant. 2007;21(2):149–58.

34. Preiksaitis JK, Brennan DC, Fishman J, Allen U. Canadian society of transplantation consensus workshop on cytomegalovirus management in solid organ transplantation final report. American Journal of Transplantation. 2005 Feb;5(2):218–27.

35. Crum NF, Riffenburgh RH, Wegner S, Agan BK, Tasker SA, Spooner KM, et al. Comparisons of Causes of Death and Mortality Rates Among HIV-Infected Persons Analysis of the Pre-, Early, and Late HAART (Highly Active Antiretroviral Therapy) Eras. Epidemiology and Social Science. 2006;41(2):194–200.

36. Munro M, Yadavalli T, Fonteh C, Arfeen S, Lobo-Chan AM. Cytomegalovirus retinitis in HIV and non-HIV individuals. Microorganisms. 2020 Jan 1;8(1):55.

37. Seitz R. Human Cytomegalovirus (HCMV)-Revised. Vol. 37, Transfusion Medicine and Hemotherapy. 2010. p. 365–75.

38. Murphy E, Yu D, Grimwood J, Schmutz J, Dickson M, Jarvis MA, et al. Coding potential of laboratory and clinical strains of human cytomegalovirus. Proceedings of the National Academy of Sciences. 2003;100(25):17976–14981.

39. Nikolich-Žugich J, van Lier RAW. Cytomegalovirus (CMV) research in immune senescence comes of age: overview of the 6th International Workshop on CMV and Immunosenescence. Geroscience. 2017 Jun 1;39(3):245–9.

40. Wang YQ, Zhao XY. Human Cytomegalovirus Primary Infection and Reactivation: Insights From Virion-Carried Molecules. Front Microbiol. 2020 Jul 14; 11: 1511.

41. Baldick CJ, Shenk T. Proteins Associated with Purified Human Cytomegalovirus Particles. J Virol. 1996;70(9):6097–105.

42. Seitz R. Human Cytomegalovirus (HCMV)-Revised. Transfusion Medicine and Hemotherapy. 2010 Dec;37(6):365–75.

43. Griffiths PD, Grundy JE. Review Article Molecular biology and immunology of cytomegalovirus. Biochemical Journal. 1987;241(2):313–24.

44. Chen DH, Jiang H, Lee M, Liu F, Zhou H. Three-dimensional visualization of tegument/capsid interactions in the intact human cytomegalovirus. Virology. 1999;260(1):10–6.

45. Landini MP, Placa M La. Humoral immune response to human cytomegalovirus proteins: A brief review. Vol. 14, Comp. lmmun. Microbiol. infect. Dis. 1991.

46. Schleiss MR. Congenital Cytomegalovirus Infection: Molecular Mechanisms Mediating Viral Pathogenesis. Infect Disord Drug Targets. 2011 Oct;11(5):449–65.

47. Landolfo S, Gariglio M, Gribaudo G, Lembo D. The human cytomegalovirus. Vol. 98, Pharmacology and Therapeutics. Elsevier Inc.; 2003. p. 269–97.

48. Navarro D, Lennette E, Tugizov S, Pereira L. Humoral immune response to functional regions of human cytomegalovirus glycoprotein B. J Med Virol. 1997 Aug;52(4):451–9.

49. Scientific Image and Illustration Software | BioRender [Internet]. [cited 2025 Jul 20]. Available from: https://www.biorender.com/

50. Hasan MohdR, Sharma P, Anzar N, Pundir CS, Pilloton R, Narang J, et al. Analytical methods for detection of human cytomegalovirus clinched biosensor a cutting-edge diagnostic tool. Biomedical Engineering Advances. 2021 Jun; 1: 100006.

51. Pignatelli S, Dal Monte P, Rossini G, Landini MP. Genetic polymorphisms among human cytomegalovirus (HCMV) wild-type strains. Rev Med Virol. 2004 Nov;14(6):383–410.

52. Dolan A, Cunningham C, Hector RD, Hassan-Walker AF, Lee L, Addison C, et al. Genetic content of wild-type human cytomegalovirus. Journal of General Virology. 2004;85(5):1301–12.

53. Ye L, Qian Y, Yu W, Guo G, Wang H, Xue X. Functional Profile of Human Cytomegalovirus Genes and Their Associated Diseases: A Review. Front Microbiol. 2020 Sep 4;11:2104.

54. Van Damme E, Van Loock M. Functional annotation of human cytomegalovirus gene products: An update. Front Microbiol. 2014 May 19;5:218.

55. Murphy E, Shenk T. Human Cytomegalovirus Genome. In: Shenk TE, Stinski MF, editors. Human Cytomegalovirus. Springer Berlin Heidelberg; 2008. p. 1–19.

56. Arav-Boger R. Strain Variation and Disease Severity in Congenital Cytomegalovirus Infection: In Search of a Viral Marker. Infect Dis Clin North Am. 2015;29(3):401–14.

57. Hage E, Wilkie GS, Linnenweber-Held S, Dhingra A, Suárez NM, Schmidt JJ, et al. Characterization of human cytomegalovirus genome diversity in immunocompromised hosts by whole-genome sequencing directly from clinical specimens. Journal of Infectious Diseases. 2017 Jun 1;215(11):1673–83.

58. Noriega V, Redmann V, Gardner T, Tortorella D. Diverse immune evasion strategies by human cytomegalovirus. Immunol Res. 2012 Dec;54(1–3):140–51.

59. Sinzger C, Digel M, Jahn G. Cytomegalovirus Cell Tropism. In: Shenk T, Stinski M, editors. Human Cytomegalovirus. Springer; 2008. p. 63–83.

60. Dong N, Cao L, Zheng D, Su L, Lu L, Dong Z, et al. Distribution of CMV envelope glycoprotein B, H and N genotypes in infants with congenital cytomegalovirus symptomatic infection. Front Pediatr. 2023 Mar 15;11.

61. Shenk T, Stinski MF. Human cytomegalovirus genome. Reddehase M, editor. Vol. 325, CURRENT TOPICS IN MICROBIOLOGY AND IMMUNOLOGY. Berlin: Springer; 2008. 1–19 p.

62. Rasmussen L, Geissler A, Winters M. Inter-and Intragenic Variations Complicate the Molecular Epidemiology of Human Cytomegalovirus. Cytomegalovirus Gene Variation • JID. 2003;187(5):809–28.

63. Sijmons S, Van Ranst M, Maes P. Genomic and functional characteristics of human cytomegalovirus revealed by next-generation sequencing. Vol. 6, Viruses. MDPI AG; 2014. p. 1049–72.

64. Li S, Xie Y, Zheng C, XU Z. The battle between host antiviral innate immunity and immune evasion by cytomegalovirus. Cellular and Molecular Life Sciences. 2024;81(5):341.

65. Vanarsdall AL, Johnson DC. Human cytomegalovirus entry into cells. Curr Opin Virol. 2012 Feb;2(1):37–42.

66. Lopper M, Compton T. Coiled-Coil Domains in Glycoproteins B and H Are Involved in Human Cytomegalovirus Membrane Fusion. J Virol. 2004 Aug;78(15):8333–41.

67. Ogawa-Goto K, Tanaka K, Gibson W, Moriishi E, Miura Y, Kurata T, et al. Microtubule Network Facilitates Nuclear Targeting of Human Cytomegalovirus Capsid. J Virol. 2003 Aug;77(15):8541–7.

68. Shan L, Li S, Meeldijk J, Blijenberg B, Hendriks A, van Boxtel KJWM, et al. Killer cell proteases can target viral immediate-early proteins to control human cytomegalovirus infection in a noncytotoxic manner. PLoS Pathog. 2020 Apr 1;16(4).

69. Collins-McMillen D, Buehler J, Peppenelli M, Goodrum F. Molecular determinants and the regulation of human cytomegalovirus latency and reactivation. Vol. 10, Viruses. MDPI AG; 2018.

70. Sinclair J, Sissons P. Latency and reactivation of human cytomegalovirus. Journal of General Virology. 2006 Jul;87(7):1763–79.

71. Kalejta RF. Tegument Proteins of Human Cytomegalovirus. Microbiology and Molecular Biology Reviews. 2008 Jun;72(2):249–65.

72. Tamrakar S, Kapasi AJ, Spector DH, McElroy AK, Yen J, Tamrakar S, et al. Human Cytomegalovirus Infection Induces Specific Hyperphosphorylation of the Carboxyl-Terminal Domain of the Large Subunit of RNA Polymerase II That Is Associated with Changes in the Abundance, Activity, and Localization of cdk9 and cdk7. J Virol. 2005;79(24):15477–93.

73. Per J, Klasse J, Bron R, Marsh M. Mechanisms of enveloped virus entry into animal cells. Adv Drug Deliv Rev. 1998;34(1):65–91.

74. Reeves MB, Macary PA, Lehner PJ, Sissons JGP, Sinclair JH. Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. Proc Natl Acad Sci U S A. 2005 Feb 1;102(11):4140–5.

75. Prichard MN, Penfold MET, Duke GM, Spaete RR, Kemble GW. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. Vol. 11, Reviews in Medical Virology. 2001. p. 191–200.

76. Chevalier MS, Daniels GM, Johnson DC. Binding of Human Cytomegalovirus US2 to Major Histocompatibility Complex Class I and II Proteins Is Not Sufficient for Their Degradation. J Virol. 2002 Aug 15;76(16):8265–75.

77. Stern-Ginossar N, Weisburd B, Michalski A, Thuy V, Le K, Hein MY, et al. Decoding human cytomegalovirus. Science (1979). 2012 Nov 23;338(6110):1088–93.

78. Jackson SE, Redeker A, Arens R, van Baarle D, van den Berg SPH, Benedict CA, et al. CMV immune evasion and manipulation of the immune system with aging. Geroscience. 2017 Jun 1;39(3):273–91.

79. Vasilieva E, Gianella S, Freeman ML. Novel strategies to combat CMV-related cardiovascular disease. Pathog Immun. 2020;5(1):240–74.

80. Wang H, Peng G, Bai J, He B, Huang K, Hu X, et al. Cytomegalovirus infection and relative risk of cardiovascular disease (ischemic heart disease, stroke, and cardiovascular death): A meta-analysis of prospective studies up to 2016. J Am Heart Assoc. 2017 Jul 1;6(7).

81. Gugliesi F, Pasquero S, Griffante G, Scutera S, Albano C, Castillo Pacheco SF, et al. Human cytomegalovirus and autoimmune diseases: Where are we? Viruses. 2021 Feb 1;13(2):260.

82. Yoo SG, Han K Do, Lee KH, La Y, Kwon DE, Han SH. Impact of cytomegalovirus disease on new-onset type 2 diabetes mellitus: Population-based matched case-control cohort study. Diabetes Metab J. 2019 Dec 1;43(6):815–29.

83. Chaemsupaphan T, Limsrivilai J, Thongdee C, Sudcharoen A, Pongpaibul A, Pausawasdi N, et al. Patient characteristics, clinical manifestations, prognosis, and factors associated with gastrointestinal cytomegalovirus infection in immunocompetent patients. BMC Gastroenterol. 2020 Jan 30;20(1).

84. Wang X, Chen J, Cao Z, Yu X. Associations between human cytomegalovirus infection and type 2 diabetes mellitus: a systematic review and meta-analysis. BMJ Open. 2023 Aug 24;13(8).

85. Kotton CN, Kumar D, Caliendo AM, Huprikar S, Chou S, Danziger-Isakov L, et al. The Third International Consensus Guidelines on the Management of Cytomegalovirus in Solid-organ Transplantation. Transplantation. 2018 Jun 1;102(6):900–31.

86. Freeman RB. The “Indirect” effects of cytomegalovirus infection: Minireview. American Journal of Transplantation. 2009 Nov;9(11):2453–8.

87. Smedbråten Y V., Sagedal S, Leivestad T, Mjøen G, Osnes K, Rollag H, et al. The impact of early cytomegalovirus infection after kidney transplantation on long-term graft and patient survival. Clin Transplant. 2014;28(1):120–6.

88. Steininger C, Puchhammer-Stöckl E, Popow-Kraupp T. Cytomegalovirus disease in the era of highly active antiretroviral therapy (HAART). Journal of Clinical Virology. 2006 Sep;37(1):1–9.

89. Gutiérrez-Delgado EM, Villanueva-Lozano H, García Rojas-Acosta MJ, Miranda-Maldonado IC, Ramos-Jiménez J. A case report of small bowel perforation secondary to cytomegalovirus related immune reconstitution inflammatory syndrome in an AIDS patient. Annals of Medicine and Surgery. 2017 Jan 1;13:20–3.

90. Gonzales Zamora J, Espinoza L. Histoplasma and Cytomegalovirus Coinfection of the Gastrointestinal Tract in a Patient with AIDS: A Case Report and Review of the Literature. Diseases. 2017 Dec 8;5(4):30.

91. Mulder R, Tichelaar YIGV, Sprenger HG, Mulder AB, Lijfering WM. Relationship between cytomegalovirus infection and procoagulant changes in human immunodeficiency virus-infected patients. Clinical Microbiology and Infection. 2011;17(5):747–9.

92. Sabin CA, Phillips AN, Lee CA, Janossy G, Emery V, Griffiths PD. The effect of CMV infection on progression of human immunodeficiency virus disease in a cohort of haemophilic men followed for up to 13 years from seroconversion. Epidemiol Infect. 1995;114(2):361–72.

93. Gianella S, Massanella M, Richman DD, Little SJ, Spina CA, Vargas M V., et al. Cytomegalovirus Replication in Semen Is Associated with Higher Levels of Proviral HIV DNA and CD4 + T Cell Activation during Antiretroviral Treatment. Hahn BH, editor. J Virol. 2014 Jul 15;88(14):7818–27.

94. Gianella S, Anderson CM, Var SR, Oliveira MF, Lada SM, Vargas M V., et al. Replication of Human Herpesviruses Is Associated with Higher HIV DNA Levels during Antiretroviral Therapy Started at Early Phases of HIV Infection. J Virol. 2016 Apr 15;90(8):3944–52.

95. Koziej S, Kowalczyk E, Niemczuk M, Jasiuk A, Wiekiera M. Congenital CMV Infection: A Complex Case of Neurological Complications and Therapeutic Approaches in Infancy. Journal of Education, Health and Sport. 2024 Jan 20;54:158–68.

96. Chiopris G, Veronese P, Cusenza F, Procaccianti M, Perrone S, Daccò V, et al. Congenital cytomegalovirus infection: Update on diagnosis and treatment. Microorganisms. 2020 Oct 1;8(10):1–17.

97. Gong Y, Moström M, Otero C, Valencia S, Tarantal AF, Kaur A, et al. Mathematical Modeling of Rhesus Cytomegalovirus Transplacental Transmission in Seronegative Rhesus Macaques. Viruses. 2023 Oct 1;15(10):2040.

98. Tagami M, Honda S, Morioka I, Iijima K, Yamada H, Nakamura M. An unusual case of congenital cytomegalovirus infection-related retinopathy. BMC Ophthalmol. 2016;16(1):81.

99. Swanson EC, Schleiss MR. Congenital Cytomegalovirus Infection. New Prospects for Prevention and Therapy. Pediatr Clin North Am. 2013 Apr;60(2):335–49.

100. Simonazzi G, Rizzo N. Congenital cytomegalovirus infection: prognostic value of maternal DNAemia at amniocentesis. Clinical Infectious Diseases Advance Access. 2016;63(3):391–2.

101. Eltayib HAA, Mudathir AA, Mohamed AAH, Ayman A. Associations of cytomegalovirus with type I diabetes mellitus among children in Khartoum State. Afr J Microbiol Res. 2014 Apr 16;8(16):1730–4.

102. Leung Ki EL, Venetz JP, Meylan P, Lamoth F, Ruiz J, Pascual M. Cytomegalovirus infection and new-onset post-transplant diabetes mellitus. Clin Transplant. 2008 Mar;22(2):245–9.

103. Atandi Batchy A, Magloire L, Boumba A, Elenga-Bongo C, Loubano Voumbi G, Pouki FS, et al. Human Cytomegalovirus Infection Associated with Low Insulin Secretion in a Type 1 Diabetic Population in Pointe Noire. Med Clin Res Open Access. 2022;3(1):1–5.

104. Du Y, Zhang G, Liu Z. Human cytomegalovirus infection and coronary heart disease: A systematic review. Virol J. 2018 Feb 6;15(1):31.

105. Melnick J, Dreesman G, McCollum C, Petrie B, Burek J, Debakey M. Cytomegalovirus antigen within human arterial smooth muscle cells. Lancet. 1983 Sep 17;2(8351):644–7.

106. Blum A, Giladi M, Weinberg M, Kaplan G, Pasternack H, Laniado S, et al. High Anti-Cytomegalovirus (CMV) IgG Antibody Titer is Associated With Coronary Artery Disease and May Predict Post-Coronary Balloon Angioplasty Restenosis. American Journal of Cardiology. 1998;81(7):866–8.

107. Kovacs A, Weber ML, Burns LJ, Jacob HS, Vercellotti GM. Cytoplasmic Sequestration of p53 in Cytomegalovirus-Infected Human Endothelial Cells. American Journal of Pathology. 1996;149(5):1531–9.

108. Nikitskaya E, Grivel J, Maryukhnich E, Lebedeva A, Ivanova I, Savvinova P, et al. Cytomegalovirus in Plasma of Acute Coronary Syndrome Patients. Acta Naturae. 2015;8(2):102–7.

109. Cristescu CV, Alain S, Rut SM. Clinical Medicine The Role of CMV Infection in Primary Lesions, Development and Clinical Expression of Atherosclerosis. J Clin Med. 2022;11(13):3832.

110. Beyaz MO, Ugurlucan M, Oztas DM, Meric M, Conkbayir C, Agacfidan A, et al. Evaluation of the relationship between plaque formation leading to symptomatic carotid artery stenosis and cytomegalovirus by investigating the virus DNA. Archives of Medical Science – Atherosclerotic Diseases. 2019 Mar 4;4(1):19–24.

111. Ross SA, Novak Z, Pati S, Boppana SB. Diagnosis of Cytomegalovirus Infections. Infect Disord Drug Targets. 2011;11(5):466–74.

112. Wang X, Li X, Hu S, Qu H, Zhang Y, Ni H, et al. Rapid detection of active human cytomegalovirus infection in pregnancy using loop-mediated isothermal amplification. Mol Med Rep. 2015 Aug 1;12(2):2269–74.

113. Valle-Arroyo J, Aguado R, Páez-Vega A, Pérez AB, González R, Fornés G, et al. Lack of cytomegalovirus (CMV)-specific cell-mediated immune response using QuantiFERON-CMV assay in CMV-seropositive healthy volunteers: fact not artifact. Sci Rep. 2020 Dec 1;10(1):7194.

114. Lazzarotto T, Gabrielli L, Rizzo R. New diagnostics and methods of assessing pregnant women at risk of cytomegalovirus. Microbiol Aust. 2015;36(4):167–9.

115. Moon SM, Sung H, Kim MN, Lee SO, Choi SH, Kim YS, et al. Diagnostic yield of the cytomegalovirus (CMV) antigenemia assay and clinical features in solid organ transplant recipients and hematopoietic stem cell transplant recipients with CMV pneumonia. Transplant Infectious Disease. 2012 Apr;14(2):192–7.

116. Hong SI, Kim T, Park SY, Jung J, Lee JY, Chong YP, et al. Sensitivity of the cytomegalovirus antigenemia assay to diagnose cytomegalovirus retinitis. Infect Chemother. 2016;48(4):302–8.

117. Razonable RR, Inoue N, Pinninti SG, Boppana SB, Lazzarotto T, Gabrielli L, et al. Clinical diagnostic testing for human cytomegalovirus infections. Journal of Infectious Diseases. 2020;221:S74–85.

118. Razonable RR, Hayden RT. Clinical utility of viral load in management of cytomegalovirus infection after solid organ transplantation. Clin Microbiol Rev. 2013;26(4):703–27.

119. Bruminhent J, Bushyakanist A, Kantachuvesiri S, Kiertiburanakul S. A Nationwide Survey of Cytomegalovirus Prevention Strategies in Kidney Transplant Recipients in a Resource-Limited Setting. Open Forum Infect Dis. 2019 Oct 5;6(Suppl 2):S641.

120. Meesing A, Razonable RR. New Developments in the Management of Cytomegalovirus Infection After Transplantation. Drugs. 2018 Jul 1;78(11):1085–103.

121. Razonable RR, Humar A. Cytomegalovirus in solid organ transplantation. American Journal of Transplantation. 2013 Mar;13(SUPPL.4):93–106.

122. Lee H, Oh EJ. Laboratory diagnostic testing for cytomegalovirus infection in solid organ transplant patients. Vol. 36, Korean Journal of Transplantation. Korean Society for Transplantation; 2022. p. 15–28.

123. Kwong AD, Rao BG, Jeang KT. Viral and cellular RNA helicases as antiviral targets. Nat Rev Drug Discov. 2005 Oct;4(10):845–53.

124. Pooja D, Sabreen J, Zhenyu Z. A Case Report of Cytomegalovirus Esophagitis in an Immune-competent Patient. Int J Curr Res Rev. 2018;10(12):8–11.

125. U.S. Food and Drug Administration C for DE and R. Cytomegalovirus in Transplantation: Developing Drugs to Treat or Prevent Disease Guidance for Industry [Internet]. Silver Spring, MD; 2020 May [cited 2024 Jun 1]. Available from: https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs

126. Faulds D, Heel CR. A Review of its Antiviral Activity, Pharmacokinetic Properties and Therapeutic Efficacy in Cytomegalovirus Infections. Drugs. 1990;39(4):597–38.

127. Mcgavin JK, Goa KL, Brennan DC. Ganciclovir An Update of its Use in the Prevention of Cytomegalovirus Infection and Disease in Transplant Recipients. 2001;61(8):1153–83.

128. AstraZeneca LP. FOSCAVIR (foscavir sodium) injection prescribing information [package insert]. 2012 [cited 2025 Aug 22]; Available from: https://www.accessdata.fda.gov/drugsafda_doc/label/2012/020068s108

129. Razonable RR. Management strategies for cytomegalovirus infection and disease in solid organ transplant recipients. Infect Dis Clin North Am. 2013 Jun;27(2):317–42.

130. De Clercq E, Holý A. Case history: Acyclic nucleoside phosphonates: A key class of antiviral drugs. Vol. 4, Nature Reviews Drug Discovery. 2005. p. 928–40.

131. Cihlar T, Fuller MD, Cherrington JM. Characterization of Drug Resistance-Associated Mutations in the Human Cytomegalovirus DNA Polymerase Gene by Using Recombinant Mutant Viruses Generated from Overlapping DNA Fragments. J Virol. 1998;72(7):5927–36.

132. Verghese PS, Schleiss MR. Letermovir Treatment of Human Cytomegalovirus Infection Antiinfective Agent. Drugs Future. 2013 May;38(5):291–8.

133. Neuber S, Wagner K, Goldner T, Lischka P, Steinbrueck L, Messerle M, et al. Mutual Interplay between the Human Cytomegalovirus Terminase Subunits pUL51, pUL56, and pUL89 Promotes Terminase Complex Formation. J Virol. 2017 Jun 15;91(12):e02384-16.

134. Corinna LR, Yoonsuh P, Dongyun Yang Qiao Z, Teodora K, Nicole K, Jerry C, et al. Cytomegalovirus Triplex vaccine in pediatric hematopoietic stem cell transplant patients at high risk for cytomegalovirus complications: evaluation of vaccine safety, immunogenicity and impact on viremia requiring antivirals. Haematologica. 2024 Jul;109(7):2303–8.

135. Tzannou I, Watanabe A, Naik S, Daum R, Kuvalekar M, Leung KS, et al. "mini" bank of only 8 donors supplies CMV-directed T cells to diverse recipients. Blood Adv. 2019 Sep 10;3(17):2571–80.

136. Gourin C, Alain S, Hantz S. Anti-CMV therapy, what next? A systematic review. Front Microbiol. 2023 Nov 20;14:1321116.

137. Kotton CN, Kamar N. New Insights on CMV Management in Solid Organ Transplant Patients: Prevention, Treatment, and Management of Resistant/Refractory Disease. Infect Dis Ther. 2023 Feb 1;12(2):333–42.

138. Stanat SC, Reardon,' JE, Erice A, Jordan MC, Drew WL, Biron' KK. Ganciclovir-Resistant Cytomegalovirus Clinical Isolates: Mode of Resistance to Ganciclovir. Antimicrob Agents Chemother. 1991;35(11):2191–7.

139. Smith IL, Shinkai M, Freeman WR, Spector SA. Polyradiculopathy Associated with Ganciclovir-Resistant Cytomegalovirus in an AIDS Patient: Phenotypic and Genotypic Characterization of Sequential Virus Isolates. Journal of Infectious Diseases. 1996 Jun;173(6):1481.

140. Wolf DG, Smith IL, Lee DJ, Freeman WR, Flores-Aguilar M, Spector SA. Mutations in Human Cytomegalovirus UL97 Gene Confer Clinical Resistance to Ganciclovir and Can Be Detected Directly in Patient Plasma. Journal of Clinical Investigation. 1995 Jan;95(1):257–63.

141. Drew WL, Miner RC, Busch DF, Follansbee SE, Gullett J, Mehalko SG, et al. Prevalence of Resistance in Patients Receiving Ganciclovir for Serious Cytomegalovirus. J Infect Dis. 1991 Apr;163(4):716–9.

142. Hakki M, Chou S. The biology of cytomegalovirus drug resistance. Curr Opin Infect Dis. 2011 Dec 24;24(6):605–11.

143. Chiereghin A, Belotti T, Borgatti EC, Fraccascia N, Piccirilli G, Fois M, et al. Off-label use of letermovir as preemptive anti-cytomegalovirus therapy in a pediatric allogeneic peripheral blood stem cell transplant. Infect Drug Resist. 2021;14:1185–90.

144. Peixoto M, Mascarenhas L, Cunha A, Dutra M, Miranda E, Silva I, et al. Recurrent and persistent cytomegalovirus infection in a kidney recipient caused by the L595S mutation in UL97 phosphotransferase gene. Antivir Ther. 2012;17(3):585–8.

145. Chou S, Song K, Wu J, Bo T, Crumpacker C. Drug Resistance Mutations and Associated Phenotypes Detected in Clinical Trials of Maribavir for Treatment of Cytomegalovirus Infection. Journal of Infectious Diseases. 2022 Aug 15;226(4):576–84.

146. Gilbert C, Azzi A, Goyette N, Lin SX, Boivin G. Recombinant phenotyping of cytomegalovirus UL54 mutations that emerged during cell passages in the presence of either ganciclovir or foscarnet. Antimicrob Agents Chemother. 2011 Sep;55(9):4019–27.

147. Torii Y, Horiba K, Kawada J ichi, Haruta K, Yamaguchi M, Suzuki T, et al. Detection of antiviral drug resistance in patients with congenital cytomegalovirus infection using long-read sequencing: a retrospective observational study. BMC Infect Dis. 2022 Dec 1;22(1):568.

148. James SH, Price NB, Hartline CB, Lanier ER, Prichard MN. Selection and recombinant phenotyping of a Novel CMX001 and cidofovir resistance mutation in human cytomegalovirus. Antimicrob Agents Chemother. 2013 Jul;57(7):3321–5.

149. Plotkin SA, Starr SE, Friedman HM, Gonczol E, Weibel RE. Protective Effects of Towne Cytomegalovirus Vaccine Against Low-Passage Cytomegalovirus Administered as a Challenge. J Infect Dis. 1989;159(5):860–5.

150. Esposito S, Chiopris G, Messina G, D'alvano T, Perrone S, Principi N. Prevention of congenital cytomegalovirus infection with vaccines: State of the art. Vaccines (Basel). 2021 May 1;9(5):523.

151. Scarpini S, Morigi F, Betti L, Dondi A, Biagi C, Lanari M. Development of a vaccine against human cytomegalovirus: Advances, barriers, and implications for the clinical practice. Vaccines (Basel). 2021 Jun 1;9(6):551.

152. Nakamura R, Rosa C La, Longmate J, Drake J, Slape C, Zhou Q, et al. Viraemia, immunogenicity, and survival outcomes of cytomegalovirus chimeric epitope

vaccine supplemented with PF03512676 (CMVPepVax) in allogeneic haemopoietic stem-cell transplantation: Randomised phase 1b trial. Lancet Haematol. 2016 Feb 1;3(2):e87–98.

153. Bernstein DI, Munoz FM, Callahan ST, Rupp R, Wootton SH, Edwards KM, et al. Safety and efficacy of a cytomegalovirus glycoprotein B (gB) vaccine in adolescent girls: A randomized clinical trial. Vaccine. 2016 Jan 12;34(3):313–9.

154. Gyurova IE, Schlums H, Sucharew H, Ambroggio L, Ochayon DE, Win HT, et al. Dynamic Changes in Natural Killer Cell Subset Frequencies in the Absence of Cytomegalovirus Infection. Front Immunol. 2019 Nov 22;10.

155. Griffiths P, Plotkin S, Mocarski E, Pass R, Schleiss M, Krause P, et al. Desirability and feasibility of a vaccine against cytomegalovirus. Vaccine. 2013 Apr 18;31(Suppl 2):B197–203.

156. DAIDS Clinical Laboratory, Oversight Team (DCLOT), Health Services L. DAIDS Good Clinical Laboratory Practice Guidelines [Internet]. Bethesda; 2021 Aug [cited 2024 Sep 1]. Available from: https://www.niaid.nih.gov/research/daids-

157. Pang X, Humar A, Preiksaitis JK. Concurrent genotyping and quantitation of cytomegalovirus gB genotypes in solid-organ-transplant recipients by use of a real-time PCR assay. J Clin Microbiol. 2008 Dec;46(12):4004–10.

158. Nelson CS, Vera Cruz D, Su M, Xie G, Vandergrift N, Pass RF, et al. Intrahost Dynamics of Human Cytomegalovirus Variants Acquired by Seronegative Glycoprotein B Vaccinees. J Virol. 2019 Mar;93(5).

159. Katoh K, Standley DM. Article Fast Track MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol [Internet]. 2013 Jan 16 [cited 2025 Aug 25];30(4):772–80. Available from: https://academic.oup.com/mbe/article/30/4/772/1073398

160. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program. Nucleic Acids Symp Ser. 1999;41:95–8.

161. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021 Jul 1;38(7):3022–7.

162. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.

163. Sahoo MK, Lefterova MI, Yamamoto F, Waggoner JJ, Chou S, Holmes SP, et al. Detection of cytomegalovirus drug resistance mutations by next-Generation sequencing. J Clin Microbiol. 2013 Nov;51(11):3700–10.

164. Vilsker M, Moosa Y, Nooij S, Fonseca V, Ghysens Y, Dumon K, et al. Genome Detective: An automated system for virus identification from high-throughput sequencing data. Bioinformatics. 2019 Mar 1;35(5):871–3.

165. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. Nucleic Acids Res. 2013 Jan 1;41(D1):D36–42.

166. Lewis PO. Society of Systematic Biologists. Syst Biol. 1995 Dec;44(4):576–7.

167. Thermo Fisher Scientific. QuantStudio 3/5 Real-Time PCR Software [Internet]. 2025 [cited 2025 Aug 22]. Available from: https://www.thermofisher.com/za/en/home/technical-resources/software-downloads/quantstudio-3-5-real-time-pcr-systems.html

168. Baloyi S, Selabe S, Kyaw T. Human cytomegalovirus gB genotype distribution among HIV infected patients from Pretoria. International Journal of Infectious Diseases. 2014 Apr;21:449–50.

169. Govender K, Parboosing R, Camiolo S, Hubáček P, Görzer I, Puchhammer-Stöckl E, et al. Complexity of Human Cytomegalovirus Infection in South African HIV-Exposed Infants with Pneumonia. Viruses. 2022 May 1;14(5):12–21.

170. Külekci B, Schwarz S, Brait N, Perkmann-Nagele N, Jaksch P, Hoetzenecker K, et al. Human cytomegalovirus strain diversity and dynamics reveal the donor lung as a major contributor after transplantation. Virus Evol. 2022 Aug 24;8(2):1–13.

171. Görzer I, Kerschner H, Jaksch P, Bauer C, Seebacher G, Klepetko W, et al. Virus load dynamics of individual CMV-genotypes in lung transplant recipients with mixed-genotype infections. J Med Virol. 2008 Aug;80(8):1405–14.

172. Vasiljevic T, Jankovic M, Tomic A, Bakrac I, Radenovic S, Miljanovic D, et al. Significance of Cytomegalovirus gB Genotypes in Adult Patients Undergoing Hematopoietic Stem Cell Transplantation: Insights from a Single-Centre Investigation. Pharmaceuticals. 2024 Apr 1;17(4):428.

173. Tarragó D, Quereda C, Tenorio A. Different cytomegalovirus glycoprotein B genotype distribution in serum and cerebrospinal fluid specimens determined by a novel multiplex nested PCR. J Clin Microbiol. 2003 Jul 1;41(7):2872–7.

174. Sarkar A, Das D, Ansari S, Chatterjee RP, Mishra L, Basu B, et al. Genotypes of glycoprotein B gene among the Indian symptomatic neonates with congenital CMV infection. BMC Pediatr. 2019 Aug 22;19(1):291.

175. Kleiboeker SB, Juranic A, et al. Simultaneous detection and genotyping of cytomegalovirus UL97 and UL54 mutations associated with ganciclovir resistance by PCR and sequencing. Journal of Clinical Microbiology. 2004;42(7): 2984–2992.

176. Hu H, Cheng Y, Peng Q, Chen K. Cytomegalovirus genotype distribution among postnatally infected infants: Association of glycoprotein B, glycoprotein N and glycoprotein H types with CMV-associated thrombocytopenia. Mediterr J Hematol Infect Dis. 2020 Aug 4;12(1):e2020057.

Appendix A: Permission from the corresponding to use Figure 1.3

Gmail — Petronela Letlonkane <pp.letlonkane@gmail.com>

Re: Permission Request to Reuse Figure 1 (CMV Life Cycle) in Thesis
1 message

Alan Zheng <zheng.alan@hotmail.com> 28 August 2025 at 17:00
To: Petronela Letlonkane <pp.letlonkane@gmail.com>

Sure, but I am not sure Springer Nature will agree to it.

You should contact them.

Dr. Chunfu Zheng
Adjunct Professor
Department of Microbiology, Immunology & Infection Diseases
University of Calgary,
Health Research Innovation Centre
3330 Hospital Drive NW, Calgary, AB T2N 4N1 Canada
https://cumming.ucalgary.ca/departments/microinfect/chunfu-zheng

From: Petronela Letlonkane <pp.letlonkane@gmail.com>
Sent: August 28, 2025 12:25 AM
To: yuchangyin68@163.com <yuchangyin68@163.com>
Cc: zheng.alan@hotmail.com <zheng.alan@hotmail.com>; docxzc@126.com <docxzc@126.com>
Subject: Permission Request to Reuse Figure 1 (CMV Life Cycle) in Thesis

Dear Dr. Li, Dr. Xie, Dr. Yu, Dr. Zheng, and Dr. Xu,

I hope this message finds you well. My name is Pholenyana Petronela Letlonkana, a postgraduate student at the University of the Witwatersrand, South Africa. I am completing my Master's dissertation titled "Cytomegalovirus diversity among patients of all ages in Johannesburg, South Africa, 2022–2023."

I recently read your article "The battle between host antiviral innate immunity and immune evasion by cytomegalovirus" (Cellular and Molecular Life Sciences, 2024), which I found to be highly valuable to my research.

I would like to kindly request your permission to reproduce Figure 1 (CMV life cycle) from your publication in my dissertation. The figure will be used strictly for academic, non-commercial purposes, and I will ensure that full credit and citation are given to your original work.

Please let me know if you require any additional information, forms, or conditions to grant permission.

Thank you very much for your time and consideration. I deeply appreciate your contribution to advancing knowledge in this important field.

Warm regards,
Pholenyana Petronela Letlonkana
Master's Candidate – Health Sciences (Pathology)
University of the Witwatersrand
Email: pp.letlonkane@gmail.com | ORCID: https://orcid.org/0000-0001-7250-4767

Appendix B: Copyright letter to use Figure 1.3

Dear PP Letlonkana,

Thank you for contacting CCC's Customer Service Department. My name is Alessandra, and I will assist you today. At CCC, we partner with many copyright owners, such as Springer Nature, to provide licensing services to reproduce or display their content.

I was unable to locate any article based on the title and DOI information you provided. However, I believe you may be referring to "The battle between host antiviral innate immunity and immune evasion by cytomegalovirus", which was published as an Open Access article under the CC BY-NC-ND 4.0 Deed.

This license permits non-commercial use, sharing, distribution, and reproduction in any medium or format, provided that you give appropriate credit to the original authors and source, include a link to the Creative Commons license, and indicate if any modifications have been made to the licensed material.

If you plan to reuse the figure exactly as it appears in your thesis, you do not need to obtain a formal permission license from the publisher. That said, I would strongly encourage you to carefully review the full terms the license: https://creativecommons.org/licenses/by-nc-nd/4.0/.

I hope this information is helpful, and if you have any further questions about CC services, please don't hesitate to reach out again.

Kind regards,
Alessandra

Alessandra Savu
Customer Account Specialist
Copyright Clearance Center
222 Rosewood Drive
Danvers, MA 01923

Appendix C: Ethics certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO: Ms PP Letlonkana
School of Pathology
Department of Virology
Medical School
University

E-mail: pp.letlonkane@gmail.com

CC: Supervisor: Drs A Glass, F Treurnicht and B Mashishi
<Allison.Glass@lancet.co.za>
and <HREC-Medical Research Office@wits.ac.za>

FROM: Mr Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 2022/03/31

REF: R14/49

PROTOCOL NO: **M210822** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)

PROJECT TITLE: *Cytomegalovirus diversity and drug resistance profiles from patients in Johannesburg, South Africa, 2021-2022*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.

MSWorks2000/Iain0007/Clearscan.wps

Appendix D: Plagiarism report

Petronela dissertation final 291124.docx

ORIGINALITY REPORT

16%	13%	12%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	1%
2	wiredspace.wits.ac.za Internet Source	1%
3	www.ncbi.nlm.nih.gov Internet Source	<1%
4	mdpi-res.com Internet Source	<1%
5	www.scribd.com Internet Source	<1%
6	"Principles and Practice of Transplant Infectious Diseases", Springer Science and Business Media LLC, 2019 Publication	<1%
7	Viral Infections of Humans, 2014. Publication	<1%
8	discovery.ucl.ac.uk Internet Source	<1%
	hdl.handle.net	