

RESEARCH ARTICLE OPEN ACCESS

A New Next-Generation Sequencing Approach in Human Cytomegalovirus for the Identification of Antiviral Resistance Mutations and Genotypic Classification

Maria Arnedo-Muñoz^{1,2}  | Ignasi Prats-Méndez¹ | Alejandra González-Sánchez¹ | Maria Carmen Martín¹ | Ariadna Rando-Segura^{1,3} | Ibai Los Arcos^{4,5} | Patricia Nadal-Barón¹ | Narcís Saubi¹ | Natalia Mendoza-Palomar⁶ | Marta Bernat-Sole¹ | Adaia Albasanz-Puig⁴ | Maria Piñana¹ | Cristina Andrés¹ | Susana Melendo-Pérez⁶ | Nieves Larrosa^{1,5} | Andrés Antón^{1,5} | Juliana Esperalba^{1,2,5}

¹Microbiology Department, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Universitat Autònoma de Barcelona, Barcelona, Spain | ²Department of Genetics and Microbiology, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain | ³CIBER Hepatic and Digestive Diseases (CIBERehd), Instituto Carlos III, Madrid, Spain | ⁴Infectious Diseases Department, Vall d'Hebron Barcelona Hospital Campus, Universitat Autònoma de Barcelona, Bellaterra, Spain | ⁵Centro de Investigación Biomédica en red de Enfermedades Infecciosas CIBERINFEC, Instituto Carlos III, Madrid, Spain | ⁶Pediatric Infectious Diseases and Immunodeficiencies Unit, Hospital Infantil. Vall d'Hebron Barcelona Hospital Campus, Institut de Recerca Vall d'Hebron, Barcelona, Spain

Correspondence: Andrés Antón (andres.anton@vallhebron.cat) | Juliana Esperalba (juliana.esperalba@vallhebron.cat)

Received: 12 February 2025 | **Revised:** 27 May 2025 | **Accepted:** 19 July 2025

Keywords: anti-herpesvirus drug < antiviral agents | antiviral agents | cytomegalovirus < virus classification | human cytomegalovirus < virus classification | mutation < genetics | resistance < infection

ABSTRACT

This study introduces a new procedure for antiviral resistance analysis and genetic classification of human cytomegalovirus (HCMV) using next-generation sequencing (NGS) adapted to existing methodologies, aiming for more targets due to the recent use of new antivirals. It expands the classical investigation of mutations in UL54 and UL97 genes, associated with resistance to (val)ganciclovir, foscarnet and cidofovir, to include UL27, UL56 and UL89 genes, which target newer antivirals like maribavir and letermovir. Additionally, it includes the genetic analysis of UL55 (glycoprotein B) for genotype classification. This new methodology involves multiplex-PCR for DNA enrichment, followed by NGS using Illumina MiSeq platform and data analysis through an in-house pipeline. Several validations were performed by firstly using genome sequence from wild-type sensitive reference strain (AD-169), secondly comparing to previously characterized samples by Sanger, and lastly the use of external quality controls. A new NGS technique based on amplicons approach has been developed. Validation using wild-type control material showed 100% identity with the reference genome across all replicates. Only one minority variant was detected in one replicate. Compared to Sanger sequencing, NGS revealed additional low-frequency mutations not detected by Sanger, without impacting resistance interpretation. The method also performed reliably in external quality assessment controls. Moreover, the detection limit of the technique was established at 17 894.60 IU/mL. Finally, this approach enabled the identification of HCMV genotypes. This approach improves the monitoring of antiviral resistance and viral diversity, enhancing early clinical decision-making in immunocompromised patients.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Journal of Medical Virology* published by Wiley Periodicals LLC.

1 | Introduction

Human Cytomegalovirus (HCMV) is a prevalent virus, with an estimated seroprevalence of 83% [1]. In immunocompetent individuals, HCMV infections are usually asymptomatic and self-limiting, due to effective immune control of the infection. However, in immunosuppressed patients, such as those undergoing solid organ transplantation (SOT) or hematopoietic stem cell transplantation (HSCT), HCMV can cause significant morbidity or mortality [2]. These patients often require antiviral drugs to prevent or treat infection and/or illness. Most of the antivirals used are ganciclovir (GCV) and valganciclovir (VGV), followed by foscarnet (FOS) and cidofovir (CDV). The recently approved drugs, maribavir (MBV) and letermovir (LMV) are prescribed to selected patients with the aim to reduce side effects associated with conventional antivirals [3].

Under selective pressure, resistance to current antiviral drugs can arise throughout the HCMV genome, even at low frequencies. The best characterized resistance-associated regions are the UL54 and UL97 genes, which encode the viral DNA polymerase (UL54) and the viral kinase (UL97), respectively. More specifically, single or cross-resistance to GCV/VGV, FOS, and CDV are related to mutations found between codons 301–987 in UL54 [4]. On the other hand, resistance to GCV/VGV or MBV are related to mutations in UL97, mainly concentrated between codons 337 and 557 [4]. Additionally, there are currently other genes of interest for resistance studies. The UL27 gene encodes a nuclear protein, and mutations in this gene only confer resistance to MBV [5]. Alternatively, UL56 and UL89 genes encode for DNA packaging terminase subunits, both involved in viral DNA maturation, and mutations in these genes can confer LMV resistance [6]. Most of the characterized mutations have been regarded primarily as single nucleotide polymorphisms, not associated with antiviral resistance. However, many mutations in these regions have not been phenotypically characterized to date [7].

Screening for HCMV mutations associated with antiviral resistance in clinical practice has traditionally been based on Sanger sequencing. Methods described in the literature predominantly focus on mutations in the UL97 and UL54 but have not been updated to include other critical regions. Although Sanger is effective, its main limitation lies in its sensitivity, as it cannot detect variants present at levels below 20%. This can be overcome by using Next Generation Sequencing (NGS), which can identify variants present at allele frequencies (AF) lower than 5% [8].

Additionally, HCMV can be classified into different genotypes based on several genes, being glycoprotein B (UL55) the most commonly used marker. By sequencing a highly variable region of this gene, covering from nucleotide positions 1051 to 1845, HCMV can be classified into four genotypes (gB1, gB2, gB3, and gB4). Although the correlation between genotype and clinical outcomes remains unclear, this highlights the significance of genetic studies [9].

This study aimed to develop and validate for its clinical use the designed NGS procedure for sequencing UL27, UL54, UL55, UL56, UL89, and UL97 genes of HCMV to enable the detection

of antiviral resistance. It also intends to facilitate the genetic classification of the virus by sequencing the UL55 gene. With this new approach we intend to adapt other previously described techniques to be able to extend them to other genes that require further study, either because mutations have been detected in these regions or new antivirals have been developed.

2 | Materials and Methods

2.1 | Primer Design—Overlapping Amplicons

Candidate primers were proposed using Primal Scheme [10], accessible at <https://primalscheme.com> and based on the reference sequence NC_006273.2 (Human herpesvirus 5 strain Merlin, complete genome). With this approach, 400–800 bp-long amplicons were generated, covering all six genes. These primer sequences were then refined through a Multiple Sequence Analysis (MSA) of all available genomic data from GenBank database until February 2023, to ensure genetic diversity coverage. Selection criteria included primers of approximately 20 nucleotides, with similar theoretical melting temperature and targeting conserved regions. Degenerate bases were added in exceptional cases to enhance variability, optimizing primer applicability for viral genome amplification.

2.2 | Enrichment Multiplex PCR

Primer sets were grouped in three different multiplex reactions to avoid primer dimerization, and PCR conditions were further optimized (information not provided). Viral DNA was amplified using Q5 High-Fidelity DNA Polymerase (M0491, New England Biolabs Inc. USA). This PCR master mix contained: primer pools described on Table 1 at a final concentration of 0.08 μ M for pool 1 or 0.1 μ M for pools 2 and 3, 1 $\times$ Q5 Reaction Buffer, 0.2 mM dNTPs, less than 10 ng of DNA template, 0.02 U/ μ L Q5 High-Fidelity DNA Polymerase, 1 $\times$ Q5 High GC Enhancer, and nuclease-free water up to a total volume of 25 μ L. The optimal PCR conditions were an initial denaturation step at 98°C for 15 min, followed by 35 cycles of 95°C for 15 s and 62°C for 5 min. A final extension step was performed at 62°C for 5 min, and then the reaction was held at 4°C.

2.3 | Library Preparation and Next-Generation Sequencing

Following amplification, PCR-products were purified using AMPureXP beads (Beckman Coulter, USA), removing non-specific DNA fragments outside the expected size range, and were quantified and normalized to 500 ng for NGS library preparation by Illumina DNAprep protocol (Illumina Inc. USA), according to the manufacturer's guidelines. Library quantification and further normalization were done using the Qubit 4 Fluorometer (ThermoFisher, USA), and pooled and loaded onto the MiSeq sequencing platform (Illumina, USA) for NGS. The sequencing was performed using a bidirectional flow cell, with paired-end reads of 101 nucleotides length. For the sequencing run, 105 samples, of which 40 belong to the validation

TABLE 1 | Main features of primers designed to perform multiplex PCR enrichment.

Gene	Name	Primer Pool	Volume (μL)	Primer sequence (5' -> 3')				Length	% GC	Theoretical melting temperature (°C)	Start-end position in whole genome NC_006273.2	Start - End on each gene	Amplicon size	Percentage of coverage of sequenced region
UL27	1_Fw	1	3	ATGAACCCCGTGGATCAGCC			20	60	59.9	34 989–35 008	1–20	757	97.04% (1773/1827)	
	1_Rv	1	3	CGGTACTTGACGGGGCAACCC			20	65	60.9	34 270–34 251	758–739			
	2_Fw	2	3	ATGCTCGATGACCGAGGAGC			20	60	59.4	34 363–34 382	627–646	827		
	2_Rv	2	3	ATCCGCTGCACCACTCGAG			20	62.5	61.2	33 574–33 555	1454–1435			
	3_Fw	3	3	GTCCACGATCTGAAGCGCAT			20	55	57.8	33 618–33 637	1372–1391	441		
	3_Rv	3	3	CTCCGACCTCGTGAGGCCGA			20	70	63.6	33 215–33 196	1813–1794			
UL54	1_Fw	1	3	TCAAYCCGTATCTGAGCGGC			20	57.5	58.3	81 896–81 915	Aug-27	791	98.71% (3681/3729)	
	1_Rv	1	3	CCTGCTGCCGCCAGTCGTAA			20	65	62.6	81 143–81 124	799–780			
	2_Fw	2	3	GARGTCCGTGTGGATCCGCT			20	62.5	60.6	81 201–81 220	703–722	833		
	2_Rv	2	3	GCGTTGCCCGCAGGTAAAGCTC			21	76	69	80 407–80 387	1536–1516			
	3_Fw	3	6	CCYGTATGCATGGCCAAAGAC			20	57.5	57.5	80 442–80 461	1462–1481	812		
	3_Rv	3	6	CTCTARCGTGACGCTGTATA			21	64.3	63.4	79 674–79 655	2274–2255			
	4_Fw	1	3	TGCCGGGYGGCGAGTACCC			18	75	64.7	79 692–79 710	2218–2237	772		
	4_Rv	1	3	CCCTTGACGAACTCGCADGC			20	61.7	59.7	78 964–78 945	2990–2971			
	5_Fw	2	3	CTGAGCATGAAGGGCGGTGGA			20	60	60	78 981–79 000	2935–2954	806		
	5_Rv	2	3	TCAACAGCAATTCGTGCGCCTT			20	55	60.2	78 214–78 194	3741–3721			
UL55	1_Fw	3	5	TTCTGGGAAGCCTCGGAACG			20	55	57.4	83 729–83 748	1070–1090	775	94.80% (751/792)	
	1_Rv	3	5	CTGACATTCTCAGTGCGGTG			21	47.6	54	82 974–82 954	1845–1825			
UL56	1_Fw	1	3	GCAACGAATACGCCATGGAGC			21	57.1	59.2	87 238–87 258	47–67	781	95.55% (2465/2553)	
	1_Rv	1	3	CTCCTCCGCCCTCCTGGATGTA			20	65	61.2	86 497–86 477	828–808			
	2_Fw	2	3	GCTGGAGTCCTTTCGGCCCT			20	65	62.2	86 535–86 554	751–770	763		
	2_Rv	2	3	CGCGCGTCCACCAARTCGAGA			21	64.3	63.4	85 811–85 791	1514–1494			
	3_Fw	3	5	CGGYTTACGGGTGCGCAACT			20	62.5	62.1	85 828–85 847	1458–1477	765		
	3_Rv	3	5	CGGCCACTCTTTGTTATACGT			21	52.4	59.2	85 102–85 082	2223–2203			
UL89	4_Fw	1	3	GACGAACTGCACCCGGACAGA			21	61.9	61.7	85 115–85 135	2170–2190	383	98.02% (5811/5928)	
	4_Rv	1	3	TTAACGCAGACTACCAAGGCAC			21	52.4	57.1	84 772–84 752	2553–2533			
	1_Fw	2	3	CTGCAGAAAGCGCAAGAGCCAC			21	61.9	61.5	139 629–139 649	49–69	788		
	1_Rv	2	3	GTCGATGCTGATCACGTTGTC			22	54.5	58.2	138 881–138 861	837–817			

(Continues)

TABLE 1 | (Continued)

Gene	Name	Primer		Volume (μL)	Primer sequence (5' -> 3')	Length	% GC	Theoretical melting temperature (°C)	Start-end position in whole genome NC_006273.2	Start - End on each gene	Amplicon size	Percentage of coverage of sequenced region
		Pool										
UL96&UL97	2_Fw	3	3	CCGCCACACCTTCGCGCGGA	18	72	66.9	138 901-139 821	777-797	859		
	2_Rv	3	3	CCGACGTGATTACGAGGTCTG	22	54.5	57.4	138 084-138 063	1636-1615			
	3_Fw	1	3	ACTGCCCTTCTGACAGGTCAG	20	54.5	58.2	138 146-138 165	1534-1553	925		
	3_Rv	1	3	GTCATGCCCTTGGCGCTCTGGAT	22	59.1	62.2	137 261-137 240	2459-2438			
	4_Fw	2	3	GCACCCCGGCTCAGACGAGG	18	72	63	137 305-137 324	2375-2394	841		
	4_Rv	2	3	GGCAGCTATCGTGACTGGGA	20	60	59.2	136 502-136 483	3216-3197			
	5_Fw	3	4	TGGTGCTGCCGGTCGCGYTT	20	67.5	66.5	136 546-136 565	3134-3153	834		
	5_Rv	3	4	GTACAGCTAAGTCATGACGC	21	55	57.1	135 750-135 731	3968-3949			
	6_Fw	1	5	TGGATCACGTCGCGGTAACG	20	60	59.8	135 759-135 778	3921-3940	915		
	6_Rv	1	5	GCCTCGTCCACCAGCAGCAA	20	65	63	134 882-134 863	4836-4817			
	7_Fw	2	3	GCATCCGAGGACAAAACCTCCA	22	50	57.6	134 884-134 905	4794-4815	753		
	7_Rv	2	3	AGGCAGGCGATGCGCACG	18	65	64.6	134 169-134 152	5547-5530			
	8_Fw	3	3	TRCACCCCTTACCTGGACGAA	20	52.5	57	134 211-134 230	5469-5489	431		
	8_Rv	3	3	TMTCGTCAACAGGTAGGTG	20	52.5	56.1	133 818-133 799	5900-5881			
	1_Fw	1	5	GTGATGCGCGTCGACCTT	18	61.11	60.63	141 327-141 344	34-51	774		96.46% (2535/2628)
	1_Rv	1	5	AAGTCGCGTGTCCAAGCAC	19	57.89	61.17	142 101-142 083	808-790			
2_Fw	2	6	CCAGGCTCACGTCGATGAAG	20	60	60.82	141 980-141 999	687-706	816			
2_Rv	2	6	CTCTCGTCGCTCATGTCCAC	20	60	60.54	142 798-142 779	1503-1486				
3_Fw	3	3	GAAAGTCAGGACAGCGCCG	19	63.2	58.9	142 488-142 506	1195-1213	700			
3_Rv	3	3	ATGAGCACGTTTCATGGGT	18	50	54.3	143 188-143 171	1895-1878				
4_Fw	1	3	GTGCCTTTTGACAGTTGGC	19	57.9	58.4	143 098-143 116	1805-1823	802			
4_Rv	1	3	GCAGTCACCGTCAAGGTCCTC	21	61.9	60	143 900-143 880	2607-2587				

process, were loaded for HCMV sequencing with this new designed technique.

2.4 | Data Processing and Analysis

An in-house pipeline was developed to analyse NGS data. Raw reads were quality trimmed with Trimmomatic v0.39 [11] ($Q \geq 30$), and primer sequences were removed with BBDMap v38.91 to minimize the risk of artifacts linked to primer binding sites. Finally, host reads (GRCh38) were excluded using Bowtie2 v2.5.1 [12].

High-quality reads were assembled *de novo* using SPAdes v3.15.2 [13]. Scaffold were identified with BLASTN v2.14.0 + [14], classified as HCMV, and assigned to UL regions. Genes linked to antiviral resistance (UL27, UL54, UL56, UL89, and UL97) were mapped to reference genome NC_006273.2 using Minimap2 [15]. Consensus sequences were generated via custom Python script. Variants with allele frequency (VAF) $> 50\%$ and depth $< 20\times$ were masked.

For UL55, mapped reads were extracted for *de novo* assembly. Genotyping was performed using BLASTN v2.14.0 + [14], followed by mapping against UL55 references (gB1, FJ616285; gB2, BK000394; gB3, GU937742; and gB4, M60926). The resulting SAM files were converted to BAM using SAMtools v1.17 [16], and variant calling was performed with LoFreq v2.1.5 [17]. Variants were filtered using BCFtools v1.17 [16] ($AF \geq 50\%$, depth $\geq 20\times$), and consensus sequences were generated with BCFtools consensus [16].

Minority variants in resistance-associated genes were analysed when VAF $\geq 5\%$, depth $\geq 100\times$, and strand bias (SB) < 30 ($SB = -10 \times \log_{10}(p \text{ value})$) ($p > 0.001$) values reported by LoFreq [17].

Quality metrics (including median depth, ambiguous bases percentage (N), and coverage percentage) were computed with custom scripts. Coverage and depth plots were generated using an in-house R script.

For antiviral resistance profiling, mutation detection and prevalence assessment were performed using minMutFinder v1.2.0 (<https://github.com/ValldHebron-Bioinformatics/minMutFinder>) [18], which also characterizes mutations based on Tilloy et al. [19].

2.5 | Variant Visualization Through Pileup Inspection

To assess variants flagged by strand bias filtering, a targeted visual inspection was performed on positions with strand bias scores equal or greater than 30 that were expected to be present in specific samples based on prior knowledge. For this purpose, the indelqual-adjusted BAM file generated using LoFreq's indelqual command was indexed and loaded into the Integrative Genomics Viewer (IGV) [20]. Pileup visualization was carried out to examine strand distribution, variant allele frequency and depth of coverage.

2.6 | Repeatability Validation of the Technique for the Amplification of Resistance-Related Genes

A wild-type reference strain from the American Type Culture Collection (ATCC, USA) was employed to validate the accuracy and replicability of the sequencing process. This reference material consists of 100 000 copies/ μL of genomic DNA from human herpesvirus 5 strain AD-169 (VR-538DQ) isolated from MRC-5 cells. Using the entire supplied volume, fourteen replicates were prepared to ensure repeatability. All replicates were processed using the NGS-designed approach, starting on the amplification step, followed by the library preparation and sequencing of each replicate. All replicates were processed in the same library preparation and sequencing run, minimizing potential biases.

2.7 | Accuracy Validation by a Comparison between Sanger and NGS

To validate the NGS-designed protocol with the one currently used in clinical practice, 21 HCMV-positive samples previously studied by Sanger sequencing of UL54 and UL97 were re-examined using this new-designed NGS protocol. Most of the samples were used in analysis of both genes, UL54 and UL97. Specifically, 18 samples were used for the UL54 gene validation, of these, 10 samples did not present any resistance-associated mutations in the UL54 gene, while 8 had resistance-associated mutations in this gene between nucleotide positions 831 to 3180. Regarding UL97 gene, 19 samples were used for the validation, of which, 10 samples did not have known genetic markers, while 9 did have between nucleotide positions 1344 to 2023.

2.8 | Accuracy Validation Using an External Quality Control

To further validate the ability of the protocol to detect specific amino acid mutations in the entire pUL54, pUL56, and pUL97 coding sequences, five external quality control samples were used within the CMVDR24 program (QAV144169_1) of Quality Control for Molecular Diagnostics (QCMD, UK). This program consists of five external specimens to assess the ability of laboratories to detect HCMV drug resistance mutations in UL54 polymerase, UL56 terminase genes, and UL97 kinase (but not in the other genome domains) using their own specific sequencing techniques.

2.9 | Phylogenetic Analysis of UL55 (gB) Region

To validate the genotyping capacity of our assay, we performed a phylogenetic analysis based on the UL55 (gB) region. The analysis included the 14 replicates of the ATCC strain, the UL55 reference sequence used by ATCC (GenBank accession number AC146999), and four representative sequences from the different gB genotypes (GenBank accession numbers: gB1 M60927, gB2 M60931, gB3 M60930.1 and gB4 M60926). All the used sequences were aligned using the MUSCLE algorithm within

MEGA v5.2 software [21]. The phylogenetic tree was constructed using a Maximum Likelihood distance method under the Kimura 2-parameter (K2P) model with a gamma distribution ($G = 2$). The topological accuracy of the internal branch was evaluated by the bootstrap method (500 replicates).

To further assess genotyping accuracy, we generated nine synthetic UL55 datasets simulating all four gB genotypes, including mixed-genotype infections. Reads were simulated with In-SilicoSeq (iss generate) using parameters mimicking an Illumina MiSeq run, consistent with our wet lab workflow [22]. Each sample included six UL regions (UL27, UL54, UL56, UL89, UL97, UL55), comprising 60% of reads, with the remaining 40% from GRCh38.p14 to simulate host background. Each UL gene contributed 10%; in coinfections, UL55 reads were equally split between genotypes. The synthetic FASTQ files were processed through our pipeline, and inferred genotypes were compared with ground truth. Assembled UL55 sequences were then used to build a phylogenetic tree as above.

2.10 | Measurement of the Limit of Detection (LoD)

To determine the LoD of the sequencing technique, a 1:3 dilution series was prepared in triplicate from clinical plasma samples, covering viral loads between 600 and 4 000 000 copies/mL. The LoD was defined as the lowest viral load at which all target genes were successfully sequenced with predefined quality metrics: N content < 5%, coverage > 95%, and median depth $\geq 100\times$, following criteria adapted from Mallory et al. [23].

Four clinical plasma samples with viral loads above 1 000 000 IU/mL were pooled after extraction, using 500 μ L input per sample and obtaining 50 μ L of eluate using EMAG platform (BioMérieux, France) to ensure sufficient volume. The resulting eluate was serially diluted 1:3 in triplicate and quantified with the AltoStar CMV PCR Kit 1.5 (Altona Diagnostics, Germany), calibrated to the 1st WHO International Standard for HCMV (NIBSC 09/162). Each dilution was sequenced in triplicate; replicate 1 was processed on 1 day, and replicates 2 and 3 on another.

Additionally, five clinical HCMV-positive plasma samples (3000–310 000 IU/mL) collected in July 2024 were sequenced to validate the method's applicability to contemporary strains to confirm technique's robustness and relevance for real-world clinical samples, addressing concerns about the use of older or long-stored material in analytical validation studies.

3 | Results

3.1 | Primer Design

A candidate primers list was obtained using PrimalScheme, followed by a comprehensive MSA from GenBank, to assess the impact of genetic diversity on primer sequences. This included 381 coding-region sequences for UL27, 583 for UL54, 526 for

UL56, 453 for UL89, and 805 for UL97. For UL55, one sequence from each of the four genotypes was selected. This process yielded 25 candidate primer pairs, divided into three pools to prevent primer dimer formation as represented in Figure 1, and main features of the primers are described in Table 1.

The primers, designed from coding regions relevant to antiviral resistance, achieved near-complete coverage: 97.04% for pUL27, 98.71% for pUL54, 95.55% for pUL56, 94.17% for pUL89, 99.00% for pUL97. For pUL55, referring to the highly variable sequenced region, the coverage was 94.8%.

3.2 | Reproducibility Evaluation of Wild-Type Control Material

Successful and reproducible results were obtained in the different replicates of the ATCC strain, as a wild-type control to validate the accuracy and replicability of the technique. A 100% coverage, 0% of N, and a uniform sequence depth > 100 $\times$ were achieved across all 14 replicates for every gene, as shown in Figure 2. In addition, consensus sequences matched the ATCC reference with 100% identity for all genes and all 14 replicates. In reference to the minority variants, there was also total concordance between the replicates and the genes, except in the UL89 gene, in which the nucleotide change G91C is detected in replicate 07, which is translated as the amino acid change A31P, which has never been previously described in the literature. This divergence is found at an AF of 10.16%, the mutation is present at a depth of 2106 $\times$ and the SB for this position was 11.

3.3 | NGS Technique Versus Sanger Sequencing

Comparison between NGS and Sanger sequencing revealed differences in UL54 and UL97, shown in Table 2. Regarding UL54, no differences were detected in most of the samples 12/18 (66.67%). In total, 7 differences were detected in the remaining 6/18 samples (33.33%). Among these changes, 5/7 (71.42%) were mutations with an AF lower than 20% and 2/7 (28.58%) had an AF greater than 20%.

On the other hand, for UL97, no differences were found between Sanger and NGS in most of the samples 16/19 (84.21%). Although the mutation 1378 at UL97 gene in Sample 01 and Sample 06 was initially filtered out in the NGS pipeline due to strand bias (SB > 30), it was detected at high allele frequency and confirmed by visual inspection of (Figure S1) showing strong support in high-quality reads. Therefore, this was considered as no difference between NGS and Sanger. A total of 3 differences were found in the remaining: 2/3 (66.67%) with an AF of less than 20% and 1/3 (33.33%) above 20%.

Despite these differences, both techniques yielded 100% concordance in antiviral susceptibility predictions (21/21). In Sample 03, both Sanger and NGS detected the A594V mutation in UL97 (conferring resistance to (Val)Ganciclovir), while NGS additionally identified the C592G mutation, which has also been associated with resistant, although this did not alter susceptibility.

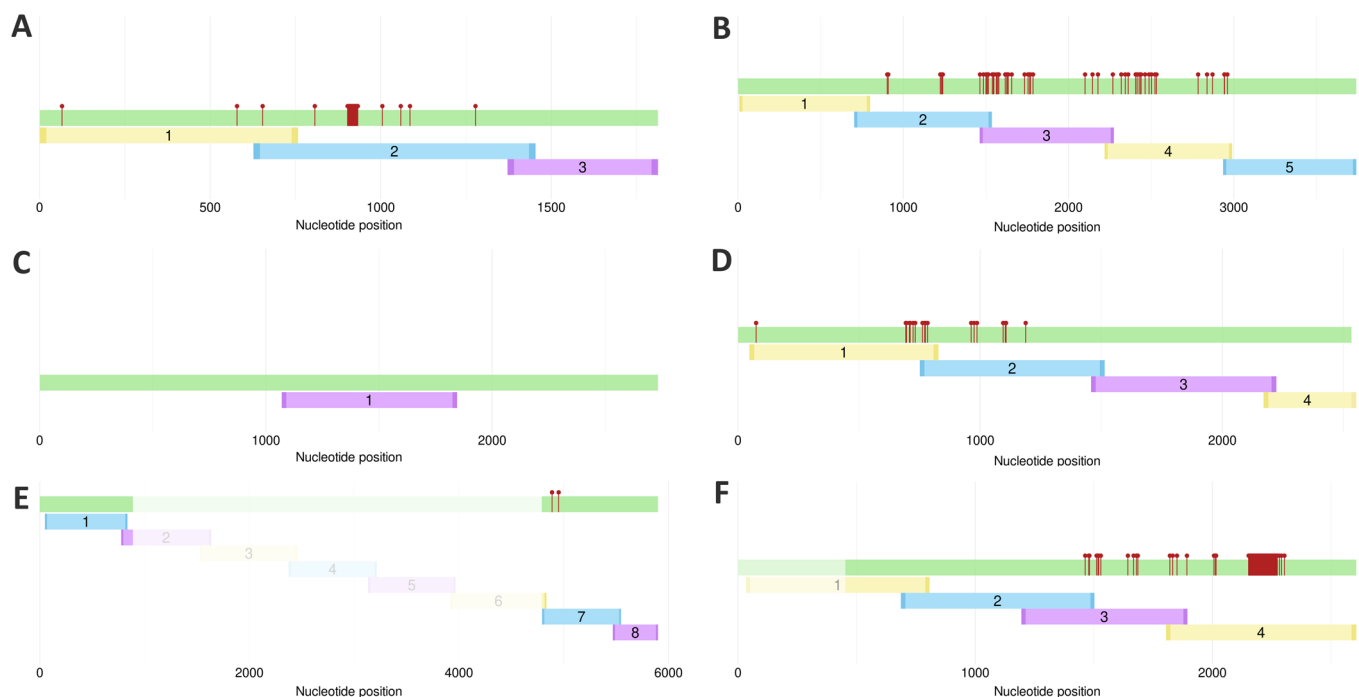


FIGURE 1 | Distribution of designed amplicons across six target genes: UL27, UL54, UL55 (Genotyping region), UL56, UL89, and UL97. Each letter represents one gene (A–F). The darker regions within each amplicon represent primer binding sites. Amplicons in yellow belong to pool 1, in blue to pool 2, and in purple to pool 3. Red lollipop markers represent nucleotide positions associated with known antiviral resistant mutations. Since UL55 gene is used for genotyping purpose, only a region with sufficient divergent information between nucleotide positions 1051 and 1845 is sequenced to determine the viral genotype [9, 24, 25]. While UL89 is sequenced in its entirety, the gray region represents an intron, which is excluded in the subsequent analysis of mutations. For UL97, since this gene does not have a sufficiently conserved region at the beginning to design the primers, it was combined with the coding region of the contiguous gene UL96 (represented in gray), which was excluded from mutation analysis.

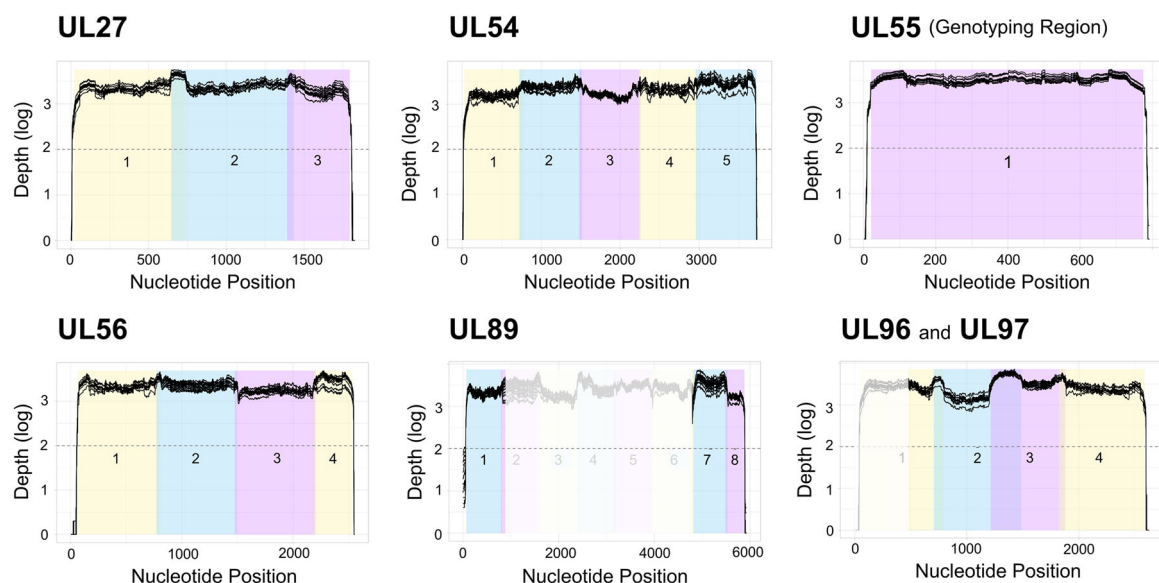


FIGURE 2 | Depth plots of the six protein coding sequences: UL27, UL54, UL55 (genotyping region), UL56, UL89 and UL96/97. The y axis represents coverage on a logarithmic scale, and the x-axis represents nucleotide position for each gene. Dashed lines indicate the minimum accepted coverage of 100x (log 2). Each black line represents a replicate. The colored regions correspond the different primer pools: yellow for pool 1, blue for pool 2, and purple for pool 3. The gray shaded areas indicate the regions that are excluded for mutation analysis.

3.4 | Implementation With External Quality Controls

Participation in CMVDR24 external quality control program (QCMD) was successfully achieved, confirming the technique's

reliability. All six genes met quality thresholds, although this program only assessed the ability to detect mutations in pUL54, pUL97, and pUL56. The detected divergences were compared with the reference sequence and are described in Table 3. Regarding to minority mutations, only the change G2686A in

TABLE 2 | Comparison between Sanger and NGS sequencing results.

Antiviral Characterization				Differences between Sanger and NGS				
Name	Gene	Sanger	NGS	Nucleotide position	Sanger nucleotide	NGS nucleotide	AF	NGS depth
Sample 01	UL54	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	GCV/VGV Resistant (M460V)	GCV/VGV Resistant (M460V)	1378 ^b	G	A	98.21%	1449
Sample 02	UL54	Sensitive	Sensitive	1463	C	G	5.97%	419×
	UL97	GCV/VGV Resistant (A594V)	GCV/VGV Resistant (A594V)	No differences	—	—	—	—
Sample 03	UL54	Sensitive	Sensitive	2407	C	T	5.96%	386×
	UL97	GCV/VGV Resistant (A594V)	GCV/VGV Resistant (A594V; C592G)	1774	T	G	22.64%	680×
	UL54	Sensitive	Sensitive	No differences	—	—	—	—
Sample 04	UL54	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	GCV/VGV Resistant (A594V)	GCV/VGV Resistant (A594V)	No differences	—	—	—	—
Sample 05	UL54	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	GCV/VGV Resistant (M460I)	GCV/VGV Resistant (M460I)	No differences	—	—	—	—
Sample 06	UL97	GCV/VGV Resistant (M460V)	GCV/VGV Resistant (M460V)	1378 ^b	G	A	99.71%	3493
Sample 07	UL54	GCV/VGV/CDV/FOS Resistant (del981-982)	GCV/VGV/CDV/FOS Resistant (del981—982)	1438	G	T	8.88%	980×
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 08	UL54	GCV/VGV/CDV/FOS Resistant (del981-982)	GCV/VGV/CDV/FOS Resistant (del981-982)	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	1412	C	A	9.47%	697×
Sample 09	UL54	GCV/VGV/CDV/FOS Resistant (del981-982)	GCV/VGV/CDV/FOS Resistant (del981-982)	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 10	UL54	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 11	UL54	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 12	UL97	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 13	UL54	Sensitive	Sensitive	1556	A	T	5.82%	636×
	UL97	Sensitive	Sensitive	2275	G	A	42.20%	550×
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
	UL97	Sensitive	Sensitive	No differences	—	—	—	—

(Continues)

TABLE 2 | (Continued)

Antiviral Characterization				Differences between Sanger and NGS				
Name	Gene	Sanger	NGS	Nucleotide position	Sanger nucleotide	NGS nucleotide	AF	NGS depth
Sample 14	UL54	GCV/VGV/CDV Resistant (A987G)	GCV/VGV/CDV Resistant (A987G)	No differences	—	—	—	—
Sample 15	UL97	GCV/VGV Resistant (H520Q)	Resistant (H520Q)	No differences	—	—	—	—
	UL54	Sensitive	Sensitive	1493	A	G	5.89%	1103X
Sample 16	UL97	GCV/VGV Resistant (C592G)	GCV/VGV Resistant (C592G)	No differences	—	—	—	—
	UL54	FOS Resistant (E756Q)	FOS Resistant (E756Q)	2171	C	G	68.02%	813X
Sample 17	UL97	Sensitive	Sensitive	1745	C	T	8.02%	499X
	UL97	Sensitive	Sensitive	No differences	—	—	—	—
Sample 18	UL54	Sensitive	Sensitive	No differences	—	—	—	—
Sample 19	UL97	Sensitive	Sensitive	No differences	—	—	—	—
	UL54	GCV/VGV/CDV Resistant (P522A)	GCV/VGV/CDV Resistant (P522A)	No differences	—	—	—	—
Sample 20	UL54	GCV/VGV/CDV Resistant (P522A)	GCV/VGV/CDV Resistant (P522A)	No differences	—	—	—	—
Sample 21	UL54	GCV/VGV/CDV Resistant (P522A)	GCV/VGV/CDV Resistant (P522A)	No differences	—	—	—	—
Sample 22	UL97	GCV/VGV Resistant (C603W)	GCV/VGV Resistant (C603W)	No differences	—	—	—	—
	UL54	GCV/VGV Resistant (C603W)	GCV/VGV Resistant (C603W)	No differences	—	—	—	—

The text shown in parentheses represents the mutation responsible for the resistance.

^aThis discrepancy is due to strand bias-based filtering. The resistance-associated mutation (UL54, position 1378) was detected by Sanger sequencing and NGS at high frequency, but was filtered out in NGS due to elevated strand bias (SB > 30). Visual inspection confirmed the presence of the mutation in high-quality reads, so there was considered as no differences.

TABLE 3 | Comparison of the variants detected in the mutation resistant sites in reference of the consensus sequence obtained from all participants in the QCMD quality control and those obtained in our laboratory.

Sample	Gene	Position	Reference nucleotide	Obtained nucleotide	Amino acid difference	Antiviral resistance
CMVDR-01	UL54	1634	T	C	L545S	GCV, CDV
		1560	G	G	H520Q ^a	GCV
		1784	T	C	L595S	GCV
		1794	T	C	G598G	Synonymous Mutation
CMVDR-02	UL54	2942-2947	ATCTGG	del	del981-982	GCV, CDV, FOS
	UL97	1780	G	C	A594P	GCV
CMVDR-03	UL54	—	—	—	—	—
	UL97	1378	A	G	M460V	GCV
		1794	T	C	G598G	Synonymous Mutation
CMVDR—04	UL54	1224	C	G	N408K	GCV, CDV
		2522	G	C	G841A	GCV, CDV, FOS
	UL97	1774	T	G	C592G	GCV
CMVDR-05	UL54	—	—	—	—	—
	UL56	975	C	G	C325W	LMV
	UL97	—	—	—	—	—

^aThe amino acid in position 520 in UL97 for CMVDR-01 was excluded from the analysis by QCMD due to lack of consensus.

UL54 was detected in CMVDR-01 sample with a depth = 1491×, an SB = 6 and an AF = 5.97%.

3.5 | Genotyping of HCMV Based on the gB Region

Phylogenetic analysis of the UL55 (gB) region represented in Figure 3 confirmed the ability of the assay to distinguish between different HCMV genotypes. All 14 replicates of the ATCC sample grouped in a single clade together with the ATCC reference sequence (AC146999) and the gB2 genotype reference (M60931). The remaining gB genotype references (gB1, gB3, and gB4) clustered separately.

Phylogenetic analysis of the synthetic FASTQs in Figure 4 confirmed the ability of the pipeline to accurately assign HCMV genotypes, even in mixed infection contexts. Each synthetic sample clustered with the expected gB genotype reference. Sample containing single genotypes (Synthetic FASTQ 1–4) grouped consistent with their corresponding reference clades (gB1–gB4). In samples simulating mixed infections (Synthetic FASTQ 5–9), multiple UL55 sequences were recovered, each clustering correctly with the relevant genotype references.

3.6 | Limit of Detection of the Technique

The last dilution accomplishing the established quality metric criteria presented a viral load of 17 894.60 IU/mL. At this

dilution level, the genes UL27, UL54, UL55, UL56, and UL97 exhibited a 0% median N content and nearly complete coverage ($\geq 99.76\%$), and median read depth ranging from 317× to 1071×. Although UL89 showed a slightly reduced coverage (95.02%) and a higher proportion of undefined bases (median N content of 4.98%), its performance remained within acceptable analytical thresholds. Table S1 shows the results for the other dilutions.

Regarding the sequencing of the five current samples, the designed technique has proven to be equally efficient and has even allowed characterization of samples with a viral load of approximately 3000 IU/mL, which is below the established detection limit. All genes demonstrated a 100% coverage, except UL89, which reached 98.47%, 0% median N percentage in most of the genes, and high sequencing depth, raging between 1376× and 4488×. Values for each sample and gene are detailed in Table S2.

4 | Discussion

Despite advances in HCMV antiviral treatments, available options remain limited, and the emergence of resistant strains, particularly in immunocompromised individuals undergoing SOT or HSCT, poses a significant clinical challenge [2]. Resistance rates range from 0.1% in general population to 18% in high-risk groups due to selective pressures from prolonged antiviral use and the host's compromised immune response [26–28]. These limitations emphasize the need for sensitive, comprehensive diagnostic methods to monitor reactivation events and detect resistance markers early [29–32].

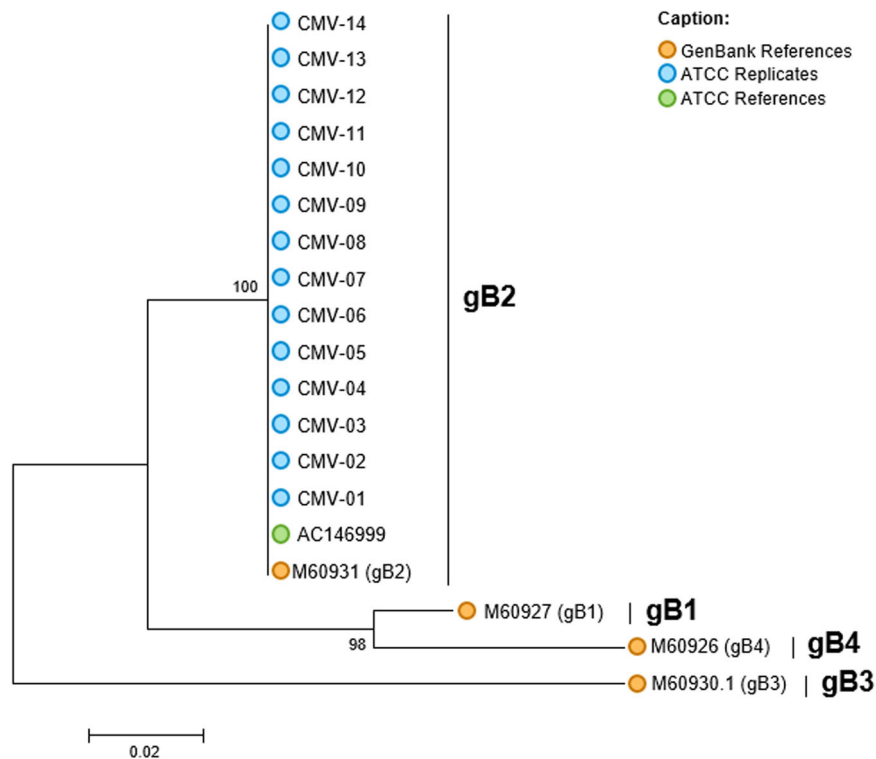


FIGURE 3 | ATCC phylogenetic tree of the UL55 (gB) region constructed by Maximum Likelihood (K2P model with $G = 2$). The 14 technical replicates of the ATCC HCMV strain (labeled CMV-01 to CMV-14) are indicated with blue circles. The ATCC reference genome (GenBank: AC146999) is shown in green, and the four gB genotype reference strains are shown in orange: gB1 (M60927), gB2 (M60931), gB3 (M60930.1), and gB4 (M60926). Scale bar represents 0.02 substitutions per site.

Traditional Sanger sequencing remains the standard in many clinical laboratories, but its limitations, particularly its inability to detect minority variants, difficult the early detection of emerging resistance [33]. The NGS-based techniques developed in this study overcomes these limitations by simultaneously targeting five resistance-associated genes and the gB region, offering both resistance detection and genotypic classification, and also allowing the identification of minority variants at allele frequencies as low as 5%, providing more detailed insight into viral heterogeneity and early resistance emerge [34].

The results obtained demonstrate a high degree of technical consistency. Across all 14 replicates of the wild-type ATCC strain, the method achieved 100% identity with the consensus genome in all target genes, with uniform sequencing depth and zero ambiguous bases. Only a single low-frequency variant (A31P in UL89) was observed in one replicate, passing all quality filters, suggesting stochastic intra-host variability rather than a technical artifact.

Comparison between NGS and Sanger sequencing in clinical samples showed full concordance in resistance interpretation. However, NGS detected additional minority variants with allele frequencies below Sange's 20% threshold, such as C592G in UL97, which despite not affecting phenotype, illustrates NGS's superior sensitivity [35, 36]. Differences between both methods were primarily found in variants with allele frequencies under 20%, likely reflecting underlying viral heterogeneity. These findings underscore NGS's added value in resolving mixed infections, detecting emerging resistance mutations not

captured by traditional methods, and providing semi-quantitative information on allele frequencies, something not achievable with Sanger [37].

While amplification-based techniques offer high sensitivity, they also introduce artifacts from polymerase errors or amplification bias. To mitigate this, the pipeline includes quality control steps such as primer trimming, strand bias filtering, and paired-end validation. Furthermore, manual inspection of minority variants ensured that only those supported by high-quality reads on both strands were retained, increasing confidence in low-frequency mutation calls.

This technique's effectiveness was demonstrated also in an external quality assessment program, achieving robust performance across both laboratory and bioinformatics workflows. The ability to analyse six resistance-associated genes simultaneously makes it a strong competitor to existing in-house and commercial methods. Notably, it identified mutations related to resistance for five of six antiviral drugs, except maribavir, which was not included in the quality control program. Validation with the ATCC strain confirmed sensitivity across all target genes, achieving high identity with reference sequences.

In addition to resistance profiling, genotyping of HCMV based on the UL55 region demonstrated that the assay could reliably classify gB genotypes. Although the clinical relevance of the gB genotypes remains uncertain, its genotyping offers valuable insights for epidemiological surveillance and viral diversity assessment. In our phylogenetic analysis, the 14 replicates of

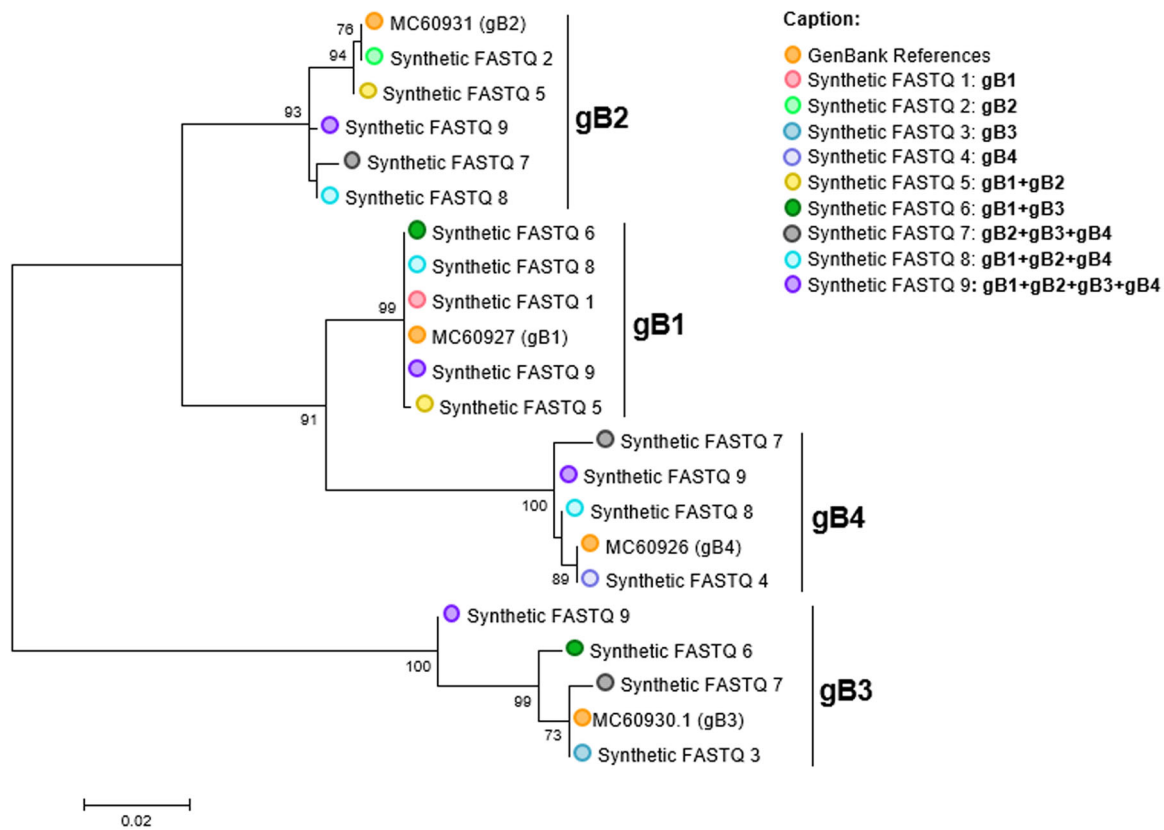


FIGURE 4 | Synthetic FASTQ phylogenetic tree of the UL55 (gB) region constructed by Maximum Likelihood (K2P model with $G = 2$). Each of the 9 synthetic FASTQ were indicated with different colors. gB reference sequences (gB1 (M60927), gB2 (M60931), gB3 (M60930.1), and gB4 (M60926)) were also showed. Synthetic samples clustered with the expected genotypes. Bootstrap values < 70 are shown at key nodes and scale bar represents 0.02 substitutions per site.

the ATCC strain clustered consistently with the ATCC reference genome, which also clustered with gB2 reference sequence. To further validate the resolution of our bio-informatics pipeline for the other genotypes and coinfections, nine synthetic datasets simulating UL55 sequences representing all four known gB genotypes, including mixed-genotype scenarios, were generated. Each one was processed through the pipeline, and the resulting UL55 assemblies were subjected to phylogenetic analysis. In all cases, the reconstructed sequences grouped correctly with their respective genotype references. Notably, in samples simulating coinfections, multiple UL55 variants were simultaneously recovered and assigned to their expected clades, highlighting the script's capacity to resolve complex viral population.

Regarding HCMV's viral load, there is no clear correlation between high viral loads and the emergence of resistance. However, patients with high viral loads over an extended period of time will also prolong treatment, thus increasing the selective pressure on resistant viruses and favoring their proliferation [38]. In particular, there have been efforts to determine the mean viral load in patients with antiviral resistance. A study involving nearly 4000 solid organ recipients estimated the mean viral load in those with resistance to be greater than 30 000 IU/mL ($4.48 \log_{10}$ IU/mL) [39]. Moreover, in our center, since 2014, a total of 235 samples with clinical suspicion of antiviral resistance have been collected, of which the median viral load was 22,028.5 IU/mL (IQR: 7912.75–84368.75)

equivalent to $4.34 \log_{10}$ IU/mL. Dilution series determined a detection limit of 17,894 IU/ml ($4.25 \log_{10}$), which aligns with published estimates of viral loads in patients with resistant HCMV strains. Therefore, in future studies it would be of interest to implement improvements that would lower the detection limit of the technique, enabling the sequencing above the established thresholds for lower viral loads. This would facilitate the routine detection of resistant variants at an early stage, while they are still emerging and before they are selected.

Interestingly, the performance of the designed methodology was also tested by sequencing five current circulating viruses, since the ATCC reference strain and the sequences from the GenBank database were relatively old. All five samples, with viral loads ranging from 3,000 to 310 000 IU/mL, yielded high-quality sequencing data. Notably, this included three of the five samples below the theoretical detection limit, which still achieved complete and reliable results, suggesting a promising performance, that may be refined in future iterations.

These results suggest that, with additional optimization and validation, the method has the potential to be implemented in clinical laboratories equipped with NGS platforms. Moreover, although HCMV is a DNA virus with a relatively low mutation rate, and our assay performed well in both reference and contemporary strains, ongoing surveillance remains essential to ensure sustained applicability over time.

In conclusion, this study presents a new validated NGS-based method that simultaneously enables resistance mutation detection and genotypic classification of HCMV. By expanding the genetic regions analysed and detecting minority variants not seen with Sanger sequencing, this approach shows potential for enhanced sensitivity and broader mutation coverage. In addition to antiviral-resistance mutation, the methodology also enables genotypic classification based on the gB, which could be contributed to improved epidemiological surveillance and strain tracking. This dual-purpose design, targeting both resistance-associated mutations and genotypic classification, highlights the assay's versatility and its potential to streamline HCMV genomic surveillance workflows. Its multi-target design and demonstrated performance across reference, synthetic, and clinical samples support its potential for routine use in diagnostic laboratories equipped with NGS platforms.

Although further clinical validation is needed, including correlation with treatment outcomes and larger sample cohorts, this methodology lays the groundwork for a future NGS-based genotyping strategy that could support the early detection and surveillance of antiviral resistance in clinical virology laboratories.

Author Contributions

The design of the technique as well as the validation and interpretation of the results was performed by M.A.M, I.P.M, A.G.S, M.C.M, M.B.S. The revision of the final work and a critical review of the manuscript has been performed by all the authors but with special mention to A.R.S, I.L.A., P.N.B, N.S, N.M.P, A.A.P, M.P, C.A, S.M.P, N.L. Finally, this study has been rigorous supervised and directed by A.A.P and J.E.E.

Acknowledgments

This study has been carried out in the framework of the doctoral program of microbiology of the "Universitat Autònoma de Barcelona". The results and the preliminary methodology designed were presented at the European congress organized by the European Society of Clinical Virology in September 2024.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. M. Zuhair, G. S. A. Smit, G. Wallis, et al., "Estimation of the Worldwide Seroprevalence of Cytomegalovirus: A Systematic Review and Meta-Analysis," *Reviews in Medical Virology* 29, no. 3 (2019): e2034, <https://doi.org/10.1002/rmv.2034>.
2. R. Razonable, "Direct and Indirect Effects of Cytomegalovirus: Can We Prevent Them?," *Enfermedades infecciosas y microbiología clínica* 28, no. 1 (2010): 1–5, <https://doi.org/10.1016/j.eimc.2009.07.008>.
3. B. A. Krishna, M. R. Wills, and J. H. Sinclair, "Advances in the Treatment of Cytomegalovirus," *British Medical Bulletin* 131, no. 1 (2019): 5–17, <https://doi.org/10.1093/bmb/ldz031>.
4. S. Chou, "Advances in the Genotypic Diagnosis of Cytomegalovirus Antiviral Drug Resistance," *Antiviral Research* 176 (2020): 104711, <https://doi.org/10.1016/j.antiviral.2020.104711>.

5. S. Chou, G. I. Marousek, A. E. Senters, M. G. Davis, and K. K. Biron, "Mutations In the Human Cytomegalovirus UL27 Gene That Confer Resistance to Maribavir," *Journal of Virology* 78, no. 13 (2004): 7124–7130, <https://doi.org/10.1128/JVI.78.13.7124-7130.2004>.
6. G. A. Perchetti, M. A. Biernacki, H. Xie, et al., "Cytomegalovirus Breakthrough and Resistance During Letermovir Prophylaxis," *Bone Marrow Transplantation* 58, no. 4 (2023): 430–436, <https://doi.org/10.1038/s41409-023-01920-w>.
7. R. Hall Sedlak, J. Castor, S. M. Butler-Wu, et al., "Rapid Detection of Human Cytomegalovirus UL97 and UL54 Mutations Directly From Patient Samples," *Journal of Clinical Microbiology* 51, no. 7 (2013): 2354–2359, <https://doi.org/10.1128/JCM.00611-13>.
8. E. M. Bunnik and K. G. Le Roch, "An Introduction to Functional Genomics and Systems Biology," *Advances in Wound Care* 2, no. 9 (2013): 490–498, <https://doi.org/10.1089/wound.2012.0379>.
9. E. Paradowska, M. Studzińska, P. Suski, et al., "Human Cytomegalovirus UL55, UL144, and US28 Genotype Distribution in Infants Infected Congenitally or Postnatally: Cmv gB, UL144, and US28 Genotypes in Infants," *Journal of Medical Virology* 87, no. 10 (2015): 1737–1748, <https://doi.org/10.1002/jmv.24222>.
10. C. Kent, A. D. Smith, J. Tyson, et al. PrimalScheme: Open-Source Community Resources for Low-Cost Viral Genome Sequencing. Published online December 22, 2024, <https://doi.org/10.1101/2024.12.20.629611>.
11. A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data," *Bioinformatics* 30, no. 15 (2014): 2114–2120, <https://doi.org/10.1093/bioinformatics/btu170>.
12. B. Langmead and S. L. Salzberg, "Fast Gapped-Read Alignment With Bowtie 2," *Nature Methods* 9, no. 4 (2012): 357–359, <https://doi.org/10.1038/nmeth.1923>.
13. A. Pribelski, D. Antipov, D. Meleshko, A. Lapidus, and A. Korobeynikov, "Using SPAdes De Novo Assembler," *Current Protocols in Bioinformatics* 70, no. 1 (2020): e102, <https://doi.org/10.1002/cpbi.102>.
14. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman Basic Local Alignment Search Tool.
15. H. Li Minimap2: Pairwise Alignment for Nucleotide Sequences. Birol I., ed. *Bioinformatics*. 2018;34(18):3094–3100, <https://doi.org/10.1093/bioinformatics/bty191>.
16. P. Danecek, J. K. Bonfield, J. Liddle, et al., "Twelve Years of SAM-tools and BCFtools," *GigaScience* 10, no. 2 (2021): giab008, <https://doi.org/10.1093/gigascience/giab008>.
17. A. Wilm, P. P. K. Aw, D. Bertrand, et al., "LoFreq: A Sequence-Quality Aware, Ultra-Sensitive Variant Caller for Uncovering Cell-Population Heterogeneity From High-Throughput Sequencing Datasets," *Nucleic Acids Research* 40, no. 22 (2012): 11189–11201, <https://doi.org/10.1093/nar/gks918>.
18. I. Prats-Méndez minMutFinder. Published online September 17, 2024, <https://github.com/ValldHebron-Bioinformatics/minMutFinder>.
19. V. Tilloy, D. Díaz-González, L. Laplace, et al., "Comprehensive Herpesviruses Antiviral Drug Resistance Mutation Database (CHARMD)," *Antiviral Research* 231 (2024): 106016, <https://doi.org/10.1016/j.antiviral.2024.106016>.
20. H. Thorvaldsdottir, J. T. Robinson, and J. P. Mesirov, "Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration," *Briefings in Bioinformatics* 14, no. 2 (2013): 178–192, <https://doi.org/10.1093/bib/bbs017>.
21. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods," *Molecular Biology and Evolution* 28, no. 10 (2011): 2731–2739, <https://doi.org/10.1093/molbev/msr121>.

22. H. Gourel, O. Karlsson-Lindsjö, J. Hayer, and E. Bongcam-Rudloff Simulating Illumina Metagenomic Data With InSilicoSeq. Hancock J., ed. *Bioinformatics*. 2019;35(3):521–522, <https://doi.org/10.1093/bioinformatics/bty630>.
23. M. A. Mallory, W. C. Hymas, K. E. Simmon, et al., “Development and Validation of a Next-Generation Sequencing Assay With Open-Access Analysis Software for Detecting Resistance-Associated Mutations in CMV,” *Journal of Clinical Microbiology* 61 (2023): e00829, <https://doi.org/10.1128/jcm.00829-23>.
24. J. Xu, “Human Cytomegalovirus Envelope Glycoprotein B, H and N Polymorphisms Among Infants of Shanghai Area in China.”
25. X. Pang, A. Humar, and J. K. Preiksaitis, “Concurrent Genotyping and Quantitation of Cytomegalovirus gB Genotypes in Solid-Organ-Transplant Recipients by Use of a Real-Time PCR Assay,” *Journal of Clinical Microbiology* 46 (2008): 4004–4010.
26. A. K. Le Page, M. M. Jager, J. M. Iwasenko, G. M. Scott, S. Alain, and W. D. Rawlinson, “Clinical Aspects of Cytomegalovirus Antiviral Resistance in Solid Organ Transplant Recipients,” *Clinical Infectious Diseases* 56, no. 7 (2013): 1018–1029, <https://doi.org/10.1093/cid/cis1035>.
27. F. El Chaer, D. P. Shah, and R. F. Chemaly, “How I Treat Resistant Cytomegalovirus Infection in Hematopoietic Cell Transplantation Recipients,” *Blood* 128, no. 23 (2016): 2624–2636, <https://doi.org/10.1182/blood-2016-06-688432>.
28. C. S. Walti, N. Khanna, R. K. Avery, and I. Helanterä, “New Treatment Options for Refractory/Resistant CMV Infection,” *Transplant International* 36 (2023): 11785, <https://doi.org/10.3389/ti.2023.11785>.
29. S. B. Kleiboeker, “Prevalence of Cytomegalovirus Antiviral Drug Resistance in Transplant Recipients,” *Antiviral Research* 215 (2023): 105623, <https://doi.org/10.1016/j.antiviral.2023.105623>.
30. W. L. Drew, C. V. Paya, and V. Emery, “Cytomegalovirus (CMV) Resistance to Antivirals,” *American Journal of Transplantation* 1, no. 4 (2001): 307–312, <https://doi.org/10.1034/j.1600-6143.2001.10403.x>.
31. I. Grgic and L. Gorenec, “Human Cytomegalovirus (HCMV) Genetic Diversity, Drug Resistance Testing and Prevalence of the Resistance Mutations: A Literature Review,” *Tropical Medicine and Infectious Disease* 9, no. 2 (2024): 49, <https://doi.org/10.3390/tropicalmed9020049>.
32. C. Gilbert and G. Boivin, “Human Cytomegalovirus Resistance to Antiviral Drugs,” *Antimicrobial Agents and Chemotherapy* 49, no. 3 (2005): 873–883, <https://doi.org/10.1128/AAC.49.3.873-883.2005>.
33. S. Kleiboeker, J. Nutt, B. Schindel, J. Dannehl, and J. Hester, “Cytomegalovirus Antiviral Resistance: Characterization of Results From Clinical Specimens,” *Transplant Infectious Disease* 16, no. 4 (2014): 561–567, <https://doi.org/10.1111/tid.12241>.
34. R. López-Aladid, A. Guiu, M. M. Mosquera, et al., “Improvement in Detecting Cytomegalovirus Drug Resistance Mutations in Solid Organ Transplant Recipients With Suspected Resistance Using Next Generation Sequencing,” *PLOS ONE* 14 (2019): e0219701, <https://doi.org/10.1371/journal.pone.0219701>.
35. P. L. Nagy and H. J. Worman, “Next-Generation Sequencing and Mutational Analysis: Implications for Genes Encoding LINC Complex Proteins.” in *Methods in Molecular Biology*, eds. G. G. Gundersen and H. J. Worman. Springer New York, 2018 1840, 321–336, https://doi.org/10.1007/978-1-4939-8691-0_22.
36. Q. Dahui, “Next-Generation Sequencing and Its Clinical Application,” *Cancer Biology & Medicine* 16, no. 1 (2019): 4–10, <https://doi.org/10.20892/j.issn.2095-3941.2018.0055>.
37. C. J. Houldcroft, M. A. Beale, and J. Breuer, “Clinical and Biological Insights From Viral Genome Sequencing,” *Nature Reviews Microbiology* 15, no. 3 (2017): 183–192, <https://doi.org/10.1038/nrmicro.2016.182>.
38. V. C. Emery and P. D. Griffiths, “Prediction of Cytomegalovirus Load and Resistance Patterns After Antiviral Chemotherapy,”

Proceedings of the National Academy of Sciences 97, no. 14 (2000): 8039–8044, <https://doi.org/10.1073/pnas.140123497>.

39. C. C. Van Leer Buter, D. W. K. de Voogd, H. Blokzijl, et al., “Antiviral-Resistant Cytomegalovirus Infections in Solid Organ Transplantation in the Netherlands,” *Journal of Antimicrobial Chemotherapy* 74, no. 8 (2019): 2370–2376, <https://doi.org/10.1093/jac/dkz196>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Sup Fig 1. Supplementary Table 1: Results of the quality metrics obtained in the sequencing of the dilution bank for each gene. **Supplementary Table 2:** Quality metrics results obtained in the sequencing of the five current samples.