Deciphering the Potential Coding of Human Cytomegalovirus: New Predicted Transmembrane Proteome

Downloaded from: https://research.chalmers.se, 2026-10-14 00:41 UTC

Citation for the original published paper (version of record):

Mancebo, F., Parras Moltó, M., García-Ríos, E. et al (2022). Deciphering the Potential Coding of Human Cytomegalovirus: New Predicted Transmembrane Proteome. International Journal of Molecular Sciences, 23(5). http://dx.doi.org/10.3390/ijms23052768

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

International Journal of
Molecular Sciences

MDPI

Article

Deciphering the Potential Coding of Human Cytomegalovirus: New Predicted Transmembrane Proteome

Francisco J. Mancebo [1,†], Marcos Parras-Moltó [2,3,4,†], Estéfani García-Ríos [1,5,*] and Pilar Pérez-Romero [1]

[1] National Center for Microbiology, Instituto de Salud Carlos III Majadahonda, 28222 Majadahonda, Spain; fj.mancebo@isciii.es (F.J.M.); mpperez@isciii.es (P.P.-R.)
[2] Department of Mathematical Sciences, Chalmers University of Technology, 41296 Gothenburg, Sweden; mparmol@gmail.com
[3] Centre for Antibiotic Resistance Research (CARe), University of Gothenburg, 41122 Gothenburg, Sweden
[4] Igenomix Foundation Research Department—INCLIVA, 46980 Valencia, Spain
[5] Department of Science, Universidad Internacional de Valencia—VIU, Calle Pintor Sorolla 21, 46002 Valencia, Spain
* Correspondence: egarcia@isciii.es; Tel.: +34-91-822-3168
† These authors contributed equally to the work.

Abstract: CMV is a major cause of morbidity and mortality in immunocompromised individuals that will benefit from the availability of a vaccine. Despite the efforts made during the last decade, no CMV vaccine is available. An ideal CMV vaccine should elicit a broad immune response against multiple viral antigens including proteins involved in virus-cell interaction and entry. However, the therapeutic use of neutralizing antibodies targeting glycoproteins involved in viral entry achieved only partial protection against infection. In this scenario, a better understanding of the CMV proteome potentially involved in viral entry may provide novel candidates to include in new potential vaccine design. In this study, we aimed to explore the CMV genome to identify proteins with putative transmembrane domains to identify new potential viral envelope proteins. We have performed in silico analysis using the genome sequences of nine different CMV strains to predict the transmembrane domains of the encoded proteins. We have identified 77 proteins with transmembrane domains, 39 of which were present in all the strains and were highly conserved. Among the core proteins, 17 of them such as UL10, UL139 or US33A have no ascribed function and may be good candidates for further mechanistic studies.

Keywords: cytomegalovirus; pangenome; proteome; transmembrane

check for updates

Citation: Mancebo, F.J.; Parras-Moltó, M.; García-Ríos, E.; Pérez-Romero, P. Deciphering the Potential Coding of Human Cytomegalovirus: New Predicted Transmembrane Proteome. *Int. J. Mol. Sci.* **2022**, *23*, 2768. https://doi.org/10.3390/ijms23052768

Academic Editor: Paolo Iadarola

Received: 3 January 2022
Accepted: 26 February 2022
Published: 2 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Human cytomegalovirus (CMV) is a large envelope worldwide prevalent betaherpesvirus, ranging from 45% to 100% in the general population based on socio-economic factors [1,2]. Although CMV generally causes asymptomatic infections in immunocompetent individuals, it is a major cause of morbidity and mortality in immunocompromised individuals such as organ transplant recipients, AIDS, and with congenital infection [3–7].

The CMV genome (236 kb) consists of a unique long (UL) and a unique short (US) region flanked by inverted repeats. CMV gene expression occurs with the expression of immediate-early genes followed by early, early-late and late transcripts [8]. In addition to the 165 canonical ORFs [9,10], CMV genome encodes for other alternative transcripts in addition to have non-canonical translation initiation sites [11,12]. Furthermore, CMV encodes for a large number of genes, many of them with unknown functions that may probably be involved in key processes during host-cell interaction [13].

CMV is able to infect a high number of cell types including fibroblasts, endothelial cells, epithelial cells and myeloid lineage cells, among others [14,15]. CMV is a highly complex virus with multiple proteins embedded in the viral envelope, with at least four

distinct types of covalently linked glycoprotein complexes required for CMV infectivity including gCI complex (gB dimer), gCII complex (gM, gN), gCIII complex (trimer gH, gL, gO) and pentameric complex (gH/gL/UL128-131) [16,17]. The therapeutic use of neutralizing antibodies, targeting glycoproteins mediating viral entry, have demonstrated to only achieve partial protection against CMV infection [18–23]. One possible explanation if that other proteins may be involved in viral entry that might be also necessary to target in order to elicit a complete protection against infection. An ideal vaccine against CMV infection should elicit a broad immune response, including both neutralizing antibody and T-cell response, against multiple viral antigens including proteins involved in virus-cell interaction and entry [24,25], which may increase efficacy compared with the previously tested vaccines [1,26–28]. In fact, despite the efforts made during the last decade, no CMV vaccine is still available [18,27,29,30]. Thus, understanding the complete repertoire of CMV proteins involved in cell entry may also help to determine the neutralizing response necessary to block infection and may provide novel candidates that could be included in new vaccines design.

In this study, we aimed to explore the CMV genome to identify proteins with putative transmembrane domains to identify new potential viral envelope proteins that may be involved in virus-cell interaction during infection and may therefore be potential targets for neutralizing antibodies for the development of novel therapeutics and preventive measures targeting viral entry and cell-to-cell spread. In order to do that, we have performed in silico analysis using the genome sequence of nine different CMV strains, to identify proteins with predicted transmembrane domains. To further characterize the identified proteins, an exhaustive systematic review of the literature and a sequence homology analysis with known proteins from other organisms were performed. Our work highlights the need to explore new experimental and computational approaches to identify and characterize the CMV proteome.

2. Results

2.1. Identification of Putative Transmembrane Proteins

To determine the transmembrane regions of the proteins encoded by the CMV genome, the genome of nine different CMV strains including both clinical and laboratory strains (Table 1) were analyzed using three different bioinformatic methods: Phobius, PureseqTM and TMHMM. A description of the methodological approach used is represented in Figure 1. CMV is known to accumulate mutations quite rapidly in cell culture during cell passaging [31]. In order to test to what level these nine selected CMV strains are representative of the 335 available CMV genomes in GenBank, the 56190 ORFs were aligned with the ORFs in our CMV dataset. We obtained 100% median percentage identity and breadth coverage (overlapping distance), representing 99.95% of the total ORFs from the Human betaherpesvirus 5 in the NCBI database (Figure S1).

Based on the first analysis, we identified 94 proteins with potential transmembrane domains (Figures 2 and S2). Seventeen of them were not considered for further analysis (Figure S2) because of the following reasons. Proteins UL74, UL115, UL47, UL49, UL76, US22, UL77, UL105, UL122 and UL89 were previously described not to be part of membrane structures [32–41]. UL47, UL49, UL76 and US22 proteins are known to be part of the tegument, UL77 is located in the capsid; UL105, UL122 and UL89 are found in the nucleus of the host cells [32–39]. In addition, UL4, UL22A and UL116, which were predicted to have one transmembrane domain, were discarded because the transmembrane domain corresponded to the sequence of signal peptide [42–44]. In addition, RL13TRL14 and US33 (TB40-E_UNC strain) or ORFL27C and ORFL49W.IORF1 (AD169-BAC20) proteins were only found in one of the CMV studied strains and were not considered for further analysis.

Table 1. Characteristics of CMV strains used in this study and their corresponding accession number at nucleotide database.

CMV Strains	Isolation Source	Number of Culture Passages	Accession Number
AD169	Adenoids of a 7-year-old girl	Many times in human fibroblasts	FJ527563.1
Towne	Urine of a 2-month-old infant with microcephaly and hepatosplenomegaly	Many times in human fibroblasts	FJ616285.1
Toledo	Urine from a congenitally infected infant	Several times in human fibroblasts	GU937742.2
TR	Vitreous humor from eye of HIV-positive male	Several times in human fibroblasts	KF021605.1
VR7863	Urine samples of a congenitally infected neonate and cultured in endothelial and epithelial cells	Cultured in endothelial and epithelial cells	KX544838.1
TB40-E_UNC	Throat swab of a bone marrow transplant patient	Cultured adapted	KX544839.1
HANSCTR4	Blood from stem cell transplant recipient (D-R+)	Sequenced directly from clinical material via target enrichment	KY123653.1
AD169-BAC20	-	-	MN920393.1
Merlin	Urine from a congenitally infected child	3 times in human fibroblasts	NC006273.2

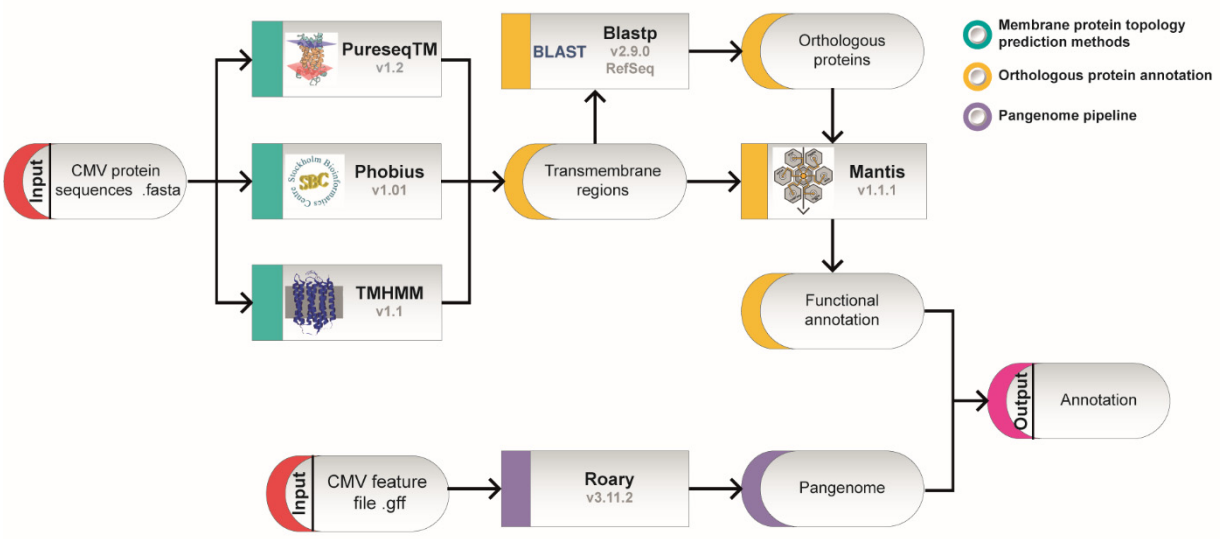

Figure 1. Schematic representation of the applied workflow. Fasta format protein sequences from nine CMV genomes were analyzed in parallel to predict transmembrane domains and to create an entire set of genes from all strains (pangenome). Transmembrane topology was studied following three different approaches: PureseqTM, Phobius and TMHMM, under default parameters. Predicted transmembrane proteins were compared with orthologous proteins identified by BLAST with the whole Mantis database for the prediction of functional annotation. Proteins that were common to all nine genome datasets formed the core protein set, and functions were annotated accordingly for each transmembrane protein.

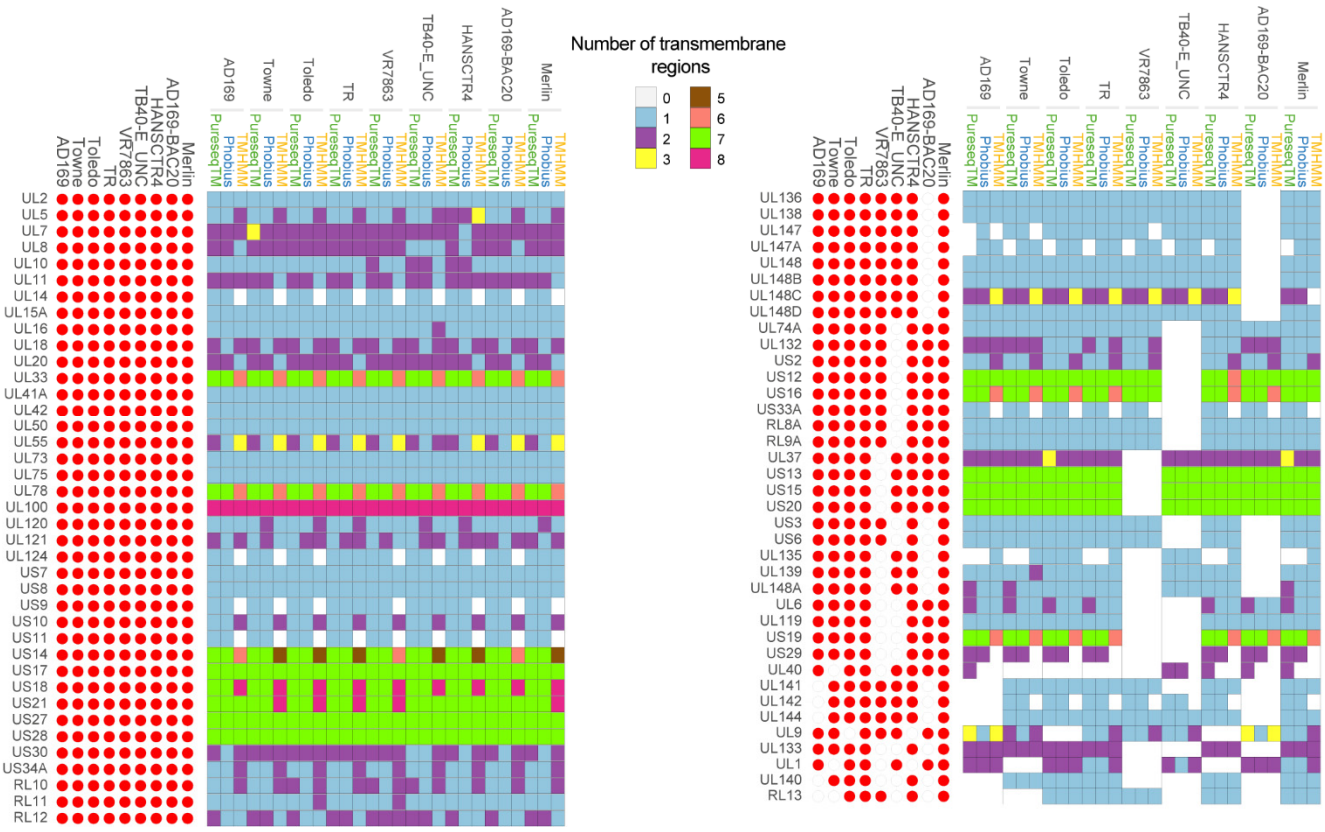

Figure 2. Predicted transmembrane proteins for the studied CMV genomes. Proteins with at least one predicted transmembrane domain with one of three tested methods were annotated as transmembrane proteins. The number of transmembrane domains for each protein was represented using the indicated chromatic scale ranging from zero to eight regions for each of the three methods used (PureseqTM, Phobius and TMHMM). For each strain, the presence of the gene was represented with the filled red circles and absent genes with empty red circles.

For further characterization of the 77 remaining proteins with predicted transmembrane (TM) domains, a systematic review was performed to search for any previous published information. For each of the proteins, information on the location, the ascribed function and the number of predicted TM domains is included in Table S1. A graphical representation of the number of predicted TM regions found for each ORF, in each of the nine strains with the indicated bioinformatics tool is shown in Figure 2. Of the 77 proteins analyzed, 33 (43%) proteins only had one TM domain, 23 (29.87%) had from 1 to 2 TM domains, 6 (7.7%) exhibited 1-3, while 15 proteins (19.48%) had from 5 to 8 TM domains. None of the analyzed proteins had four TM domains.

Nineteen out of the 77 proteins (UL2, UL6, UL9, UL14, UL15A, UL74A, UL120, UL121, UL140, UL148C, UL148D, US13, US15, US19, US29, US30, US33A, RL8A, RL9A and RL10) have no previously described function, 13 (UL1, UL5, UL8, UL10, UL20, UL42, UL78, UL124, UL139, UL147, US34A, RL12 and RL13) have been partially studied, 1 (UL41A) has previously been shown not to code for a protein [10] and the other 43 proteins have a previously described function (Table 2).

Table 2. CMV predicted transmembrane proteins indicating the cellular localization based on biotool Uniprop, the ascribed functions based on a bibliographic search and the number of predicted domains using the three different tools. (*) indicates unknown or non-verified function.

Gene	Localization	Function	Number of TM Domains	References
UL1 *	VM	Unknown. pUL1 could modulate CMV host cell tropism.	1–2	[45]
UL2 *	HM	Unknown.	1	-
UL5 *	V	Unknown. It is suggested to be involved in efficient viral assembly, propagation and replication.	1–3	[46,47]
UL6 *	HM	Unknown.	1–2	-
UL7		UL7 is involved in immunomodulation.	2–3	[48–50]
UL8 *	HM	UL8 decreases the release of a large number of pro-inflammatory factors later after infection of THP-1 myeloid cells. UL8 may exert an immunosuppressive role key for CMV survival in the host.	1–2	[51]
UL9 *	HM	Unknown function. Its deletion mutation cause enhanced growth in HFFs cells.	1–3	[9]
UL10 *	M	Unknown. Potential role in immunomodulation.	1–2	[52]
UL11	HM, ERM	pUL11 interacts with CD45 phosphatase on T cells, inducing the IL-10 secretion.	1–2	[53]
UL14 *	HM	Unknown.	0–1	-
UL15A *	HM	Unknown.	1	-
UL16	HM	Immunoevasion and inhibition of the activation of NK cells.	1	[54]
UL18	HM	Immunomodulation and immunoevasion.	1–2	[55]
UL20 *	ERM	Unknown. UL20 could be destined to sequester cellular proteinases not known to date for degradation in lysosomes.	1–2	[56]
UL33	HM	UL33 has homology with GPCR which activates different ligand-independent signalling pathways and also involved in virus dissemination.	6–7	[57–59]
UL37	ERM, GM, MM	Viral replication.	2–3	[60,61]
UL40	HM	Immunomodulation.	0–2	[62]
UL41A *	VM	Unknown. UL41A not to code for proteins.	1	[10]
UL42 *	HM, C	Unknown. Potential role in immunoevasion.	1	[63,64]
UL50	HNM	Assembly, maturation and egress of virions.	1	[65]
UL55	VM, HM, GM	Glycoprotein B participates in viral entry.	1–3	[66]
UL73	VM, HM, GM	Glycoprotein N is involved in the binding of the virus to the host cell, viral spread and virion morphogenesis.	1	[67]
UL74A *	VM	Unknown	1	-

Table 2. *Cont.*

Gene	Localization	Function	Number of TM Domains	References
UL75	HM, VM	Glycoprotein H participates in viral entry. It is part of the trimeric and pentameric complexes.	1	[68]
UL78 *	HM, ERM	Unknown. UL78 is a G protein-coupled receptor.	6–7	[69,70]
UL100	HM, VM	Envelope glycoprotein M participates in viral entry.	8	[71,72]
UL119	VM	Immunoevasion.	1	[73]
UL120 *	HM	Unknown.	1–2	-
UL121 *	HM	Unknown.	1–2	-
UL124 *	HM	Potential role in latency.	0–1	[74]
UL132	VM	Essential for CMV assembly compartment formation and the efficient production of infectious particles.	1–2	[75]
UL133	GM	UL133 forms a complex with UL138 and UL136. It is involved in the establishment of CMV latency.	2	[76]
UL135	HM, GM	Immunomodulation. Post entry Tropism in Endothelial Cells.	0–1	[77,78]
UL136	HM	Replication, latency, and dissemination. Post entry Tropism in Endothelial Cells.	1	[76,78–80]
UL138	GM	Latency and DNA replication.	1	[81,82]
UL139 *	HM	Unknown. Potential role in immunomodulation.	1–2	[83]
UL140 *	HM	Unknown.	1	-
UL141	ERM	Immunomodulation and DNA replication.	1	[84–86]
UL142	ERM	Immunomodulation.	0–1	[87]
UL144	HM	Inhibition of T-cell activation and latency.	1	[88,89]
UL147 *	EXR	Unknown. Potential role in immunomodulation.	0–1	[90]
UL147A	HM	Immunomodulation.	0–1	[91]
UL148	ERM	Viral ER-resident glycoprotein that interacts with UL116 promoting the incorporation of gH/gL complexes into virions.	1	[92]
UL148A	HM	Immunoevasion of NK cells.	1–2	[93]
UL148B *	HM	Unknown.	1	[94]
UL148C *	HM	Unknown.	0–3	[94]
UL148D *	HM	Unknown.	1	[94]
US2	ERM	Immunomodulation.	1–2	[95]
US3	ERM	Immunoevasion.	1	[96]
US6	ERM	Immunomodulation.	1	[97]
US7	ERM	Immunoevasion.	1	[98]
US8	ERM, GM	Immunomodulation.	1	[98]
US9	ERM, GM, CK	Glycoprotein US9 is an antagonist of IFN signalling to persistently evade host innate antiviral responses.	0–1	[99]

Table 2. *Cont.*

Gene	Localization	Function	Number of TM Domains	References
US10	ERM	Inhibition of the host immune response.	1–2	[100]
US11	ERM	Inhibition of the host immune response.	0–1	[101]
US12	HM	Inmunomodulation of NK cells activation.	6–7	[102]
US13 *	HM	Unknown.	7	-
US14	HM	Inmunomodulation of NK cells activation. Potential role in virions maturation and egress.	5–7	[102,103]
US15 *	HM	Unknown.	7	-
US16	HM, C	Tropism in endothelial and epithelial cells.	6–7	[104]
US17	HM	Immunomodulation.	7	[105]
US18	HM.	Immunoevasion of NK cell.	7–8	[106]
US19 *	HM	Unknown. Its delection affect NK cell activation.	6–7	[102]
US20	M	Inhibition NK cell activation. Also participates in the viral replication process in endothelial cells.	7	[106,107]
US21	HM	Viroporin that modulates calcium homeostasis and protects cells against apoptosis.	7–8	[108]
US27	V, HM	Immunomodulation. Also is required for efficient viral spread by the extracellular route.	7	[109–111]
US28	HM	Immunomodulation. Lytic and latent CMV infection. Possible role in regulation of the actin cytoskeleton or cytoskeletal remodelling.	7	[112,113]
US29 *	HM	Unknown.	0–2	-
US30 *	HM	Unknown.	1–2	-
US33A *	-	Unknown.	0–1	[114]
US34A *	HM.	Unknown. Potential target of SUMO complex.	1–2	[115]
RL8A *	HM	Unknown.	1	-
RL9A *	HM	Unknown.	1	-
RL10 *	VM	Unknown.	1–2	-
RL11	HM	Immunomodulation. RL11 is a type I transmembrane glycoproteins which bind immunoglobulin G Fc. I	1–2	[116]
RL12 *	VM	Unknown. RL12 is a Fc binding protein.	1–2	[117]
RL13 *	VM	Unknown. Potential role in replication, immunoevasión and viral spread by cell-free or cell-to-cell mechanisms.	1	[118–120]

* indicates unknown or non-verified function. CK: Cytoskeleton C: Cytoplasm, ERM: Host endoplasmic reticulum membrane, EXR: Extracellular region, GM: Golgi reticulum membrane, HM: host membrane, HMN: Host nucleus membrane, M: Membrane, MM: Mitochondrion membrane, V: Virion, VM: Virion membrane.

The number of predicted domains differed in some of the studied proteins when using different methods. The results obtained TMHMM method were the most divergent of the

three tested methods. On the contrary, a group of proteins encoded by the genes UL33, UL78 and the genes from the unique short (US) region US12-US21, US27 and US28 proteins were predicted to have more than five TM regions by all three methods. In fact, TM regions of these genes, such as the members of US12 family and the proteins with homology to the chemokine receptor family of G protein-coupled receptors (GPCRs): US27 and US28, have been previously described supporting our results [108,109,121].

A validation experiment was carried out using as an example UL2 and UL124, two of the identified proteins with unknown function. The ORF encoding for these two proteins were cloned into a eukaryotic expression plasmid that included a Myc tag sequence in the 5´end of the clone products. After transfecting the HEK 293T mammalian cell line, plasma membrane proteins were extracted and the cytoplasmic (C) and plasma membrane (PM) protein fractions were tested by Western Blot using an anti Myc antibody. As shown in Figure S3A, both UL4 and UL124 proteins were only detected in the PM fractions confirming their location in the membrane. A loading control using the stain free technology is shown in Figure S3B.

2.2. Homology Analysis of thePredicted Transmembrane Proteins

In addition to the exhaustive systematic review of the literature, further analysis of sequence homology with known proteins from other organisms was performed using Mantis software (Table S1). Based on this analysis, we found homologies for two of the proteins with unknown function. UL139 had some level of homology (*e*-value = 5.1×10^{-28}) with proteins involved in cell adhesion, while UL15A had some homology (*e* value = 1.53×10^{-4}) with a biotin permease protein. UL15A ORF was identified in all 9 CMV strains analyzed, while UL139 that was only present in the TR strain.

In addition, Mantis analysis shed an association of UL1 with a carcinoembryonic antigen-related cell adhesion molecule, which is a cell adhesion receptor of the immunoglobulin-like superfamily. UL78 was also identified by Mantis as seven transmembrane receptor from the rhodopsin family. UL147 has been proposed by Mantis to be involved in immune response and chemokine activity and US33A seems to have a von Willebrand A (VWA) domain. However, US33A was present exclusively in Towne, Toledo, TR and VR7863 strains.

2.3. Sequence Differences among Strains

The analysis of the generated pangenome (Figures 2 and 3), revealed wide differences among strains [13]. Clinical isolate VR7863 and TB40-E_UNC strains lacked a large number of genes compared to the other strains. The VR7863 strain lacked some genes with unknown functions such as UL1, UL6, UL139, UL140, US13, US15, US19 and US29 and other genes involved in immunoevasion, DNA packaging, latency or viral replication such as UL37, UL40, UL119, UL133, UL135, UL148A and US20 [39,60–62,73,76,77,93,106]. The TB40-E_UNC strain lacked some genes with unknown function such as UL6, UL74A, UL140, US12, US19, US29, US33A, RL8A, RL9A and RL13 and other genes involved in host immune response evasion, CMV assembly, tropism, or latency such as UL119, UL132, UL133, US2, US3 and US16 [73,75,95,96,102,104,118,119].

AD169 strain have a deletion of a genomic region that included UL140, UL141, UL142, UL144 genes and RL13 gene (known to have TM domains) [117]. All of them have functions related to the evasion of the host's immune system [84–88]. In addition, the AD169 BACmid (Table 1) widely used for research, lacked several genes encoding TM proteins [117]. Some of them such as UL140, UL141, UL142, UL144 and RL13 were also deleted in AD169 strain. While the AD169 BACmid also lacked other genes such as UL135, UL136, UL138, UL148 and US3-US6 genes, involved in DNA replication, latency, virulence, tropism, evasion of the immune response and other genes such as UL139, UL147, UL148B, UL148C and UL148D with uncharacterized function [77,79,81,92,96,97,117]. The Toledo strain lacked RL13, UL9 and UL128 genes, while the Towne strain lacked RL13, UL1 and UL40 genes. The TR and Merlin strains, widely used in research included all the analyzed genes.

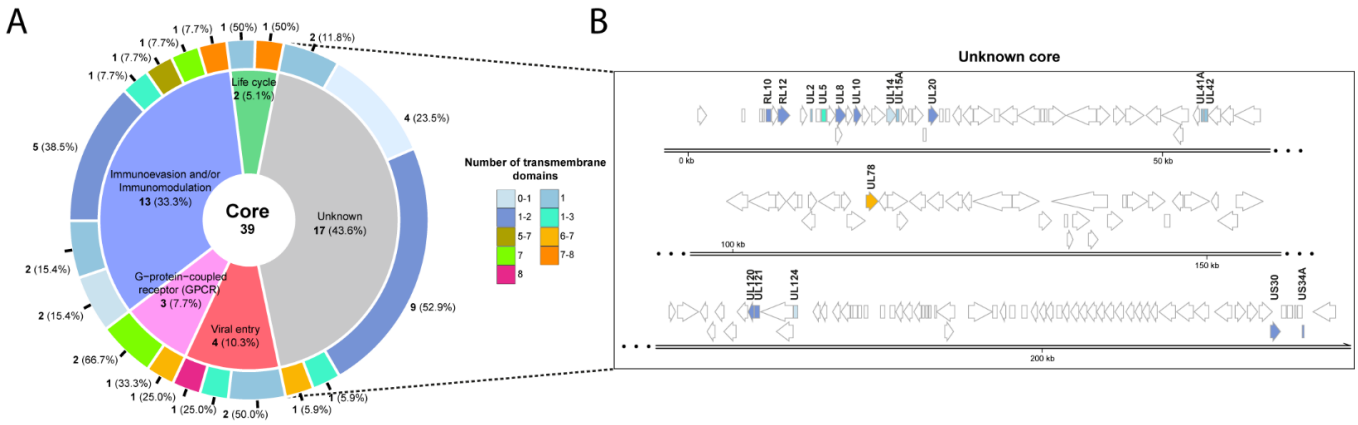

Figure 3. Functional analysis of the 39 core proteins. (**A**) Pie chart of the proteins found in all the studied strains were grouped based on their functions. For each group, the number of predicted transmembrane domains is also indicated. When the number of transmembrane domains predicted was different using the three methods, a range of values is shown. The number of proteins in each section is marked in blue and the percentage between brackets. (**B**) Genomic location of the 17 proteins with non-described function.

2.4. Sequence Similarities among Strains: Core Proteins

Based on the differences among different strains, we searched for those genes with predicted TM domains that were present in all the studied CMV genomes. Of the 77 initially predicted TM proteins, 39 (50.64%) met the criteria and were designated as the core TM proteome (Figures 2 and 3). A representation of the 39 proteins grouped according to their function is shown in Figure 3A. The number of predicted TM domains is also indicated in each group. No association was found between the number of TM domains and the function in each group.

Most of the 39 core proteins were highly conserved among all nine strains with a high percentage of sequence similarity (94.33 ± 7.3), except for the RL12 gene (48.31 ± 24.17, Figures 4 and S4). The similarity matrix for each protein in each indicated strain is shown in Figure S4 and an example of heat map depicting the similarity of the core proteins comparing all the strains with AD169, as a reference strain, is shown in Figure 4. As expected, AD169-derived BAC was almost identical to the AD169 strain. Despite the fact that sequence similarity was overall high (with an average above 94%), the sequences of the core proteins from the Merlin, Towne and TR strains differed the most. In addition, most of the core proteins with unknown function (highlighted in red in Figure 4), tended to have lower similarity values compared with other core proteins with known functions.

Some of the 39 identified core proteins were involved in cell entry (4, 9.52%) and immunomodulation/immunoevasion mechanisms (13, 33.33%). However, a significant percentage of them (17, 43.58%) had no described function (UL2, UL8, UL14, UL15A, UL41A, ULL120, UL121, UL124, US30, RL10 and RL12) or were poorly studied (UL5, UL10, UL20, UL42, UL78 and US34A). Figure 3B shows the location for each of the 18 core proteins with uncharacterized function in the CMV genome.

In addition, in order to test to what extent the core proteins were found among the CMV population, the 335 available genomes in the database were aligned to our set using blastp to all proteins present in our pangenome core. All genomes had representative proteins related to proteins in our set in different proportions with a high number of genomes containing all 39 core proteins, indicating that our pangenome could be extended to all annotated genomes (Figure S1B).

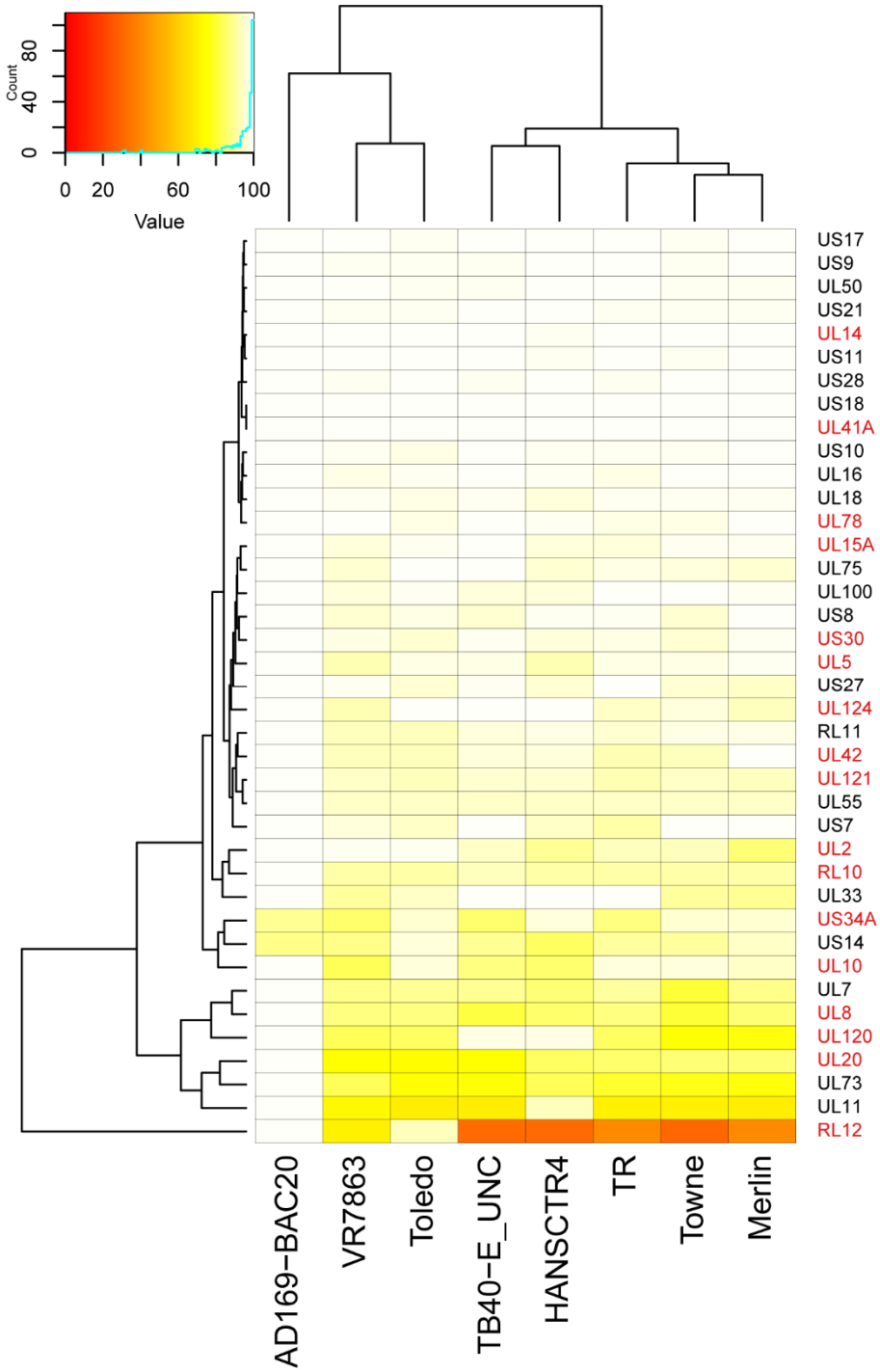

Figure 4. Percentage identity heatmap among the 39 core proteins from all the studied strains using AD169 as a reference strain. Color scales range from red (0% identity) to white (100% identity). Strains and genes were clustered following a hierarchical clustering method (HCL). Genes were clustered from top to bottom of the figure based on their similarity within the indicated strains. Core proteins with unknown functions are highlighted in red.

2.5. CMV Gene Families with Transmembrane Domains

The genes encoded by the CMV genome are grouped into several gene families [122], five of which are important families that include genes with TM domains. Thus, we next analyzed which of the 39 identified core proteins are part of these families.

The RL11 family share the RL11D central domain that includes three conserved residues (one tryptophan and two cysteines) and several potential N-linked glycosylation sites associated with immunomodulatory properties. The RL11 family consists of 14 genes (RL5A, RL6, RL11-13, UL1 and UL4-11 genes), encoding for proteins with transmembrane domains, except for RL5A and RL6 [51], most of them with unknown specific function [123]. Based on our analysis, we identified three of the genes belonging to this family (UL5, UL8 and UL10) that have not been well characterized, although they are suggested to be involved in the viral cycle and immunomodulation mechanisms (Table S1) [42,45,52].

The US12 family includes US12-US21 genes [102]. US12, US14, US18 and US20 had been shown to be involved in the inhibition of natural killer cells [107]. US16 had been shown to interact with UL130 participating in CMV tropism [104]. US21 encodes a viroporin that modulates calcium homeostasis and protects the cells against apoptosis [108]. US18 and US20 were described to play a role in cell tropism mediating viral replication in specific cells while US13, US15 and US19 have no described function or are poorly characterized [102,107].

The GPCR-7 TM family includes four genes coding for proteins with seven predicted TM domains: US27, US28, UL33 and UL78. This family includes G protein-coupled surface receptors with an important role in immunomodulation that transmit an intracellular signal when binding to an extracellular ligand [109,124]. Within this family only UL78 (protein from the unknown core proposed in this work) remains uncharacterized.

UL120 and UL121 are proposed to form a family and both of them were identified in our unknown core analysis. However, little is known about their function and further studies are needed to shed light into their biological relevance [125].

3. Discussion

The lack of knowledge of an important part of the genes encoded by the CMV genome, the variability between laboratory and clinical strains and the complex regulation of the virus over the host immune system, have probably limited the design of new preventive and prophylactic measures against CMV infection [1,122,126]. CMV proteins involved in the interaction with the target cells during entry located in the viral envelope may be considered possible targets to develop new treatments and vaccines against CMV since blocking a combination of these proteins may block the infection in the different target cells [24]. However, of the 165 proteins potentially encoded by the CMV genome, only around 60 proteins have been functionally characterized [12,112,127,128]. In this sense, some authors have reported the association between γ marker, human leucocyte antigens and killer immunoglobulin-like receptors and the natural course of CMV infection [129–132].

We performed the present study using in silico analysis of the genomic sequences of nine CMV strains (representative of the 335 available CMV genomes in GenBank) with three different bioinformatic tools, to identify proteins that have putative TM domains, to identify new potential envelop proteins, that may help to better understand CMV interaction with the cells and with the host immune system.

As a result, we gained knowledge about the performance of the three bioinformatics tools used in this study. While the Phobius and TMHMM tools are well-established methods to study TM regions (using sequence-based approaches that use hidden Markov), PureseqTM is a novel alternative based on a machine learning model algorithm (DeepCNF), which has demonstrated to increase the number of results obtained. These tools have already been used to study membrane proteins in other organisms such as humans, *Plasmodium falciparum*, or *E. coli* [133–135]. Although these tools are quick in silico approaches to analyze the genome of a given organism, some of the results obtained are not always accurate. We obtained some variability of the number of predicted TM domain as a function of the used method. These discordances between methods could be explained due to differences in the algorithm and the threshold values. By default, each applied method that predicts TM regions uses their own algorithms and thresholds, probably detecting slight differences for the same region resulting in different thresholds. The analysis of the

signal peptide was also important to exclude false positives since the signal peptides of type I TM proteins are usually hydrophobic and are often predicted to be TM domains. In addition, the nucleotide sequence variability of the different CMV strains with a significant number of polymorphisms (such as genes gN, and UL21) may also explain this variability [56,136,137]. These results highlight the importance of applying different bioinformatics methods for predicting domains in silico.

Proteins that are evolutionarily related are commonly referred to as homologues and very close homologues often have similar functions [138]. Homology-based methods have been previously used in different scenarios such as the identification of oncogenes from retrovirus proteins; cancer metastasis; herpesviruses; and *Hepadnaviridae* [139] among others. Using a homology-based method, we proposed functions for 23 CMV proteins that were previously characterized, which confirmed the performance of the method. We were also able to propose putative functions to several unknown CMV proteins such as UL1, UL15A, UL139, UL78, UL147, US13 and US33A. UL15A was proposed as a biotin transport system permease protein, while UL139 could be involved in cell adhesion. Furthermore, based on the homologies, UL1 was identified to be involved in the viral-cell adhesion and potentially modulating the immune system of the host. Most of those carcinoembryonic antigen-related cell adhesion molecules are modulators of general cellular processes such as proliferation, motility, apoptosis as well as cell-cell interaction that binds to pathogens enhancing their capacity to colonize the host [140,141]. Our analysis identified UL78 as a member of the rhodopsin family which is large group of evolutionarily-related proteins that are cell surface receptors, detecting molecules outside the cell and activate cellular responses [142]. However, given the functional homology among the other members of the family, it is plausible to think that UL78 exhibits similar functions and may have an immunomodulatory role [113], although its ligand is still unknown. UL147 arises with a potential role in immune response and chemokine activity, likely due to its homology with UL146 which is already characterized [90,143]. Finally, US33A, which is present in Towne, Toledo, TR and VR strains, showed a von Willebrand A (VWA) domain that is well characterized to be involved in cell adhesion with extracellular matrix proteins and integrin receptors [144] and could be likely be involved in signaling.

4. Materials and Methods

4.1. Transmembrane Region Analysis

The nine CMV genome sequences AD169, Towne, Toledo, TR, VR7863, TB40-E_UNC, HANSCTR4, AD169-BAC20 and Merlin (Table 1) were available at the Nucleotide database. In order to test to what level these nine selected CMV strains are representative of the available CMV genomes in GenBank and the ORF from the 335 available CMV genomes containing 56217 ORFs (annotated in the NCBI database as human betaherpesvirus 5 complete genomes) were downloaded. Sequences were aligned using blastp with default parameters. Additionally, the complete set of ORFs coding for CMV proteins was aligned with the ORFs of the protein core in the pangenome. Blast results were filtered for sequence percentage identity $\geq$75% and breadth coverage $\geq$90%.

These nine genome sequences were analyzed to predict transmembrane domains within the open reading frames (ORFs). The analysis was carried out using the default parameters of three different bioinformatics approaches: PureseqTM (v1.2) [145], Phobius (v1.01, Stockholm Bioinformatics Center, Sweden) [146] and TMHMM (DTU Health Tech, Lyngby, Denmark) (v1.1) [147]. Phobius and TMHMM are based on sequence methods as hidden Markov model (HMM) approach, while PureseqTM adds an extra layer of prediction based on deep learning. All methods were expected to show similar output with differences in the predictive threshold for the same set of analyzed ORFs. The tool SignalP6.0 (DTU Health Tech, Lyngby, Denmark) was used for signal peptide analysis [148].

4.2. Functional Annotation of the Transmembrane ORFs

Predicted TM proteins were analyzed using Mantis protein function annotation v1.1.1 [149] which is a stand-alone tool that uses HMMER or Diamond to match sequences against multiple reference datasets to produce high-quality consensus driven protein annotations, under default parameters. This analysis uses the information from the following different available protein function databases: KOfam [150], Pfam [151], eggnog [152], NCBI protein family models (NPFM) [153], and TIGRfams [154] and sets a consensus result. In parallel, we searched for sequence homology in the RefSeq database using blast with Blastp (v2.9.0) [155] (*e*-value < 10^{-3}) looking for orthologous proteins. Orthologous proteins are proteins found in other species that maintain the same or close functionality as the studied protein. Proteins with unknown functions could be related to their orthologous protein function this way.

Additional systematic review was performed to retrieve related articles published on the PubMed database website (https://pubmed.ncbi.nlm.nih.gov/, accessed on 15 December 2021) until July 2021. For each of the genes studied, articles were identified using the following search terms: "HHV-5" AND or "CMV".

4.3. Viral Pangenome Construction

A pangenome was created for all nine sequences with Roary (v3.11.2, Wellcome Genome Campus, Hinxton, UK) [156] under default parameters. Pangenome representation and TM information from all genomes were plotted using ggplot2 (v3.3.3) [157] and ggcorrplot (v0.1.3) R packages.

4.4. Similarity Analysis

Core proteins for all nine CMV strains were aligned using Clustal Omega (v1.2.1) [158] with -percent-id and –full parameters to obtain identity matrices. Thirty-nine matrices were obtained and results were plotted as individual heatmaps using heatmap.2 from gplots package (v3.1.1). In parallel, a percentage identity heatmap was plotted for the comparison between AD169 and the other studied strains 39 core proteins, following the same parameters.

4.5. Protein Analysis

UL2 and UL124 ORFs were amplified by PCR using CMV BADrUL131-Y4 BAC as a template, with specific primers (Table S2), and the Phusion DNA polymerase (Thermo Scientific). PCR products and the pcDNATM3.1/myc-His (-) (5.5 kb) vector (Invitrogen) were digested with the appropriate restriction enzyme (FastDigest enzymes, Thermo Scientific, Waltham, MA, USA), ligated (Ligase, Thermo Scientific, Waltham, MA, USA) and transformed into the XL10 Gold chemically competent *E. coli* cells. The constructed plasmids were verified by sequencing (Table S2). The pcDNATM3.1/myc-His (Thermo Scientific, Waltham, MA, USA) constructs containing the UL2 and UL124 ORFs were transfected into the HEK 293T human cell line, using the $CaCl_2$ transfection method. Briefly, the day before transfection, 1,500,000 HEK 293T cells were seeded in a 10-cm plate and the next day transfections were carried out with approximately 70–80% of cell confluence. Four hours before the transfection, fresh medium was added to cells. For the transfection, 750 µL of $CaCl_2$ were mixed with 40 µg of the DNA construct. 750 µL of HBS saline buffer (140 mM NaCl, 1.5 mM Na_2HPO_4, 50 mM HEPES) were added and the mixture was incubated for 15 min at room temperature. Subsequently, the mixture was added dropwise to each plate containing the 293T cell monolayer, and gently mixed. Transfected cells were incubated during 48 h at 37 °C with 5% CO_2. Twenty-four hours post-transfection, fresh medium was added to cells. Transfected cells were pelleted and treated with Mem-PER^{TM}Plus Membrane Protein Extraction Kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. Cytoplasmic and plasma membrane protein fractions were obtained and quantified by Bradford. Ten µg of protein lysates were separated on a gradient 4–20% pre-cast SDS gel (BioRad, Hercules, CA, USA), transferred to 0.2 µm

nitrocellulose membranes. Detection of the expressed proteins was performed using an anti-Myc monoclonal antibody (MA1-21316, Invitrogen, Waltham, MA, USA) at a 1:1000 dilution in blocking buffer (1X PBS + 0.1% Tween 20 + 5% skim milk) and incubated at 4 °C overnight, with a secondary horseradish peroxidase (HRP)-labeled anti-mouse IgG antibody (diluted 1:2000; 05/2019, Cell Signaling, Danvers, MA, USA). Stain free technology (BioRad, Hercules, CA, USA) of the acrylamide gel was used in parallel as a loading control.

5. Conclusions

Our results highlight the utility of using these bioinformatic tools to gain knowledge of previously uncharacterized proteins that may be useful to select potential targets of the immune system. Our work also suggests that differences among the strains may be crucial for CMV tropism, replication, latency or the evasion of the host immune response [13]. Among the 77 identified proteins with predicted TM domains, only 39 (designated as a core TM proteome) were shared by all analyzed strains most of which were highly conserved which may have potential clinical relevance for the design of new therapeutics and preventives measures. This group of proteins was highly conserved in all nine strains analyzed, which may have solved the limitation of the variability among laboratory strains and clinical isolates, facilitating the extrapolation of the results.

In addition, the study of the role of the previously uncharacterized proteins may provide novel candidates for new preventive or prophylactic measures against CMV infection. Furthermore, if their location in the viral envelope is confirmed, they could be used as viral antigens for developing a vaccine or monoclonal antibodies. It is noteworthy to highlight protein candidates such as UL139 and US33A, which seem to be involved in the adhesion with the host cell or UL10 in which its location in the membrane has been demonstrated but its function has not been fully characterized [52].

In conclusion, using a complex in silico analysis we have predicted CMV proteins with TM domains that could be of interest because of their possible role in virus-cell interaction and entry. Our approach has been very useful to search for new potential candidates for a more rational design of new treatment targets and vaccines as well as to increase our knowledge of CMV.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23052768/s1.

Author Contributions: Conceptualization, E.G.-R.; Data curation, E.G.-R. and P.P.-R.; Formal analysis, E.G.-R. and P.P.-R.; Funding acquisition, P.P.-R.; Investigation, F.J.M., M.P.-M. and E.G.-R.; Methodology, F.J.M., M.P.-M., E.G.-R. and P.P.-R.; Software, M.P.-M.; Supervision, E.G.-R. and P.P.-R.; Validation, P.P.-R.; Visualization, P.P.-R.; Writing—original draft, F.J.M., M.P.-M. and E.G.-R.; Writing—review and editing, F.J.M., M.P.-M., E.G.-R. and P.P.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Spanish Ministry of Science, Innovation and University, Instituto de Salud Carlos III Grant/Award Numbers: PI17CIII-00014 (MPY110/18); DTS18CIII/00006 (MPY127/19); PI20-009 (MPY303/20). E.G-R is supported by the Sara Borrell Program (CD18CIII/00007), Instituto de Salud Carlos III, Ministerio de Ciencia, Innovación y Universidades. FJ.M is supported by the PFIS Program (F18III/00013), Instituto de Salud Carlos III, Ministerio de Ciencia, Innovación y Universidades.MP-M is supported by the Swedish Research Council (VR) (2019–03482).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Plotkin, S.A.; Boppana, S.B. Vaccination against the human cytomegalovirus. *Vaccine* **2019**, *37*, 7437–7442. [CrossRef] [PubMed]
2. Zuhair, M.; Smit, G.S.A.; Wallis, G.; Jabbar, F.; Smith, C.; Devleesschauwer, B.; Griffiths, P. Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. *Rev. Med. Virol.* **2019**, *29*, e2034. [CrossRef] [PubMed]
3. Murphy, E.; Shenk, T.E. *Human Cytomegalovirus Genome*; Shenk, T.E., Stinski, M.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 325, ISBN 9783540773481.
4. Emery, V.C. Investigation of CMV disease in immunocompromised patients. *J. Clin. Pathol.* **2001**, *54*, 84–88. [CrossRef]
5. Boeckh, M.; Geballe, A.P. Science in medicine Cytomegalovirus: Pathogen, paradigm, and puzzle. *J. Clin. Investig.* **2011**, *121*, 1673–1680. [CrossRef] [PubMed]
6. Seitz, R. Human Cytomegalovirus (HCMV)-Revised. *Transfus. Med. Hemotherapy* **2010**, *37*, 365–375. [CrossRef]
7. Griffiths, P.; Baraniak, I.; Reeves, M. The pathogenesis of human cytomegalovirus. *J. Pathol.* **2015**, *235*, 288–297. [CrossRef]
8. Stenberg, R.M.; Witte, P.R.; Stinski, M.F. Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate-early region 1. *J. Virol.* **1985**, *56*, 665–675. [CrossRef]
9. Dunn, W.; Chou, C.; Li, H.; Hai, R.; Patterson, D.; Stolc, V.; Zhu, H.; Liu, F. Functional profiling of a human cytomegalovirus genome. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14223–14228. [CrossRef]
10. Van Damme, E.; Van Loock, M. Functional annotation of human cytomegalovirus gene products: An update. *Front. Microbiol.* **2014**, *5*, 218. [CrossRef]
11. Balázs, Z.; Tombácz, D.; Szucs, A.; Csabai, Z.; Megyeri, K.; Petrov, A.N.; Snyder, M.; Boldogkoi, Z. Long-Read Sequencing of Human Cytomegalovirus Transcriptome Reveals RNA Isoforms Carrying Distinct Coding Potentials. *Sci. Rep.* **2017**, *7*, 15989. [CrossRef]
12. Stern-Ginossar, N.; Weisburd, B.; Michalski, A.; Le, V.T.K.; Hein, M.Y.; Huang, S.X.; Ma, M.; Shen, B.; Qian, S.B.; Hengel, H.; et al. Decoding human cytomegalovirus. *Science* **2012**, *338*, 1088–1093. [CrossRef]
13. Galitska, G.; Coscia, A.; Forni, D.; Steinbrueck, L.; De Meo, S.; Biolatti, M.; De Andrea, M.; Cagliani, R.; Leone, A.; Bertino, E.; et al. Genetic Variability of Human Cytomegalovirus Clinical Isolates Correlates with Altered Expression of Natural Killer Cell-Activating Ligands and IFN-γ. *Front. Immunol.* **2021**, *12*, 15989. [CrossRef] [PubMed]
14. Li, G.; Kamil, J.P. Viral Regulation of Cell Tropism in Human Cytomegalovirus. *J. Virol.* **2016**, *90*, 626–629. [CrossRef]
15. Gerna, G.; Kabanova, A.; Lilleri, D. Human Cytomegalovirus Cell Tropism and Host Cell Receptors. *Vaccines* **2019**, *7*, 70. [CrossRef] [PubMed]
16. Tandon, R.; Mocarski, E.S. Viral and host control of cytomegalovirus maturation. *Trends Microbiol.* **2012**, *20*, 392–401. [CrossRef] [PubMed]
17. Das, S.; Ortiz, D.A.; Gurczynski, S.J.; Khan, F.; Pellett, P.E. Identification of Human Cytomegalovirus Genes Important for Biogenesis of the Cytoplasmic Virion Assembly Complex. *J. Virol.* **2014**, *88*, 9086–9099. [CrossRef] [PubMed]
18. Bernstein, D.I.; Munoz, F.M.; Callahan, S.T.; Rupp, R.; Wootton, S.H.; Edwards, K.M.; Turley, C.B.; Stanberry, L.R.; Patel, S.M.; Mcneal, M.M.; et al. Safety and efficacy of a cytomegalovirus glycoprotein B (gB) vaccine in adolescent girls: A randomized clinical trial. *Vaccine* **2016**, *34*, 313–319. [CrossRef]
19. Pass, R.F. Development and evidence for efficacy of CMV glycoprotein B vaccine with MF59 adjuvant. *J. Clin. Virol.* **2009**, *46*, S73–S76. [CrossRef]
20. Nelson, C.S.; Herold, B.C.; Permar, S.R. A new era in cytomegalovirus vaccinology: Considerations for rational design of next-generation vaccines to prevent congenital cytomegalovirus infection. *NPJ Vaccines* **2018**, *3*, 38. [CrossRef]
21. Baraniak, I.; Kropff, B.; Ambrose, L.; McIntosh, M.; McLean, G.R.; Pichon, S.; Atkinson, C.; Milne, R.S.B.; Mach, M.; Griffiths, P.D.; et al. Protection from cytomegalovirus viremia following glycoprotein B vaccination is not dependent on neutralizing antibodies. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6273–6278. [CrossRef]
22. Nelson, C.S.; Huffman, T.; Jenks, J.A.; De La Rosa, E.C.; Xie, G.; Vandergrift, N.; Pass, R.F.; Pollara, J.; Permar, S.R. HCMV glycoprotein B subunit vaccine efficacy mediated by nonneutralizing antibody effector functions. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6267–6272. [CrossRef]
23. Vincenti, F.; Budde, K.; Merville, P.; Shihab, F.; Ram Peddi, V.; Shah, M.; Wyburn, K.; Cassuto-Viguier, E.; Weidemann, A.; Lee, M.; et al. A randomized, phase 2 study of ASP0113, a DNA-based vaccine, for the prevention of CMV in CMV-seronegative kidney transplant recipients receiving a kidney from a CMV-seropositive donor. *Am. J. Transplant.* **2018**, *18*, 2945–2954. [CrossRef]
24. Sandonís, V.; García-Ríos, E.; McConnell, M.J.; Pérez-Romero, P. Role of Neutralizing Antibodies in CMV Infection: Implications for New Therapeutic Approaches. *Trends Microbiol.* **2020**, *28*, 900–912. [CrossRef]
25. Pérez-Romero, P.; Blanco, P.; Giménez, E.; Solano, C.; Navarro, D. An update on the treatment of cytomegalovirus infection after allogeneic hematopoietic stem cell transplantation. *Expert Rev. Hematol.* **2015**, *12*, 937–945. [CrossRef]
26. Krause, P.R.; Bialek, S.R.; Boppana, S.B.; Griffiths, P.D.; Laughlin, C.A.; Ljungdahl, P.O.; Mocarski, E.S.; Pass, R.F.; Read, J.S.; Schleiss, M.R.; et al. Priorities for CMV vaccine development. *Vaccine* **2013**, *32*, 4–10. [CrossRef] [PubMed]
27. Gomes, A.C.; Griffiths, P.D.; Reeves, M.B. The Humoral Immune Response Against the gB Vaccine: Lessons Learnt from Protection in Solid Organ Transplantation. *Vaccines* **2019**, *7*, 67. [CrossRef]
28. Caposio, P.; van den Worm, S.; Crawford, L.; Perez, W.; Kreklywich, C.; Gilbride, R.M.; Hughes, C.M.; Ventura, A.B.; Ratts, R.; Marshall, E.E.; et al. Characterization of a live-attenuated HCMV-based vaccine platform. *Sci. Rep.* **2019**, *9*, 19236. [CrossRef] [PubMed]

29. Schleiss, M.R. Cytomegalovirus vaccines under clinical development. *J. Virus Erad.* **2016**, *2*, 198–207. [CrossRef]
30. Cui, X.; Cao, Z.; Wang, S.; Lee, R.B.; Wang, X.; Murata, H.; Adler, S.P.; McVoy, M.A.; Snapper, C.M. Novel trimeric human cytomegalovirus glycoprotein B elicits a high-titer neutralizing antibody response. *Vaccine* **2018**, *36*, 5580–5590. [CrossRef]
31. Dargan, D.J.; Douglas, E.; Cunningham, C.; Jamieson, F.; Stanton, R.J.; Baluchova, K.; McSharry, B.P.; Tomasec, P.; Emery, V.C.; Percivalle, E.; et al. Sequential mutations associated with adaptation of human cytomegalovirus to growth in cell culture. *J. Gen. Virol.* **2010**, *91*, 1535–1546. [CrossRef]
32. Cappadona, I.; Villinger, C.; Schutzius, G.; Mertens, T.; von Einem, J. Human Cytomegalovirus pUL47 Modulates Tegumentation and Capsid Accumulation at the Viral Assembly Complex. *J. Virol.* **2015**, *89*, 7314–7328. [CrossRef] [PubMed]
33. Zhu, F.; Yuan, J.; Li, H.J.; Zeng, Z.F.; Luo, Z.W.; Li, S.Q.; He, C.Q.; Jia, X.F.; Zhang, X.; Zuo, H.; et al. Human cytomegalovirus UL49 encodes an early, virion-associated protein essential for virus growth in human foreskin fibroblasts. *Arch. Virol.* **2016**, *161*, 1273–1284. [CrossRef] [PubMed]
34. Zhang, W.; Yao, Y.; Chen, J.; Wang, M. Unconserved C terminal of human cytomegalovirus tegument protein pUL76 elicits nuclear aggresome formation and induces DNA damage in transfected cells. *J. Biomed. Sci.* **2015**, *22*, 95. [CrossRef] [PubMed]
35. Köppen-Rung, P.; Dittmer, A.; Bogner, E. Intracellular Distribution of Capsid-Associated pUL77 of Human Cytomegalovirus and Interactions with Packaging Proteins and pUL93. *J. Virol.* **2016**, *90*, 5876–5885. [CrossRef] [PubMed]
36. Adair, R.; Douglas, E.R.; MacLean, J.B.; Graham, S.Y.; Aitken, J.D.; Jamieson, F.E.; Dargan, D.J. The products of human cytomegalovirus genes UL23, UL24, UL43 and US22 are tegument components. *J. Gen. Virol.* **2002**, *83*, 1315–1324. [CrossRef]
37. Luo, J.; Chen, J.; Yang, E.; Shen, A.; Gong, H.; Pei, Z.; Xiao, G.; Lu, S.; Liu, F. Modulation of the Cellular Distribution of Human Cytomegalovirus Helicase by Cellular Factor Snapin. *J. Virol.* **2013**, *87*, 10628–10640. [CrossRef]
38. Pizzorno, M.C.; Mullen, M.A.; Chang, Y.N.; Hayward, G.S. The functionally active IE2 immediate-early regulatory protein of human cytomegalovirus is an 80-kilodalton polypeptide that contains two distinct activator domains and a duplicated nuclear localization signal. *J. Virol.* **1991**, *65*, 3839–3852. [CrossRef]
39. Wang, Y.; Mao, L.; Kankanala, J.; Wang, Z.; Geraghty, R.J. Inhibition of Human Cytomegalovirus pUL89 Terminase Subunit Blocks Virus Replication and Genome Cleavage. *J. Virol.* **2017**, *91*, e02152-16. [CrossRef]
40. Connolly, S.A.; Jardetzky, T.S.; Longnecker, R. The structural basis of herpesvirus entry. *Nat. Rev. Microbiol.* **2021**, *19*, 110–121. [CrossRef]
41. Nguyen, C.C.; Kamil, J.P. Pathogen at the gates: Human cytomegalovirus entry and cell tropism. *Viruses* **2018**, *10*, 704. [CrossRef]
42. Gao, S.; Ruan, S.; Ma, Y.; Li, M.; Wang, L.; Zheng, B.; Qi, Y.; Sun, Z.; Huang, Y.; Ruan, Q. Newly identified transcripts of UL4 and UL5 genes of human cytomegalovirus. *Acta Biochim. Pol.* **2015**, *62*, 97–101. [CrossRef] [PubMed]
43. Chang, C.P.; Vesole, D.H.; Nelson, J.; Oldstone, M.B.; Stinski, M.F. Identification and expression of a human cytomegalovirus early glycoprotein. *J. Virol.* **1989**, *63*, 3330–3337. [CrossRef] [PubMed]
44. Vezzani, G.; Amendola, D.; Yu, D.; Chandramouli, S.; Frigimelica, E.; Maione, D.; Merola, M. The Human Cytomegalovirus UL116 Glycoprotein Is a Chaperone to Control gH-Based Complexes Levels on Virions. *Front. Microbiol.* **2021**, *12*, 630121. [CrossRef] [PubMed]
45. Shikhagaie, M.; Mercé-Maldonado, E.; Isern, E.; Muntasell, A.; Albà, M.M.; López-Botet, M.; Hengel, H.; Angulo, A. The Human Cytomegalovirus-Specific UL1 Gene Encodes a Late-Phase Glycoprotein Incorporated in the Virion Envelope. *J. Virol.* **2012**, *86*, 4091–4101. [CrossRef] [PubMed]
46. Anselmi, G.; Giuliani, M.; Vezzani, G.; Ferranti, R.; Gentile, M.; Cortese, M.; Amendola, D.; Pacchiani, N.; D'Aurizio, R.; Bruno, L.; et al. Characterization of pUL5, an HCMV protein interacting with the cellular protein IQGAP1. *Virology* **2020**, *540*, 57–65. [CrossRef]
47. Gonzalez-Perez, A.C.; Stempel, M.; Wyler, E.; Urban, C.; Piras, A.; Hennig, T.; Ganskih, S.; Wei, Y.; Heim, A.; Landthaler, M.; et al. The Zinc Finger Antiviral Protein ZAP Restricts Human Cytomegalovirus and Selectively Binds and Destabilizes Viral UL4/UL5 Transcripts. *MBio* **2021**, *12*, e02683-20. [CrossRef]
48. Engel, P.; Pérez-Carmona, N.; Albà, M.M.; Robertson, K.; Ghazal, P.; Angulo, A. Human cytomegalovirus UL7, a homologue of the SLAM-family receptor CD229, impairs cytokine production. *Immunol. Cell Biol.* **2011**, *89*, 753–766. [CrossRef]
49. MacManiman, J.D.; Meuser, A.; Botto, S.; Smith, P.P.; Liu, F.; Jarvis, M.A.; Nelson, J.A.; Caposio, P. Human Cytomegalovirus-Encoded pUL7 Is a Novel CEACAM1-Like Molecule Responsible for Promotion of Angiogenesis. *MBio* **2014**, *5*, e02035. [CrossRef]
50. Crawford, L.B.; Kim, J.H.; Collins-McMillen, D.; Lee, B.J.; Landais, I.; Held, C.; Nelson, J.A.; Yurochko, A.D.; Caposio, P. Human cytomegalovirus encodes a novel FLT3 receptor ligand necessary for hematopoietic cell differentiation and viral reactivation. *MBio* **2018**, *9*, e00682-18. [CrossRef]
51. Pérez-Carmona, N.; Martínez-Vicente, P.; Farré, D.; Gabaev, I.; Messerle, M.; Engel, P.; Angulo, A. A Prominent Role of the Human Cytomegalovirus UL8 Glycoprotein in Restraining Proinflammatory Cytokine Production by Myeloid Cells at Late Times during Infection. *J. Virol.* **2018**, *92*, e02229-17. [CrossRef]
52. Bruno, L.; Cortese, M.; Monda, G.; Gentile, M.; Caló, S.; Schiavetti, F.; Zedda, L.; Cattaneo, E.; Piccioli, D.; Schaefer, M.; et al. Human cytomegalovirus pUL10 interacts with leukocytes and impairs TCR-mediated T-cell activation. *Immunol. Cell Biol.* **2016**, *94*, 849–860. [CrossRef] [PubMed]
53. Zischke, J.; Mamareli, P.; Pokoyski, C.; Gabaev, I.; Buyny, S.; Jacobs, R.; Falk, C.S.; Lochner, M.; Sparwasser, T.; Schulz, T.F.; et al. The human cytomegalovirus glycoprotein pUL11 acts via CD45 to induce T cell IL-10 secretion. *PLoS Pathog.* **2017**, *13*, e1006454. [CrossRef] [PubMed]

54. Rölle, A.; Mousavi-Jazi, M.; Eriksson, M.; Odeberg, J.; Söderberg-Nauclér, C.; Cosman, D.; Kärre, K.; Cerboni, C. Effects of Human Cytomegalovirus Infection on Ligands for the Activating NKG2D Receptor of NK Cells: Up-Regulation of UL16-Binding Protein (ULBP)1 and ULBP2 Is Counteracted by the Viral UL16 Protein. *J. Immunol.* **2003**, *171*, 902–908. [CrossRef] [PubMed]
55. Kim, Y.; Park, B.; Cho, S.; Shin, J.; Cho, K.; Jun, Y.; Ahn, K. Human cytomegalovirus UL18 utilizes US6 for evading the NK and T-cell responses. *PLoS Pathog.* **2008**, *4*, 1000123. [CrossRef] [PubMed]
56. Jelcic, I.; Reichel, J.; Schlude, C.; Treutler, E.; Sinzger, C.; Steinle, A. The polymorphic HCMV glycoprotein UL20 is targeted for lysosomal degradation by multiple cytoplasmic dileucine motifs. *Traffic* **2011**, *12*, 1444–1456. [CrossRef] [PubMed]
57. Casarosa, P.; Gruijthuijsen, Y.K.; Michel, D.; Beisser, P.S.; Holl, J.; Fitzsimons, C.P.; Verzijl, D.; Bruggeman, C.A.; Mertens, T.; Leurs, R.; et al. Constitutive Signaling of the Human Cytomegalovirus-encoded Receptor UL33 Differs from That of Its Rat Cytomegalovirus Homolog R33 by Promiscuous Activation of G Proteins of the Gq, Gi, and Gs Classes. *J. Biol. Chem.* **2003**, *278*, 50010–50023. [CrossRef]
58. Van Senten, J.R.; Bebelman, M.P.; Fan, T.S.; Heukers, R.; Bergkamp, N.D.; Van Gasselt, P.; Langemeijer, E.V.; Slinger, E.; Lagerweij, T.; Rahbar, A.; et al. The human cytomegalovirus-encoded G protein- coupled receptor UL33 exhibits oncomodulatory properties. *J. Biol. Chem.* **2019**, *294*, 16297–16308. [CrossRef]
59. van Senten, J.R.; Bebelman, M.P.; van Gasselt, P.; Bergkamp, N.D.; van den Bor, J.; Siderius, M.; Smit, M.J. Human cytomegalovirus-encoded G protein-coupled receptor UL33 facilitates virus dissemination via the extracellular and cell-to-cell route. *Viruses* **2020**, *12*, 594. [CrossRef]
60. Xi, Y.; Harwood, S.; Wise, L.M.; Purdy, J.G. Human Cytomegalovirus pUL37x1 Is Important for Remodeling of Host Lipid Metabolism. *J. Virol.* **2019**, *93*, e00843-19. [CrossRef]
61. Brune, W.; Andoniou, C.E. Die another day: Inhibition of cell death pathways by cytomegalovirus. *Viruses* **2017**, *9*, 249. [CrossRef]
62. Prod'homme, V.; Tomasec, P.; Cunningham, C.; Lemberg, M.K.; Stanton, R.J.; McSharry, B.P.; Wang, E.C.Y.; Cuff, S.; Martoglio, B.; Davison, A.J.; et al. Human Cytomegalovirus UL40 Signal Peptide Regulates Cell Surface Expression of the NK Cell Ligands HLA-E and gpUL18. *J. Immunol.* **2012**, *188*, 2794–2804. [CrossRef] [PubMed]
63. Koshizuka, T.; Inoue, N. Activation of c-Jun by human cytomegalovirus UL42 through JNK activation. *PLoS ONE* **2020**, *15*, e0232635. [CrossRef] [PubMed]
64. Fu, Y.Z.; Guo, Y.; Zou, H.M.; Su, S.; Wang, S.Y.; Yang, Q.; Luo, M.H.; Wang, Y.Y. Human cytomegalovirus protein UL42 antagonizes cGAS/MITA-mediated innate antiviral response. *PLoS Pathog.* **2019**, *15*, e1007691. [CrossRef]
65. Häge, S.; Sonntag, E.; Svrlanska, A.; Borst, E.M.; Stilp, A.C.; Horsch, D.; Müller, R.; Kropff, B.; Milbradt, J.; Stamminger, T.; et al. Phenotypical characterization of the nuclear egress of recombinant cytomegaloviruses reveals defective replication upon orf-ul50 deletion but not pul50 phosphosite mutation. *Viruses* **2021**, *13*, 165. [CrossRef]
66. Weiler, N.; Paal, C.; Adams, K.; Calcaterra, C.; Fischer, D.; Stanton, R.J.; Stöhr, D.; Sampaio, K.L.; Sinzger, C. Role of envelope glycoprotein complexes in cell-associated spread of human cytomegalovirus. *Viruses* **2021**, *13*, 614. [CrossRef]
67. Mach, M.; Osinski, K.; Kropff, B.; Schloetzer-Schrehardt, U.; Krzyzaniak, M.; Britt, W. The Carboxy-Terminal Domain of Glycoprotein N of Human Cytomegalovirus Is Required for Virion Morphogenesis. *J. Virol.* **2007**, *81*, 5212–5224. [CrossRef] [PubMed]
68. Wu, Y.; Prager, A.; Boos, S.; Resch, M.; Brizic, I.; Mach, M.; Wildner, S.; Scrivano, L.; Adler, B. Human cytomegalovirus glycoprotein complex gH/gL/gO uses PDGFR-α as a key for entry. *PLoS Pathog.* **2017**, *13*, e1006281. [CrossRef]
69. van Senten, J.R.; Fan, T.S.; Siderius, M.; Smit, M.J. Viral G protein-coupled receptors as modulators of cancer hallmarks. *Pharmacol. Res.* **2020**, *156*, 104804. [CrossRef]
70. O'Connor, C.M.; Shenk, T. Human Cytomegalovirus pUL78 G Protein-Coupled Receptor Homologue Is Required for Timely Cell Entry in Epithelial Cells but Not Fibroblasts. *J. Virol.* **2012**, *86*, 11425–11433. [CrossRef]
71. Krzyzaniak, M.A.; Mach, M.; Britt, W.J. HCMV-encoded glycoprotein M (UL100) interacts with rab11 effector protein FIP4. *Traffic* **2009**, *10*, 1439–1457. [CrossRef]
72. Krzyzaniak, M.; Mach, M.; Britt, W.J. The Cytoplasmic Tail of Glycoprotein M (gpUL100) Expresses Trafficking Signals Required for Human Cytomegalovirus Assembly and Replication. *J. Virol.* **2007**, *81*, 10316–10328. [CrossRef] [PubMed]
73. Pandey, J.P.; Namboodiri, A.M.; Radwan, F.F.; Nietert, P.J. The decoy Fcγ receptor encoded by the cytomegalovirus UL119-UL118 gene has differential affinity to IgG proteins expressing different GM allotypes. *Hum. Immunol.* **2015**, *76*, 591–594. [CrossRef] [PubMed]
74. Landini, M.P.; Lazzarotto, T.; Xu, J.; Geballe, A.P.; Mocarski, E.S. Humoral Immune Response to Proteins of Human Cytomegalovirus Latency-Associated Transcripts. *Biol. Blood Marrow Transplant.* **2000**, *6*, 100–108. [CrossRef]
75. Wu, H.; Kropff, B.; Mach, M.; Britt, W.J. Human cytomegalovirus envelope protein gpul132 regulates infectious virus production through formation of the viral assembly compartment. *MBio* **2020**, *11*, e02044-20. [CrossRef]
76. Mlera, L.; Moy, M.; Maness, K.; Tran, L.N.; Goodrum, F.D. The Role of the Human Cytomegalovirus UL133-UL138 Gene Locus in Latency and Reactivation. *Viruses* **2020**, *12*, 714. [CrossRef] [PubMed]
77. Stanton, R.J.; Prod'Homme, V.; Purbhoo, M.A.; Moore, M.; Aicheler, R.J.; Heinzmann, M.; Bailer, S.M.; Haas, J.; Antrobus, R.; Weekes, M.P.; et al. HCMV pUL135 remodels the actin cytoskeleton to impair immune recognition of infected cells. *Cell Host Microbe* **2014**, *16*, 201–214. [CrossRef] [PubMed]
78. Bughio, F.; Umashankar, M.; Wilson, J.; Goodrum, F. Human Cytomegalovirus UL135 and UL136 Genes Are Required for Postentry Tropism in Endothelial Cells. *J. Virol.* **2015**, *89*, 6536–6550. [CrossRef] [PubMed]

79. Caviness, K.; Bughio, F.; Crawford, L.B.; Streblow, D.N.; Nelson, J.A.; Caposio, P.; Goodrum, F. Complex interplay of the UL136 isoforms balances cytomegalovirus replication and Latency. *MBio* **2016**, *7*, e01986. [CrossRef]
80. Bughio, F.; Elliott, D.A.; Goodrum, F. An Endothelial Cell-Specific Requirement for the UL133-UL138 Locus of Human Cytomegalovirus for Efficient Virus Maturation. *J. Virol.* **2013**, *87*, 3062–3075. [CrossRef]
81. Han, L.; Ma, Y.; Liu, Z.; Liu, C.; Lu, Y.; Qi, Y.; Huang, Y.; Sun, Z.; Ruan, Q. Transcriptional regulation and influence on replication of the human cytomegalovirus UL1381.4 kb transcript. *Mol. Med. Rep.* **2017**, *16*, 5649–5658. [CrossRef]
82. Gelbmann, C.B.; Kalejta, R.F. The Golgi sorting motifs of human cytomegalovirus UL138 are not required for latency maintenance. *Virus Res.* **2019**, *270*, 197646. [CrossRef] [PubMed]
83. Qi, Y.; Mao, Z.Q.; Ruan, Q.; He, R.; Ma, Y.P.; Sun, Z.R.; Ji, Y.H.; Huang, Y. Human cytomegalovirus (HCMV) UL139 open reading frame: Sequence variants are clustered into three major genotypes. *J. Med. Virol.* **2006**, *78*, 517–522. [CrossRef] [PubMed]
84. Zou, F.; Lu, Z.T.; Wang, S.; Wu, S.; Wu, Y.Y.; Sun, Z.R. Human cytomegalovirus UL141 protein interacts with CELF5 and affects viral DNA replication. *Mol. Med. Rep.* **2018**, *17*, 4657–4664. [CrossRef] [PubMed]
85. Tomasec, P.; Wang, E.C.Y.; Davison, A.J.; Vojtesek, B.; Armstrong, M.; Griffin, C.; McSharry, B.P.; Morris, R.J.; Llewellyn-Lacey, S.; Rickards, C.; et al. Downregulation of natural killer cell-activating ligang CD155 by human cytomegalovirus UL141. *Nat. Immunol.* **2005**, *6*, 181–188. [CrossRef] [PubMed]
86. Nemčovičová, I.; Benedict, C.A.; Zajonc, D.M. Structure of Human Cytomegalovirus UL141 Binding to TRAIL-R2 Reveals Novel, Non-canonical Death Receptor Interactions. *PLoS Pathog.* **2013**, *9*, e1003224. [CrossRef]
87. Ashiru, O.; Bennett, N.J.; Boyle, L.H.; Thomas, M.; Trowsdale, J.; Wills, M.R. NKG2D Ligand MICA Is Retained in the cis -Golgi Apparatus by Human Cytomegalovirus Protein UL142. *J. Virol.* **2009**, *83*, 12345–12354. [CrossRef]
88. Bitra, A.; Nemčovičová, I.; Picarda, G.; Doukov, T.; Wang, J.; Benedict, C.A.; Zajonc, D.M. Structure of human cytomegalovirus UL144, an HVEM orthologue, bound to the B and T cell lymphocyte attenuator. *J. Biol. Chem.* **2019**, *294*, 10519–10529. [CrossRef]
89. Poole, E.; Walther, A.; Raven, K.; Benedict, C.A.; Mason, G.M.; Sinclair, J. The Myeloid Transcription Factor GATA-2 Regulates the Viral UL144 Gene during Human Cytomegalovirus Latency in an Isolate-Specific Manner. *J. Virol.* **2013**, *87*, 4261–4271. [CrossRef]
90. Lüttichau, H.R. The cytomegalovirus UL146 gene product vCXCL1 targets both CXCR1 and CXCR2 as an agonist. *J. Biol. Chem.* **2010**, *285*, 9137–9146. [CrossRef]
91. Seidel, E.; Dassa, L.; Schuler, C.; Oiknine-Djian, E.; Wolf, D.G.; Le-Trilling, V.T.K.; Mandelboim, O. The human cytomegalovirus protein UL147A downregulates the most prevalent MICA allele: MICA*008, to evade NK cell-mediated killing. *PLoS Pathog.* **2021**, *17*, e1008807. [CrossRef] [PubMed]
92. Siddiquey, M.N.A.; Schultz, E.P.; Yu, Q.; Amendola, D.; Vezzani, G.; Yu, D.; Maione, D.; Lanchy, J.-M.; Ryckman, B.J.; Merola, M.; et al. The human cytomegalovirus protein UL116 interacts with the viral ER resident glycoprotein UL148 and promotes the incorporation of gH/gL complexes into virions. *J. Virol.* **2021**, *95*, e0220720. [CrossRef] [PubMed]
93. Dassa, L.; Seidel, E.; Oiknine-Djian, E.; Yamin, R.; Wolf, D.G.; Le-Trilling, V.T.K.; Mandelboim, O. The Human Cytomegalovirus Protein UL148A Downregulates the NK Cell-Activating Ligand MICA to Avoid NK Cell Attack. *J. Virol.* **2018**, *92*, e00162-18. [CrossRef] [PubMed]
94. Ji, Y.H.; Rong Sun, Z.; Ruan, Q.; Guo, J.J.; He, R.; Qi, Y.; Ma, Y.P.; Mao, Z.Q.; Huang, Y.J. Polymorphisms of human cytomegalovirus UL148A, UL148B, UL148C, UL148D genes in clinical strains. *J. Clin. Virol.* **2006**, *37*, 252–257. [CrossRef]
95. Gabor, F.; Jahn, G.; Sedmak, D.D.; Sinzger, C. In vivo Downregulation of MHC Class I Molecules by HCMV Occurs during All Phases of Viral Replication but Is Not Always Complete. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 283. [CrossRef]
96. Park, B.; Kim, Y.; Shin, J.; Lee, S.; Cho, K.; Früh, K.; Lee, S.; Ahn, K. Human Cytomegalovirus Inhibits Tapasin-Dependent Peptide Loading and Optimization of the MHC Class I Peptide Cargo for Immune Evasion. *Immunity* **2004**, *20*, 71–85. [CrossRef]
97. Dugan, G.E.; Hewitt, E.W. Structural and Functional Dissection of the Human Cytomegalovirus Immune Evasion Protein US6. *J. Virol.* **2008**, *82*, 3271–3282. [CrossRef] [PubMed]
98. Park, A.; Ra, E.A.; Lee, T.A.; Choi, H.; Lee, E.; Kang, S.; Seo, J.Y.; Lee, S.; Park, B. HCMV-encoded US7 and US8 act as antagonists of innate immunity by distinctively targeting TLR-signaling pathways. *Nat. Commun.* **2019**, *10*, 4670. [CrossRef]
99. Choi, H.J.; Park, A.; Kang, S.; Lee, E.; Lee, T.A.; Ra, E.A.; Lee, J.; Lee, S.; Park, B. Human cytomegalovirus-encoded US9 targets MAVS and STING signaling to evade type i interferon immune responses. *Nat. Commun.* **2018**, *9*, 125. [CrossRef]
100. Park, B.; Spooner, E.; Houser, B.L.; Strominger, J.L.; Ploegh, H.L. The HCMV membrane glycoprotein US10 selectively targets HLA-G for degradation. *J. Exp. Med.* **2010**, *207*, 2033–2041. [CrossRef]
101. Zimmermann, C.; Kowalewski, D.; Bauersfeld, L.; Hildenbrand, A.; Gerke, C.; Schwarzmüller, M.; Le-Trilling, V.T.K.; Stevanovic, S.; Hengel, H.; Momburg, F.; et al. HLA-B locus products resist degradation by the human cytomegalovirus immunoevasin US11. *PLoS Pathog.* **2019**, *15*, e1008040. [CrossRef]
102. Fielding, C.A.; Weekes, M.P.; Nobre, L.V.; Ruckova, E.; Wilkie, G.S.; Paulo, J.A.; Chang, C.; Suárez, N.M.; Davies, J.A.; Antrobus, R.; et al. Control of immune ligands by members of a cytomegalovirus gene expansion suppresses natural killer cell activation. *Elife* **2017**, *6*, e22206. [CrossRef] [PubMed]
103. Das, S.; Pellett, P.E. Members of the HCMV US12 family of predicted heptaspanning membrane proteins have unique intracellular distributions, including association with the cytoplasmic virion assembly complex. *Virology* **2007**, *361*, 263–273. [CrossRef] [PubMed]

104. Luganini, A.; Cavaletto, N.; Raimondo, S.; Geuna, S.; Gribaudo, G. Loss of the Human Cytomegalovirus US16 Protein Abrogates Virus Entry into Endothelial and Epithelial Cells by Reducing the Virion Content of the Pentamer. *J. Virol.* **2017**, *91*, 205–222. [CrossRef] [PubMed]
105. Gurczynski, S.J.; Das, S.; Pellett, P.E. Deletion of the Human Cytomegalovirus US17 Gene Increases the Ratio of Genomes per Infectious Unit and Alters Regulation of Immune and Endoplasmic Reticulum Stress Response Genes at Early and Late Times after Infection. *J. Virol.* **2014**, *88*, 2168–2182. [CrossRef]
106. Charpak-Amikam, Y.; Kubsch, T.; Seidel, E.; Oiknine-Djian, E.; Cavaletto, N.; Yamin, R.; Schmiedel, D.; Wolf, D.; Gribaudo, G.; Messerle, M.; et al. Human cytomegalovirus escapes immune recognition by NK cells through the downregulation of B7-H6 by the viral genes US18 and US20. *Sci. Rep.* **2017**, *7*, 8661. [CrossRef]
107. Cavaletto, N.; Luganini, A.; Gribaudo, G. Inactivation of the Human Cytomegalovirus US20 Gene Hampers Productive Viral Replication in Endothelial Cells. *J. Virol.* **2015**, *89*, 11092–11106. [CrossRef]
108. Luganini, A.; Di Nardo, G.; Munaron, L.; Gilardi, G.; Pla, A.F.; Gribaudo, G. Human cytomegalovirus US21 protein is a viroporin that modulates calcium homeostasis and protects cells against apoptosis. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E12370–E12377. [CrossRef]
109. Boeck, J.M.; Stowell, G.A.; O'Connor, C.M.; Spencer, J.V. The Human Cytomegalovirus US27 Gene Product Constitutively Activates Antioxidant Response Element-Mediated Transcription through G β γ, Phosphoinositide 3-Kinase, and Nuclear Respiratory Factor 1. *J. Virol.* **2018**, *92*, e00644-18. [CrossRef]
110. Tu, C.C.; O'Connor, C.M.; Spencer, J.V. Identification of a novel signaling complex containing host chemokine receptor CXCR4, Interleukin-10 receptor, and human cytomegalovirus US27. *Virology* **2020**, *548*, 49–58. [CrossRef]
111. O'Connor, C.M.; Shenk, T. Human Cytomegalovirus pUS27 G Protein-Coupled Receptor Homologue Is Required for Efficient Spread by the Extracellular Route but Not for Direct Cell-to-Cell Spread. *J. Virol.* **2011**, *85*, 3700–3707. [CrossRef]
112. Nobre, L.; Nightingale, K.; Ravenhill, B.J.; Antrobus, R.; Soday, L.; Nichols, J.; Davies, J.; Seirafian, S.; Wang, E.C.Y.; Davison, A.J.; et al. Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions. *Elife* **2019**, *8*, e49894. [CrossRef] [PubMed]
113. Krishna, B.A.; Miller, W.E.; O'Connor, C.M. US28: HCMV's swiss army knife. *Viruses* **2018**, *10*, 445. [CrossRef]
114. Gatherer, D.; Seirafian, S.; Cunningham, C.; Holton, M.; Dargan, D.J.; Baluchova, K.; Hector, R.D.; Galbraith, J.; Herzyk, P.; Wilkinson, G.W.G.; et al. High-resolution human cytomegalovirus transcriptome. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19755–19760. [CrossRef] [PubMed]
115. Kim, E.T.; Kim, Y.E.; Kim, Y.J.; Lee, M.K.; Hayward, G.S.; Ahn, J.H. Analysis of human cytomegalovirus-encoded SUMO targets and temporal regulation of SUMOylation of the immediate-early proteins IE1 and IE2 during infection. *PLoS ONE* **2014**, *9*, 103308. [CrossRef] [PubMed]
116. Lilley, B.N.; Ploegh, H.L.; Tirabassi, R.S. Human Cytomegalovirus Open Reading Frame TRL11/IRL11 Encodes an Immunoglobulin G Fc-Binding Protein. *J. Virol.* **2001**, *75*, 11218–11221. [CrossRef]
117. Wilkinson, G.W.G.; Davison, A.J.; Tomasec, P.; Fielding, C.A.; Aicheler, R.; Murrell, I.; Seirafian, S.; Wang, E.C.Y.; Weekes, M.; Lehner, P.J.; et al. Human cytomegalovirus: Taking the strain. *Med. Microbiol. Immunol.* **2015**, *204*, 273–284. [CrossRef]
118. Cortese, M.; Calò, S.; D'Aurizio, R.; Lilja, A.; Pacchiani, N.; Merola, M. Recombinant Human Cytomegalovirus (HCMV) RL13 Binds Human Immunoglobulin G Fc. *PLoS ONE* **2012**, *7*. [CrossRef] [PubMed]
119. Wang, G.; Ren, G.; Cui, X.; Lu, Z.; Ma, Y.; Qi, Y.; Huang, Y.; Liu, Z.; Sun, Z.; Ruan, Q. Human cytomegalovirus RL13 protein interacts with host NUDT14 protein affecting viral DNA replication. *Mol. Med. Rep.* **2016**, *13*, 2167–2174. [CrossRef]
120. Schultz, E.P.; Lanchy, J.-M.; Day, L.Z.; Yu, Q.; Peterson, C.; Preece, J.; Ryckman, B.J. Specialization for Cell-Free or Cell-to-Cell Spread of BAC-Cloned Human Cytomegalovirus Strains Is Determined by Factors beyond the UL128-131 and RL13 Loci. *J. Virol.* **2020**, *94*, e00034-20. [CrossRef] [PubMed]
121. Lesniewski, M.; Das, S.; Skomorovska-Prokvolit, Y.; Wang, F.Z.; Pellett, P.E. Primate cytomegalovirus US12 gene family: A distinct and diverse clade of seven-transmembrane proteins. *Virology* **2006**, *354*, 286–298. [CrossRef]
122. Mozzi, A.; Biolatti, M.; Cagliani, R.; Forni, D.; Dell'Oste, V.; Pontremoli, C.; Vantaggiato, C.; Pozzoli, U.; Clerici, M.; Landolfo, S.; et al. Past and ongoing adaptation of human cytomegalovirus to its host. *PLoS Pathog.* **2020**, *16*, e1008476. [CrossRef] [PubMed]
123. Davison, A.J.; Akter, P.; Cunningham, C.; Dolan, A.; Addison, C.; Dargan, D.J.; Hassan-Walker, A.F.; Emery, V.C.; Griffiths, P.D.; Wilkinson, G.W.G. Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J. Gen. Virol.* **2003**, *84*, 657–663. [CrossRef] [PubMed]
124. Scarborough, J.A.; Paul, J.R.; Spencer, J.V. Evolution of the ability to modulate host chemokine networks via gene duplication in Human Cytomegalovirus (HCMV). *Infect. Genet. Evol.* **2017**, *51*, 46–53. [CrossRef] [PubMed]
125. 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**, *11*, 2104. [CrossRef] [PubMed]
126. Patro, A.R.K. Subversion of immune response by human cytomegalovirus. *Front. Immunol.* **2019**, *10*, 1155. [CrossRef]
127. Weekes, M.P.; Tomasec, P.; Huttlin, E.L.; Fielding, C.A.; Nusinow, D.; Stanton, R.J.; Wang, E.C.Y.; Aicheler, R.; Murrell, I.; Wilkinson, G.W.G.; et al. Quantitative temporal viromics: An approach to investigate host-pathogen interaction. *Cell* **2014**, *157*, 1460–1472. [CrossRef]

128. Couté, Y.; Kraut, A.; Zimmermann, C.; Büscher, N.; Hesse, A.M.; Bruley, C.; De Andrea, M.; Wangen, C.; Hahn, F.; Marschall, M.; et al. Mass spectrometry-based characterization of the virion proteome, phosphoproteome, and associated kinase activity of human cytomegalovirus. *Microorganisms* **2020**, *8*, 820. [CrossRef] [PubMed]
129. Di Bona, D.; Scafidi, V.; Plaia, A.; Colomba, C.; Nuzzo, D.; Occhino, C.; Tuttolomondo, A.; Giammanco, G.; De Grazia, S.; Montalto, G.; et al. HLA and killer cell immunoglobulin-like receptors influence the natural course of CMV infection. *J. Infect. Dis.* **2014**, *210*, 1083–1089. [CrossRef]
130. Aiello, A.; Accardi, G.; Candore, G.; Caruso, C.; Colomba, C.; Di Bona, D.; Duro, G.; Gambino, C.M.; Ligotti, M.E.; Pandey, J.P. Role of immunogenetics in the outcome of HCMV infection: Implications for ageing. *Int. J. Mol. Sci.* **2019**, *20*, 685. [CrossRef]
131. Di Bona, D.; Accardi, G.; Aiello, A.; Bilancia, M.; Candore, G.; Colomba, C.; Caruso, C.; Duro, G.; Gambino, C.M.; Macchia, L.; et al. Association between γ marker, human leucocyte antigens and killer immunoglobulin-like receptors and the natural course of human cytomegalovirus infection: A pilot study performed in a Sicilian population. *Immunology* **2018**, *153*, 523–531. [CrossRef]
132. Pandey, J.P.; Kistner-Griffin, E.; Radwan, F.F.; Kaur, N.; Namboodiri, A.M.; Black, L.; Butler, M.A.; Carreón, T.; Ruder, A.M. Immunoglobulin genes influence the magnitude of humoral immunity to cytomegalovirus Glycoprotein B. *J. Infect. Dis.* **2014**, *210*, 1823–1826. [CrossRef] [PubMed]
133. Aguttu, C.; Okech, B.A.; Mukisa, A.; Lubega, G.W. Screening and characterization of hypothetical proteins of Plasmodium falciparum as novel vaccine candidates in the fight against malaria using reverse vaccinology. *J. Genet. Eng. Biotechnol.* **2021**, *19*, 103. [CrossRef] [PubMed]
134. Almén, M.S.; Nordström, K.J.V.; Fredriksson, R.; Schiöth, H.B. Mapping the human membrane proteome: A majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* **2009**, *7*, 50. [CrossRef] [PubMed]
135. Orfanoudaki, G.; Economou, A. Proteome-wide subcellular topologies of E. coli polypeptides database (STEPdb). *Mol. Cell. Proteom.* **2014**, *13*, 3674–3687. [CrossRef] [PubMed]
136. Dal Monte, P.; Pignatelli, S.; Rossini, G.; Landini, M.P. Genomic variants among human cytomegalovirus (HCMV) clinical isolates: The glycoprotein n (gN) paradigm. *Hum. Immunol.* **2004**, *65*, 387–394. [CrossRef] [PubMed]
137. Mattick, C.; Dewin, D.; Polley, S.; Sevilla-Reyes, E.; Pignatelli, S.; Rawlinson, W.; Wilkinson, G.; Dal Monte, P.; Gompels, U.A. Linkage of human cytomegalovirus glycoprotein gO variant groups identified from worldwide clinical isolates with gN genotypes, implications for disease associations and evidence for N-terminal sites of positive selection. *Virology* **2004**, *318*, 582–597. [CrossRef] [PubMed]
138. Whisstock, J.C.; Lesk, A.M. Prediction of protein function from protein sequence and structure. *Q. Rev. Biophys.* **2003**, *36*, 307–340. [CrossRef]
139. Gilbert, C.; Feschotte, C. Genomic fossils calibrate the long-term evolution of hepadnaviruses. *PLoS Biol.* **2010**, *8*, e1000495. [CrossRef]
140. Tchoupa, A.K.; Schuhmacher, T.; Hauck, C.R. Signaling by epithelial members of the CEACAM family—Mucosal docking sites for pathogenic bacteria. *Cell Commun. Signal.* **2014**, *12*, 27. [CrossRef]
141. Chan, C.-M.; Chu, H.; Wang, Y.; Wong, B.H.-Y.; Zhao, X.; Zhou, J.; Yang, D.; Leung, S.P.; Chan, J.F.-W.; Yeung, M.-L.; et al. Carcinoembryonic Antigen-Related Cell Adhesion Molecule 5 Is an Important Surface Attachment Factor That Facilitates Entry of Middle East Respiratory Syndrome Coronavirus. *J. Virol.* **2016**, *90*, 9114–9127. [CrossRef]
142. Terakita, A. The opsins. *Genome Biol.* **2005**, *6*, 213. [CrossRef] [PubMed]
143. Arav-Boger, R.; Foster, C.B.; Zong, J.C.; Pass, R.F. Human cytomegalovirus-encoded α-chemokines exhibit high sequence variability in congenitally infected newborns. *J. Infect. Dis.* **2006**, *193*, 788–791. [CrossRef] [PubMed]
144. Whittaker, C.A.; Hynes, R.O. Distribution and Evolution of von Willebrand/Integrin A Domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. *Mol. Biol. Cell* **2002**, *13*, 4100–4109. [CrossRef] [PubMed]
145. Wang, Q.; Ni, C.; Li, Z.; Li, X.; Han, R.; Zhao, F.; Xu, J.; Gao, X.; Wang, S. PureseqTM: Efficient and accurate prediction of transmembrane topology from amino acid sequence only. *bioRxiv* **2019**. [CrossRef]
146. Käll, L.; Krogh, A.; Sonnhammer, E.L.L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **2004**, *338*, 1027–1036. [CrossRef]
147. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [CrossRef]
148. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. [CrossRef]
149. Queirós, P.; Delogu, F.; Hickl, O.; May, P.; Wilmes, P. Mantis: Flexible and consensus-driven genome. *Giga Sci.* **2021**, *10*, giab042. [CrossRef]
150. Aramaki, T.; Blanc-Mathieu, R.; Endo, H.; Ohkubo, K.; Kanehisa, M.; Goto, S.; Ogata, H. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **2020**, *36*, 2251–2252. [CrossRef]
151. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427–D432. [CrossRef]
152. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [CrossRef] [PubMed]

153. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [CrossRef] [PubMed]
154. Haft, D.H.; Selengut, J.D.; Richter, R.A.; Harkins, D.; Basu, M.K.; Beck, E. TIGRFAMs and genome properties in 2013. *Nucleic Acids Res.* **2013**, *41*, 387–395. [CrossRef] [PubMed]
155. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [CrossRef]
156. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef] [PubMed]
157. Wickham, H. Ggplot2: Elegant Graphics for Data Analysis. In *Ggplot2*; Springer: New York, NY, USA, 2009; pp. 1–7, ISBN 9780387981413.
158. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef]