

UL146 variability among clinical isolates of Human Cytomegalovirus from Japan

Francisco Aguayo^{1,*}, Tsugiya Murayama² and Yoshito Eizuru³

¹ Molecular Pathology and Epidemiology laboratory, Centro de Investigaciones Médicas, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile

² Department of Microbiology and Immunology, Faculty of Pharmaceutical Sciences, Hokuriku University, 3 Kanagawa-machi, Kanazawa 920-1181, Japan

³ Division of Persistent and Oncogenic Viruses, Center for Chronic Viral Diseases, Faculty of Medicine, Kagoshima University, Sakuragaoka 8-35-1, Kagoshima 890-8520, Japan

ABSTRACT

Human Cytomegalovirus (HCMV) is a herpesvirus associated with serious diseases in immunocompromised subjects. The region between ORF UL133 and UL151 from HCMV, named ULb' is frequently deleted in attenuated AD169 and in highly passaged laboratory strains. However, this region is conserved in low-passaged and more virulent HCMV, like the Toledo strain. The UL146 gene, which is located in the ULb' region, encodes a CXC-chemokine analogue. The diversity of UL146 gene was evaluated among fifty-six clinical isolates of HCMV from Japan. Results show that UL146 gene was successfully amplified by the polymerase chain reaction (PCR) in only 17/56 strains (30%), while the success rate for UL145/UL147 gene was 18/56 strains (32%). After DNA sequencing, the 35 amplified strains were classified into 8 groups. When compared, variability of UL146 ranged from 25.1% to 52.9% at the DNA level and from 34.5% to 67% at the amino acid level. Seven groups had the interleukin-8 (IL-8) motif ERL (Glu-Leu-Arg) CXC and one group had only the CXC motif, suggesting the absence of the IL-8 function of UL146. In conclusion, we found that UL146 gene of HCMV is hypervariable in clinical strains from Japan suggesting the possibility of a different function in each sequence group.

Key terms: a) Cytomegalovirus, b) UL146, c) Sequence, d) V.

INTRODUCTION

Human Cytomegalovirus (HCMV) is a widespread ubiquitous herpesvirus that usually produces asymptomatic infections in immunocompetent hosts. However, serious HCMV-associated diseases might occur in immunocompromised individuals such as organ transplant recipients or in patients with acquired immunodeficiency syndrome (AIDS) and in congenitally infected newborns (Britt WJ, 1996). The severity of HCMV diseases varies widely ranging from stillborn or neonatal death to no abnormalities in congenital infections and sight- and life-threatening diseases in immunocompromised hosts. Furthermore, the prevalence of HCMV retinitis in patients with AIDS is higher than among drug users (Brantsaeter et al., 2002). Clinical and epidemiological factors associated with severe disease have been identified (Puchhammer-Stockl & Gorzer, 2006), however little is known about viral factors associated with the variability in susceptibility to, or severity of HCMV-associated disease.

HCMV has a large and complex genome; containing approximately 235 Kb of DNA and more than 200 open reading frames (ORFs) (Dolan et al., 2004). Interestingly, in laboratory-adopted attenuated HCMV strains, such as AD169, 19 ORFs between ORF UL133 and UL151 are deleted (Cha et al., 1996), suggesting that proteins encoded by these genes are not required for HCMV replication in cell culture. Nonetheless, this region is present in the Toledo strain of HCMV (HCMV_{Tol}) and also in low-passaged laboratory isolates, suggesting that these genes or their encoded proteins play a role in the wild-type HCMV strains and may be essential for viral infection or pathogenicity *in vivo*. The

UL146 gene, which encodes for a protein similar to α (CXC)-chemokine (Penfold et al., 1999), is located between UL133 and UL151. The UL146 product, vCXCL-1, is a 117-amino acid secreted glycoprotein that functions as a selective agonist for CXCR1 and CXCR2, although with differential affinity and potency (Luttichau, 2010). Its expression correlates with the ability to attract human neutrophils. In fact, recombinant vCXCL-1 has full chemokine functions, such as the induction of calcium mobilization, chemotaxis and human neutrophils degranulation (Penfold et al., 1999). The UL146 gene has been previously reported to be hypervariable among clinical isolates of HCMV. Additionally, phylogenetic analysis reveals that different unrelated strains cluster in defined sequence groups (Arav-Boger et al., 2006; Hassan-Walker et al., 2004; Lurain et al., 2006). In this report, we described the variability of UL146 gene and the phylogenetic relatedness among clinical isolates of HCMV from Japan.

MATERIALS AND METHODS

Cells and viruses

Human embryonic lung fibroblasts (HEL) were grown in Eagle's minimum essential medium (Nissui Pharmaceutical Co., LTD., Tokyo) supplemented with 10% fetal bovine serum (Flow Laboratories Inc., Irvine, VA), 60 μ g/mL of kanamycin and 0.12 % NaHCO₃. HEL cells were used for HCMV isolation and propagation. Thirty-three clinical isolates were obtained from urine, 2 from amniotic fluid, 2 from saliva, 4 from lymphocytes, 1 from mother's milk, 1 from throat swab, 1 from tears and 9 from other human

* Corresponding author: Francisco Aguayo González, Ph.D., Programa de Virología, Instituto de Ciencias Biomédicas (ICBM), Facultad de Medicina, Universidad de Chile, Fono: 56-2-9786317, Fax : 56-2-9786124, E-mail: faguayo@med.uchile.cl

sources. Almost all the specimens used in this study were obtained from subjects with congenital HCMV infection, except for one liver and four renal post-transplanted patients. All the clinical strains were passaged as a cell-associated virus 5 to 8 times before preparation of viral DNA.

Preparation of viral DNA

Viral DNA was prepared from the infected cells following the Hirt's procedure (Hirt, 1967). The Hirt's supernatant was treated with Proteinase K (50 µg/mL) at 56°C for 2 hours, followed by phenol/chloroform extraction and ethanol precipitation. Finally, the DNA was dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA).

Polymerase chain reaction for the UL146 sequence

The Hirt's supernatant DNA was diluted in sterilized water to optimal concentrations for use as templates for PCR amplification of specific regions of the HCMV genome. All the samples were shown to be HCMV positive through IE1 region amplification.

In the initial study of the region containing 19 ORFs homologous to Toledo strain sequence (Cha et al., 1996), a 354 bp portion (nucleotide 1 to 354) of UL146 coding sequence was amplified using the primer pair: forward 5'-CCATGCATATGCGATTAATTTTTGGT-3'; reverse 5'-GCTCTAGATTATCCTCTAACACCTAT-3'. The entire UL145 gene (303 bp) and UL147 (480 bp) were amplified using the following primer pairs: forward (UL145F) 5'-ATGTGCACGGACCCGAGA-3'; reverse (UL145R) 5'-TCAATCATCACTTCCACCCATCGAG-3' and Forward (UL147) 5'-ATGGTGCTAACATGGTTGCACC-3'; reverse (UL147R) 5'-TCACCAGCGCAGTCTGAAGTGG-3', respectively. When UL146 gene was not amplified with the former primer pair, the region covering from UL145 to UL147 (1481 bp) was amplified with the UL145F and UL147R primers. The conditions of amplification were as follows: denaturation at 94°C for 4 min followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min and elongation at 72°C for 1 min followed by a final extension at 72°C for 7 min.

DNA sequencing

The amplicons were purified from agarose gels using QIAEX II Gel extraction Kit (Qiagen GmmH, Hilden, Germany) according to the manufacturer's instructions. The purified PCR products were sequenced with the BigDye Terminal Cycle sequencing Kit (PE applied Biosystems, Foster City, CA), according to the manufacturer's instructions, and analyzed using the ABI310 automated sequencer (PE Applied Biosystems, Foster City, CA).

Phylogenetic and protein analysis

A phylogenetic tree inferred from the HCMV UL146 ORF was constructed by the neighbor-joining method with 354 bp nucleotides of this gene. The Clustal W and Jalview software servers were used. TreeView software was used to display the phylogenetic tree. NetNGlyc 1.0 Server was used to find

glycosylation sites and NetPhos 2.0 Server was used to find phosphorylation sites in the vCXCL-1 protein.

RESULTS

UL146 gene amplification and sequencing in clinical strains

Only 17/56 (30.4%) strains were successfully amplified by PCR when the primer pair to HCMV UL146_{Tol} ORF was used. Another 18 strains were amplified using UL145F and UL147R primer pair. Thus, a total of 21 clinical strains were not successfully amplified when the herein proposed experimental strategy was used. Nonetheless, all the clinical strains were positive for IE1 gene amplification, demonstrating the presence of viral DNA (data not shown). Amplicons were sequenced, confirming that they corresponded to HCMV isolates. Phylogenetic analysis clustered the 35 isolated HCMV strains into 8 distinct groups (Fig.1). The UL146_{Tol}, group 1 proved to be the most prevalent among the Japanese clinical isolates (14 strains), followed by group 8 (4 strains), group 3 (4 strains), group 5 (4 strains), group 6 (3 strains), group 4 (3 strains), group 2 (2 strains) and group 7 (1 strain). Further analysis revealed that the ratios of sequence variability among the different groups ranged from 25.1% to 52.9% (Table I).

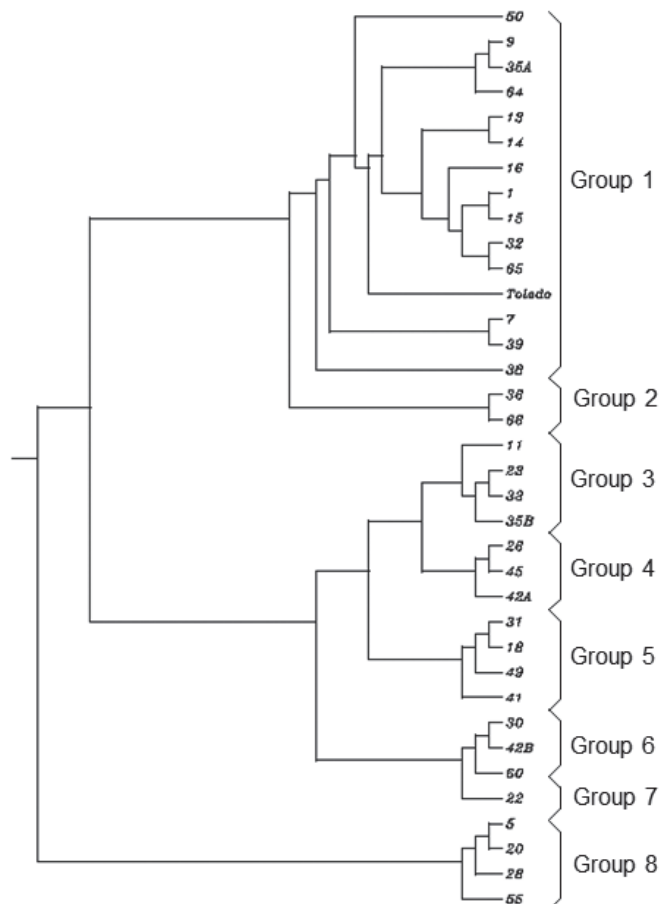


Figure 1: Phylogenetic tree based on UL146 DNA sequences.

Table I
Nucleotide and amino-acid variability among different UL146 groups

	G1	G2	G3	G4	G5	G6	G7	G8
G1	0.80* (2.6) [§]	28.8 (38.3)	46.1 (58.5)	44.8 (53.5)	45.7 (54.3)	48.7 (58.8)	52.1 (62.9)	50.5 (67.0)
G2		0 (0)	46.2 (57.0)	45.9 (57.0)	44.8 (56.6)	45.9 (66.7)	49.2 (64.2)	52.8 (64.0)
G3			0.60 (0.80)	33.0 (41.9)	31.9 (40.7)	38.4 (53.6)	36.4 (43.5)	52.4 (60.4)
G4				0.60 (1.50)	28.7 (42.1)	32.8 (44.1)	28.9 (38.1)	49.3 (63.2)
G5					0 (0)	37.6 (50.5)	32.7 (40.6)	50.3 (62.4)
G6						0 (0)	25.1 (34.5)	52.9 (54.5)
G7							2.1 (2.80)	52.1 (61.7)
G8								0 (0)

* Nucleotide variability; [§] Amino acid variability

The vCXCL-1 sequence analysis

Sequence analysis further revealed that the ratios of amino acid variability among the 8 distinct groups ranged from 34.5% to 67%. The predicted amino acid sequences were compared among 8 representative strains of each individual group (Table II). The estimated size of the vCXCL-1 protein ranged from 97 to 125 amino acids (Fig. 2). The UL146_{Toledo} group contained the ELRCRC motif. Two strains belonging to group 2 contained the ELRCCK motif. All 16 strains belonging to groups 3, 4, 5, 6 and 7 contained the ELRCPC motif, although the variability at the amino acid level among these groups ranged from 34.5% to 66.7%. The 4 strains belonging to group 8 contained the CXC motif, but lacked the ELR motif. The variability between group 8 and other groups at the amino acid level ranged from 54.5% to 67%, suggesting a clear difference in the protein sequence between group 8 and the other groups of this study. The bioinformatics analysis of post-translational modifications showed that the sequence of group 1 and 2 exhibit three possible glycosylation sites; groups 3 to 6 do not have any predictable glycosylation sites, while groups 7 and 8 have only one possible glycosylation site. The analysis for possible serine, threonine and tyrosine phosphorylation sites were different in all 8 groups (Table III).

DISCUSSION

It was previously reported that the HCMV Toledo strain has 19 ORF between UL133 and UL151 in the UL/b' region. Strikingly, these OFR are not present in AD169, even in highly passaged attenuated laboratory viral strains (Cha et al., 1996). To us, these observations suggest that these additional genes might be candidate pathogenic markers of HCMV infection. Inside the UL/b' region, the UL146 gene that encodes a CXC-chemokine is highly variable in epidemiological studies. Thus, we were interested in analyzing the diversity of UL146 of clinical isolates of HCMV from Japan using PCR followed by UL146 sequencing. The phylogenetic analysis of the UL146 gene showed that is highly polymorphic and variants can be clustered in eight sequence groups (Fig. 1). These groups, except for four

strains corresponding to group 8, exhibited the conserved ELR/CXC domain (Fig. 2). Group 8 displayed the CTC domain, suggesting that the physiological function of these proteins may differ from that of the other seven groups. Previously, Dolan et al. clustered UL146 sequences in 14 groups (Dolan et al., 2004), so we aligned these sequences with those reported here. Seven UL146 sequences at the amino acid level reported by Dolan can be clustered in seven groups reported in this study (group 1: Toledo; group 2: KSG, group 3: KM; group 5: Towne; group 6: FS; group 7: TB40; group 8: Davis). In addition, another six sequences cannot be clustered in the groups reported in this study (RK, AL, 6397, ML1, NT and Merlin).

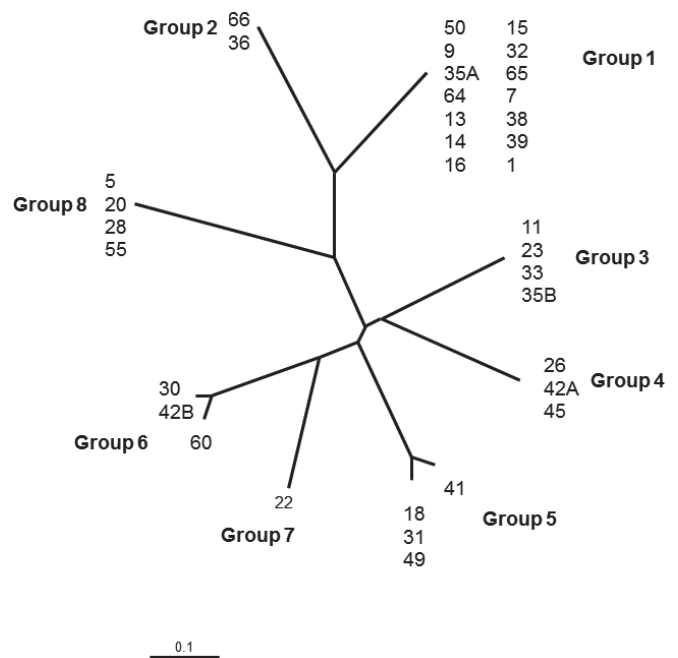


Figure 2: Non-rooted phylogenetic tree inferred from vCXCL-1 amino acid sequences.

Table II

Alignment of the HCMV-encoded chemokine vCXCL-1 obtained from 35 HCMV strains from Japan

		25	50	
Toledo	M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
G1	15 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	32 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	13 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	14 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	1 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	65 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	16 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	9 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H K - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	35A M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	64 M R L I F G A L I - I F L A Y V Y R Y E V N G T E L R - - C R C L H K - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	38 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	39 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	7 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
	50 M R L I F G A L I - I F L A Y V Y H Y E V N G T E L R - - C R C L H R - K W P P N K I I L G N Y W L H R D - P R G P G C D K N E H L L Y P D			
G2	36 M R F I F G L L I - I F L A Y M Y Y Y E V N G T E L R - - C K C P G D K K L P R T R I M L G D F W T H R E - S G G P G C N G Y Q Y L L Y F N			
	66 M R F I F G L L I - I F L A Y M Y Y Y E V N G T E L R - - C K C P G D K K L P R T R I M L G D F W T H R E - S G G P G C N G Y Q Y L L Y F N			
G3	23 M R F I F G L L I G L V I V Y T Y Y Y E V Q S T E L R - - C P C T N - - - G L H D P L Y G I F Y A G R D P P R P P G C E K D Q Y Y L K P P			
	35B M R F I F G L L I G L V I V Y T Y Y Y E V Q S T E L R - - C P C T N - - - G L H D P L Y G I F Y A G R D P P R P P G C E K D Q Y Y L K P P			
	33 M R F I F G L L I G L V I V Y T Y Y Y E V Q S T E L R - - C P C T N - - - G L H D P L Y G I F Y A G R D P P R P P G C E K D Q Y Y L K P P			
	11 M R F I F G L L I G L V I V Y T Y Y Y E V Q S T E L R - - C P C T N - - - G L H D P L Y G I F Y A G R D P P R P P G C E K D Q Y Y L K P P			
G4	26 M R L I F G S L I S I L I V C - - Y Y G V E S V E L R - - C P C G S N - - - G L R D P L S G V F L V G R D P P R P P G C T K Y Q H Y L V P P			
	45 M R L I F G S L I S I L I V C - - Y Y G V E S V E L R - - C P C G S N - - - G L R D P L S G V F L V G R D P P R P P G C T K Y Q H Y L V P P			
	42A M R L I F G S L I S I L I V C - - Y Y G V E S V E L R - - C P C G S N - - - G L R D P L S G V F L V G R D P P R P P G C T K Y Q H Y L V P P			
G5	49 M R L I F S S L I S L M A F M Y Y H G V H S R E L R - - C P C T H K - - - A L H R P I G G L F W V G R D P P N P P E C D K P Q H Y L L P P			
	31 M R L I F S S L I S L M A F M Y Y H G V H S R E L R - - C P C T H K - - - A L H R P I G G L F W V G R D P P N P P E C D K P Q H Y L L P P			
	18 M R L I F S S L I S L M A F M Y Y H G V H S R E L R - - C P C T H K - - - A L H H P I G G L F W V G R D P P N P P E C D K P Q H Y L L P P			
	41 M R L I F S S L I S L M A F M Y D H G V H S R E L R - - C P C T H K - - - A L H H P I G G L F W V G R D P P N P P E C D K P Q H C L L P P			
G6	30 M R F I F S L F G - - L L I A L Y Y K V E S V E L R - - C P C G S N - - - G L S L P I G G F F W I G Y N P P D P P K C E K P Q H F L L P P			
	42B M R F I F S L F G - - L L I A L Y Y K V E S V E L R - - C P C G S N - - - G L S L P I G G F F W I G Y N P P D P P K C E K P Q H F L L P P			
	60 M R L I F S L F G - - L L I A L D Y K V E S V E L R - - C P C G S N - - - G S S L P I R G S F W I G Y N P P H P P K C E K P Q H F L L P P			
G7	22 M R F I F S L F G - - L L I A L C Y K V E S V E L R - - C R C S - N - - G S N H P V F G V F W V G Y K P P D P - T C D K T E H F L L P P			
G8	5 M R L I F G L L I - I F I V T D T C N G G F G T E G N G R C T C I G Y H R L L G - Q L P R G T F W L G H L - P P G S H C P K G Q V M I K I G			
	20 M R L I F G L L I - I F I V T D T C N G G F G T E G N G R C T C I G Y H R L L G - Q L P R G T F W L G H L - P P G S H C P K G Q V M I K I G			
	28 M R L I F G L L I - I F I V T D T C N G G F G T E G N G R C T C I G Y H R L L G - Q L P R G T F W L G H L - P P G S H C P K G Q V M I K I G			
	55 M R L I F G L L I - I F I V T D T C N G G F G T E G N G R C T C I G Y H R L L G - Q L P R G T F W L G H L - P P G S H C P K G Q V M I K I G			
		75	100	125
Toledo	G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - G P R R I N I T L I G V R G - - - - -			
G1	15 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	32 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	13 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	14 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	1 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	65 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	16 G K K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	9 G R K P P P G H G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	35A G R K P P P G P G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	64 G R K P P P G P G V C L S P D H L F S K W L D K R N D N R W Y N V N I T K S P - - - - - - - E P R R I N I T L I G V R G - - - - -			
	38 G R K P P P G P G V C L S P D H L F P - - - - - - - - - - - - - G P R R I N I T L I G V K G - - - - -			
	39 G R K P P P G P G V C L S P D H L F S K W L H K Y N D N R W Y N V N I T K S P - - - - - - - G P R R R N I T L I G V K G - - - - -			
	7 G R K P P P G P G V C L S P D H L F S K W L D K Y N D N R W Y N V N I T K S P - - - - - - - G P R R I N I T L I G V K G - - - - -			
	50 G R K P P P G P G V C L S P D H L F S K W L D K Y N D N R W Y N V N I T K S P - - - - - - - G P R R L N I T L I G V R G - - - - -			
G2	36 N G R H G R G V C L A P D H H I S K W L N T H N D G R W Y N V N I T K Q P G R R T G G R G P G Q V N I T L I A V K Q - - - - -			
	66 N G R H G R G V C L A P D H H I S K W L N T H N D G R W Y N V N I T K Q P G R R T G G R G P G Q V N I T L I A V K Q - - - - -			
G3	23 - - - - K G K A V C L G P H H H L S I W L N G Q N S S L W H K V L V T G K N G - - - - - N G P H V T K K G D F P R G R K N I M I -			
	35B - - - - K G K A V C L G P H H H L S I W L N G Q N S S L W H K V L V T G K N G - - - - - N G P H V T K K G D F P R G R K N I M I -			
	33 - - - - K G K A V C L G P H H H L S I W L N G Q N S S L W H K V L V T G K N G - - - - - N G P H V T K K G D F P R G R K N I M I -			
	11 - - - - K G K A V Y L G P H H H L S I W L N G Q N S S L W H K V L V T G K N G - - - - - N G P H V T K K G D F P R G R K N I M I -			
G4	26 - - - - N G R R A C L N P D H H L S K W L D A Q N S N S W Y K V V V T T G G G - - - - - S G P H V D K K A E I K K R L S - - - - -			
	45 - - - - N G R R A C L N P D H H L S K W L D A Q N S N S W Y K V V V T T G G G - - - - - S G P H V D K K A E I K K R L S - - - - -			
	42A - - - - N G R R A C L N P D H H L S K W L D A Q N S N S W Y K V V V T T G G G - - - - - S G P H V D K K A E I K K R L S - - - - -			
G5	49 - - - - R G K P V C L A P D H H L S K W L D G K K D N S W H K V L V K L K D S - - - - - N G P H V E E S A V T N K R P R W K - - - - -			
	31 - - - - R G K P V C L A P D H H L S K W L D G K K D N S W H K V L V K L K D S - - - - - N G P H V E E S A V T N K R P R W K - - - - -			
	18 - - - - R G K P V C L A P D H H L S K W L D G K K D N S W H K V L V K L K D S - - - - - N G P H V E E S A V T N K R P R W K - - - - -			
	41 - - - - R G K P V C L A P D H H L S K W S D G K K D N S W H K V L V K L K D M - - - - - N G P H V E G - - - - -			
G6	30 - - - - K G K P V C L S P D H V L S K W L H G K L S N T W H K V L L R T K G G - - - - - D G P H V E E R T A S N G R P P W K L K F			
	42B - - - - K G K P V C L S P D H V L S K W L H G K S S N T W H K V L L R T K G G - - - - - D G P H V E E R T A S N G R P P W K L K F			
	60 - - - - K G K P V C L S P D H V L S K W L H G K S S N T W H K V L L R T K G G - - - - - D G P H V E Q R T A S N G R P P W K L K F			
G7	22 - - - - R Q T P V C L S P D H Y L S K W V D G K R S N W W H K V L I K K N T D - - - - - N G P H I E D K S D T N R H P P W R L -			
G8	5 - - - - Q G P I V C L S D H H P L S K W M Y G N H K S G S E T W L Q I K M E - - - - - G P R N A T V V Q R S N T R P - - - - -			
	20 - - - - Q G P I V C L S D H H P L S K W M Y G N H K S G S E T W L Q I K M E - - - - - G P R N A T V V Q R S N T R P - - - - -			
	28 - - - - Q G P I V C L S D H H P L S K W M Y G N H K S G S E T W L Q I K M E - - - - - G P R N A T V V Q R S N T R P - - - - -			
	55 - - - - Q G P I V C L S D H H P L S K W M Y G N H K S G S E T W L Q I K M E - - - - - G P R N A T V V Q R S N T R P - - - - -			

Table III

Glycosylation and phosphorylation site in eight vCXCL-1 groups.
Group Glycosylation Phosphorylation

		Ser	Thr	Tyr
1	3	1	0	0
2	3	1	2	1
3	0	0	1	3
4	0	2	0	1
5	0	2	1	0
6	0	2	0	0
7	1	0	3	1
8	1	2	1	1

The ELR/CXC chemokine motif plays an important role in neutrophils recruitment and promotes angiogenesis (Strieter et al., 2005). The specific absence of the ELR domain has been reported to be angiostatic (Strieter et al., 1995). In this report, we did not find clinical evidence suggesting association between ELR/CXC motif presence and outcome or progression of HCMV infection. In agreement with this observation, Arav- Boger et al. in 2006 reported that UL146 gene is polymorphic in congenitally infected newborns and no association between specific UL146 variant and the outcome of HCMV infection was found (Arav-Boger et al., 2006). Other studies have analysed the association between the variation of glycoprotein genes (gB) and the outcome of HCMV infection, however, results remain highly controversial (Rasmussen et al., 1997; Sarcinella et al., 2002). In our study, we observed differences in glycosylation and phosphorylation sites in the vCXCL-1 protein among the eight groups, suggesting that these proteins might have different functions *in vivo*. However, it was reported that variations in the ELR-CXC motif in UL146 ORF do not alter the functionality of the vCXCL-1 protein (Miller-Kittrell et al., 2007).

The variability of other genes in the UL133-UL151 region of HCMV has been analysed and the function of these genes is currently being investigated. It is suggested that the UL133-UL151 region is important in immune response evasion and might be involved in some of the pathogenic properties ascribed to HCMV (Wang et al., 2005). For instance, the sequence of the UL144 ORF, reported to encode a tumor necrosis factor analogue, amplified from clinical samples clustered in five defined phylogenetic groups (Lurain et al., 1999; Murayama et al., 2005; Picone et al., 2005). The UL139 gene was shown to be highly polymorphic and the UL139 sequences clustered in six groups, in which various nucleotide insertions and non-synonymous substitutions were detected (Qi et al., 2006). UL139 function is unknown; however the protein product had sequence similarity with human CD24, a signal transducer modulating B-cell activation responses (Qi et al., 2006). On the other hand, it was found that UL141 is highly conserved in clinical strains from China (Ma et al., 2006). This protein downregulates a natural killer cell-activating ligand CD155, suggesting that might be involved in NK cell evasion function (Tomasec et al., 2005). In addition, a previous study reported that the UL147 gene is highly polymorphic and the encoded protein displays a CXC motif,

similar to the UL146 gene (Lurain et al., 2006). Albeit variation between HCMV strains was lower when compared to UL146 (He et al., 2006). However, some evidence established a link between these two genes, although no recombination was shown (Lurain et al., 2006). Although the HCMV genome is well conserved, the UL133-UL151 region is polymorphic and especially the UL146 gene. In the case of this protein, the variability among different groups has been reported to be very stable *in vitro* and *in vivo* and no mutations occur easily in infected individuals (Lurain et al., 2006; Stanton et al., 2005). It has been proposed that the variability in the viral chemokines helps to fine tune the neutrophil response, increasing the efficiency of HCMV dissemination (Miller-Kittrell et al., 2007).

In summary, in this report we describe that the HCMV UL146 gene amplified from Japanese clinical isolates is highly polymorphic and phylogenetic analysis clustered the gene sequences in 8 defined groups, suggesting a different function for the vCXCL-1 protein in each distinct group. Additional studies are necessary to characterize the UL146 function in infected cells and the eventual association with outcome or severity of the disease in immunocompromised patients.

ACKNOWLEDGEMENTS

This work was supported by Grants-in-Aid for Scientific Research on Priority Areas of the Ministry of Education, Culture, Sports, Science and Technology, Japan. We thank to Dr. Marcelo Lopez-Lastra from Pontificia Universidad Catolica de Chile for proofreading the manuscript.

REFERENCES

- ARAV-BOGER, R., FOSTER, C. B., ZONG, J. C. & PASS, R. F. (2006). Human cytomegalovirus-encoded alpha -chemokines exhibit high sequence variability in congenitally infected newborns. *J Infect Dis* 193, 788-91.
- BRANTSÆTER, A. B., LIESTOL, K., GOPLEN, A. K., DUNLOP, O. & BRUUN, J. N. (2002). CMV disease in AIDS patients: incidence of CMV disease and relation to survival in a population-based study from Oslo. *Scand J Infect Dis* 34, 50-5.
- BRITT WJ, A. C. (1996). Cytomegalovirus. In *Fields Virology* (ed. K. D. Fields BN, Howly PM), pp. 2493-2523. Lippincott-Raven, Philadelphia.
- CHA, T. A., TOM, E., KEMBLE, G. W., DUKE, G. M., MOCARSKI, E. S. & SPAETE, R. R. (1996). Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J Virol* 70, 78-83.
- DOLAN, A., CUNNINGHAM, C., HECTOR, R. D., HASSAN-WALKER, A. F., LEE, L., ADDISON, C., DARGAN, D. J., MCGEOCH, D. J., GATHERER, D., EMERY, V. C., GRIFFITHS, P. D., SINZGER, C., MCSHARRY, B. P., WILKINSON, G. W. & DAVISON, A. J. (2004). Genetic content of wild-type human cytomegalovirus. *J Gen Virol* 85, 1301-12.
- HASSAN-WALKER, A. F., OKWUADI, S., LEE, L., GRIFFITHS, P. D. & EMERY, V. C. (2004). Sequence variability of the alpha-chemokine UL146 from clinical strains of human cytomegalovirus. *J Med Virol* 74, 573-9.
- HE, R., RUAN, Q., QI, Y., MA, Y. P., HUANG, Y. J., SUN, Z. R. & JI, Y. H. (2006). Sequence variability of human cytomegalovirus UL146 and UL147 genes in low-passage clinical isolates. *Intervirology* 49, 215-23.
- HIRT, B. (1967). Selective extraction of polyoma DNA from infected mouse cell cultures. *J Mol Biol* 26, 365-9.
- LURAIN, N. S., FOX, A. M., LICHY, H. M., BHORADE, S. M., WARE, C. F., HUANG, D. D., KWAN, S. P., GARRITY, E. R. & CHOU, S. (2006). Analysis of the human cytomegalovirus genomic region from UL146 through UL147A reveals sequence hypervariability, genotypic stability, and overlapping transcripts. *Virol J* 3, 4.

- LURAIN, N. S., KAPELL, K. S., HUANG, D. D., SHORT, J. A., PAINTSIL, J., WINKFIELD, E., BENEDICT, C. A., WARE, C. F. & BREMER, J. W. (1999). Human cytomegalovirus UL144 open reading frame: sequence hypervariability in low-passage clinical isolates. *J Virol* 73, 10040-50.
- LUTTICHAU, H. R. (2010). The cytomegalovirus UL146 gene product vCXCL1 targets both CXCR1 and CXCR2 as an agonist. *J Biol Chem* 285, 9137-46.
- MA, Y. P., RUAN, Q., HE, R., QI, Y., SUN, Z. R., JI, Y. H., HUANG, Y. J., LIU, Q., CHEN, S. R. & WANG, J. D. (2006). Sequence variability of the human cytomegalovirus UL141 Open Reading Frame in clinical strains. *Arch Virol* 151, 827-35.
- MILLER-KITTRELL, M., SAI, J., PENFOLD, M., RICHMOND, A. & SPARER, T. E. (2007). Functional characterization of chimpanzee cytomegalovirus chemokine, vCXCL-1(CCMV). *Virology* 364, 454-65.
- MURAYAMA, T., TAKEGOSHI, M., TANUMA, J. & EIZURU, Y. (2005). Analysis of human cytomegalovirus UL144 variability in low-passage clinical isolates in Japan. *Intervirology* 48, 201-6.
- PENFOLD, M. E., DAIRAGHI, D. J., DUKE, G. M., SAEDERUP, N., MOCARSKI, E. S., KEMBLE, G. W. & SCHALL, T. J. (1999). Cytomegalovirus encodes a potent alpha chemokine. *Proc Natl Acad Sci U S A* 96, 9839-44.
- PICONE, O., COSTA, J. M., CHAIX, M. L., VILLE, Y., ROUZIQUX, C. & LERUEZ-VILLE, M. (2005). Human cytomegalovirus UL144 gene polymorphisms in congenital infections. *J Clin Microbiol* 43, 25-9.
- PUCHHAMMER-STOCKL, E. & GORZER, I. (2006). Cytomegalovirus and Epstein-Barr virus subtypes the search for clinical significance. *J Clin Virol* 36, 239-48.
- QI, Y., MAO, Z. Q., RUAN, Q., HE, R., MA, Y. P., SUN, Z. R., JI, Y. H. & HUANG, Y. (2006). Human cytomegalovirus (HCMV) UL139 open reading frame: Sequence variants are clustered into three major genotypes. *J Med Virol* 78, 517-22.
- RASMUSSEN, L., HONG, C., ZIPETO, D., MORRIS, S., SHERMAN, D., CHOU, S., MINER, R., DREW, W. L., WOLITZ, R., DOWLING, A., WARFORD, A. & MERIGAN, T. C. (1997). Cytomegalovirus gB genotype distribution differs in human immunodeficiency virus-infected patients and immunocompromised allograft recipients. *J Infect Dis* 175, 179-84.
- SARCINELLA, L., MAZZULLI, T., WILLEY, B. & HUMAR, A. (2002). Cytomegalovirus glycoprotein B genotype does not correlate with outcomes in liver transplant patients. *J Clin Virol* 24, 99-105.
- STANTON, R., WESTMORELAND, D., FOX, J. D., DAVISON, A. J. & WILKINSON, G. W. (2005). Stability of human cytomegalovirus genotypes in persistently infected renal transplant recipients. *J Med Virol* 75, 42-6.
- STRIETER, R. M., BURDICK, M. D., GOMPERS, B. N., BELPERIO, J. A. & KEANE, M. P. (2005). CXC chemokines in angiogenesis. *Cytokine Growth Factor Rev* 16, 593-609.
- STRIETER, R. M., POLVERINI, P. J., KUNKEL, S. L., ARENBERG, D. A., BURDICK, M. D., KASPER, J., DZUIBA, J., VAN DAMME, J., WALZ, A., MARRIOTT, D. & ET AL. (1995). The functional role of the ELR motif in CXC chemokine-mediated angiogenesis. *J Biol Chem* 270, 27348-57.
- TOMASEC, P., WANG, E. C., DAVISON, A. J., VOJTESEK, B., ARMSTRONG, M., GRIFFIN, C., MCSHARRY, B. P., MORRIS, R. J., LLEWELLYN-LACEY, S., RICKARDS, C., NOMOTO, A., SINZGER, C. & WILKINSON, G. W. (2005). Downregulation of natural killer cell-activating ligand CD155 by human cytomegalovirus UL141. *Nat Immunol* 6, 181-8.
- WANG, W., TAYLOR, S. L., LEISENFELDER, S. A., MORTON, R., MOFFAT, J. F., SMIRNOV, S. & ZHU, H. (2005). Human cytomegalovirus genes in the 15-kilobase region are required for viral replication in implanted human tissues in SCID mice. *J Virol* 79, 2115-23.