Human cytomegalovirus *UL145* gene is highly conserved among clinical strains

ZHENGRONG SUN[1], YING LU[2], QIANG RUAN*[,1], YAOHUA JI[1], RONG HE[1], YING QI[1], YANPING MA[1] and YUJING HUANG[1]

[1] *Virus Laboratory, Affiliated ShengJing Hospital, China Medical University, Shenyang 110004, P. R. China*
[2]*Department of Microbiology and Immunology, Jinzhou Medical University, Jinzhou 121001, P. R. China*

**Corresponding author (Fax, +86-24-23892617; Email, ruanq@cmu2h.com)*

Human cytomegalovirus (HCMV), a ubiquitous human pathogen, is the leading cause of birth defects in newborns. A region (referred to as UL/b′) present in the Toledo strain of HCMV and low-passage clinical isolates) contains 22 additional genes, which are absent in the highly passaged laboratory strain AD169. One of these genes, *UL145* open reading frame (ORF), is located between the highly variable genes *UL144* and *UL146*. To assess the structure of the *UL145* gene, the *UL145* ORF was amplified by PCR and sequenced from 16 low-passage clinical isolates and 15 non-passage strains from suspected congenitally infected infants. Nine *UL145* sequences previously published in the GenBank were used for sequence comparison. The identities of the gene and the similarities of its putative protein among all strains were 95.9–100% and 96.6–100%, respectively. The post-translational modification motifs of the UL145 putative protein in clinical strains were conserved, comprising the protein kinase C phosphorylation motif (PKC) and casein kinase II phosphorylation site (CK-II). We conclude that the structure of the *UL145* gene and its putative protein are relatively conserved among clinical strains, irrespective of whether the strains come from patients with different manifestations, from different areas of the world, or were passaged or not in human embryonic lung fibroblast (HELF) cells.

[Sun Z, Lu Y, Ruan Q, Ji Y, He R, Qi Y, Ma Y and Huang Y 2007 Human cytomegalovirus *UL145* gene is highly conserved among clinical strains; *J. Biosci.* **32** 1111–1118]

1. Introduction

Human cytomegalovirus (HCMV) is an important pathogen capable of causing infections that persist lifelong, and normally remain asymptomatic. However, severe diseases may develop resulting in a wide variety of clinical manifestations (Ho 1990). The mechanisms that determine the type, duration and severity of clinical manifestations are poorly understood. The study on genetic and immunological variability, as well as differences in growth characteristics *in vitro* is the basic project, since strain differences may affect HCMV virulence (Bale *et al* 2001; Brown *et al* 1995; Woodroffe *et al* 1997).

Clinical strains of HCMV display genetic polymorphisms, possibly related to strain-specific tissue tropism and HCMV-induced immunopathogenesis. A substantial portion of the HCMV genome encodes proteins with the potential to affect virulence through cell tropism, immune evasion, molecular mimicry, or interference with host chemokines (Hahn *et al* 2004; Penfold *et al* 1999). The large coding capacity and slow cell-associated replication cycle of HCMV suggest that pathogenesis involves a complex interaction of

Keywords. Human cytomegalovirus; *UL145* ORF

Abbreviations used: HCMV, human cytomegalovirus; ORF, open reading frame;gB, glycoprotein B; PCR, polymerase chain reaction; HD, Hirschsprung disease; HELF, human embryonic lung fibroblasts; PKC, protein kinase C phosphorylation motif; CK II, casein kinase II phosphorylation site; TNF, tumour necrosis factor

viral proteins with multiple host cell targets (Rasmussen *et al* 2003). It has been indicated that HCMV disease and cell tropism may be associated with inter-strain sequence variability, based mainly on the four glycoprotein B (gB) genotypes (Wirgart *et al* 1998). However, the linkage between the outcome of virus infection and gB genotype is not conclusive, suggesting that other gene products of HCMV may also have a role to play in pathogenesis. The UL/b′ region of the HCMV genome was found in the Toledo and several other low-passage clinical isolates, but it was absent in the laboratory AD169 strain (X17403) (Cha *et al* 1996). The fact that AD169 shows attenuated virulence and different tropism for endothelial cells from low-passage isolates suggests that the predicted products of *UL/b'* genes may determine the manifestations of HCMV infection *in vivo*. Several genes such as *UL139*, *UL141* and *UL149* ORFs have been analysed for genomic variation in clinical isolates (Qi *et al* 2006; Ma *et al* 2006; Ji *et al* 2006).

Putative functions of these UL/b′ ORFs have so far been ascribed to *UL144*, *UL146* and *UL147* (Lurain *et al* 1999; Benedict *et al* 1999; Penfold *et al* 1999). The *UL144* gene encodes a structural homologue of the herpesvirus entry mediator, a viral tumour necrosis factor (TNF) receptor family member, and the *UL146* and *UL147* genes encode viral CXC (a) chemokines. The *UL146* protein is a neutrophil attractant and its recombinant protein can serve as a fully functional chemokine able to induce calcium mobilization, chemotaxis and neutrophil degranulation. To determine the structure of the *UL145* gene, the *UL145* ORF was sequenced and analysed by using 31 clinical strains obtained from infants with suspected congenital HCMV infection in the present study.

2. Materials and methods

2.1 *HCMV strains*

The study material consisted of 16 HCMV clinical isolates (named as low-passage isolates) and 15 HCMV strains in original samples (named as non-passage strains) from infants suspected to be congenitally infected. The infants suffered from jaundice (n=8 samples), microcephaly (n=4 samples) and megacolon (Hirschsprung disease [HD], n=19 samples) (table 1). Among the 19 strains from infants with HD, 15 were non-passage strains and 4 were low-passage isolates. The main manifestation of infants in the jaundice group was hepatic dysfunction, such as elevation of direct bilirubin, accompanied by increased transaminase levels. Besides the specific clinical manifestations, infants with microcephaly and HD showed a rise in indirect bilirubin. Strains from infants with HD were obtained from the abnormal colon tissue or urine samples, and the other strains were isolated from urine samples. The patients were 1–14 months of age, with an average age of 5.5+3.8 months. Clinical isolates were passaged less than 10 times in human embryonic lung fibroblasts (HELF). All clinical strains were proved to contain detectable HCMV DNA by the diagnostic fluorescent-quantified PCR method (He *et al* 2001). The clinical strains were stored at –70°C until use.

2.2 *Amplification of UL145 genes*

HCMV *UL145*-specific primers were designed on the basis of the Toledo sequence (Accession number: U33331, Cha *et al* 1996). The sequence of primers used for PCR amplification was designated as *UL145*-F (forward) 5′CCGTGTCTGAGTATCTGTGG3′ and *UL145*-R (reverse) 5′CGTAATATCCGGTATTCCC3′. Virus DNA was exposed by boiling the infected cells with lysis buffer for 15 min. The *UL145* gene was amplified as follows: 5 μl DNA was added to a PCR reaction mixture containing 1X buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 150 ng of primers and 0.5 U of Taq polymerase (Promega Company) to a final volume of 50 μl. Amplification was carried out with a PE Applied Biosystems Geneamp PCR System 2400 (Perkin-Elmer, USA), and negative controls were included in each round of amplification. A 754 bp product that encompassed the *UL145* ORF together with the flanking sequence was obtained.

2.3 *Purification of PCR products and sequencing*

PCR products were monitored by agarose gel electrophoresis, and the appropriate band was purified from low melting-point agarose gels (Promega Company). Briefly, amplified DNA fragments were separated on a 1.5% agarose gel. DNA bands were visualized by ethidium bromide staining and ultraviolet light. The appropriate band was excised and the amplicon was eluted in 30 μl of nuclease-free water. Concentrated PCR products containing the whole length of the *UL145* ORF were sequenced directly with the BigDye Terminators Cycle Kit. The sequenced products were analysed on an ABI 3700 automated sequencer. The *UL145* sequences of all strains were sequenced bilaterally to ensure each nucleotide.

2.4 *Sequence analysis*

Analysis of nucleotide and deduced amino acid sequences was performed with the BioEdit 7.0 software. Similarity was calculated by scoring for both identical amino acids and conservative amino acid substitutions. The consensus sequence was established on a simple majority of aligned reading, and nucleotides that differed from the consensus sequence were examined individually. To generate the

Table 1. Detailed information about human cytomegalovirus (HCMV) strains

Strain	Source*	Location	Passage	GenBank	Reference accession No.
1M, 2J, 8J, 9J, 10J, 13J	urine from congenitally infected infants	China	low-passage	AF489966–AF489971 AF489973, AF489974	this study
16M, 21M, 20M, 33J	urine from congenitally infected infants	China	low-passage	AY941166, AY941169 AY941170, AY941172	this study
27C, 29C, 32C, 51C	tissue from congenitally infected infants	China	low-passage	AY941167, AY941168, AY941171, AF489972	this study
U253, U257,	urine from congenitally infected infants	China	non-passage	DQ381584, DQ381585	this study
T8, T9, T11, T16, T25, T27, T28 T38, T39,T45, T49, T50, T51	tissue from congenitally infected infants	China	non-passage	DQ381572–DQ381583, DQ381586	this study
3301	urine from a congenitally infected infant	Cardiff	non-passage	AY446864	Davison *et al* 2003
W	tissue from an infected AIDS patient	London	non-passage	AY446865	Davison *et al* 2003
Toledo	urine from a congenitally infected infant	unknown	low-passage	AY446871	Dolan *et al* 2004
Merlin	urine from a congenitally infected infant	Cardiff	low-passage	AY446894	Dolan *et al* 2004
6397	urine from a congenitally infected infant	Cardiff	low-passage	AY446870	Davison *et al* 2003
3157	urine from a congenitally infected infant	Cardiff	low-passage	AY446867	Davison *et al* 2003
Towne	urine from a congenitally infected infant	unknown	high-passage	AY446869	Plotkin *et al* 1975
Davis	liver biopsy from a congenitally infected infant	unknown	high-passage	AY446868	Craig *et al*1957
TB40	Throat washing from a bone marrow recipient unknown	unknown	high-passage	AY446866	Sinzger *et al* 2000

*All the strains were from different patients

phylogenetic tree, a neighbour-joining bootstrap analysis was performed with the PHYLIP software package, and the final amino acid tree was rendered with Treeview. Bootstrap calculations were based on 1000 repetitions. Functional motifs of the predicted *UL145* protein were identified from the PROSITE database. Nine previously published *UL145* sequences in GenBank were used for sequence comparison with those of the clinical strains. Their accession numbers in GenBank are: AY446864–AY446871, AY446894.

3. Results

3.1 *Presence of the UL145 genes in clinical strains*

UL145 genes were successfully amplified and sequenced in all of the 31 clinical strains. The results indicated that the *UL145* gene exists in all of the low-passage isolates and non-passage strains. The sequences obtained have been deposited with GenBank under accession numbers AF489966–AF489974, AY941166–AY941172 and DQ381572–DQ381586. Detailed descriptions of the HCMV strains used in this paper are listed in table 1.

3.2 *Sequence analysis of the UL145 genes of clinical strains*

As in those of previously published strains, all of the *UL145* ORFs of the clinical strains in this study were identical in size, composed of 393 nucleotides. The distance of each paired sequence was calculated from the nucleotide alignments to evaluate the level of variation. The *UL145* sequence of strains obtained in this study shared 95.9–100% identity, while those of previously published strains shared 96.1–98.2% identity. Compared with the consensus sequence, the number of mutant nucleotides within the whole coding region of individual strains ranged from

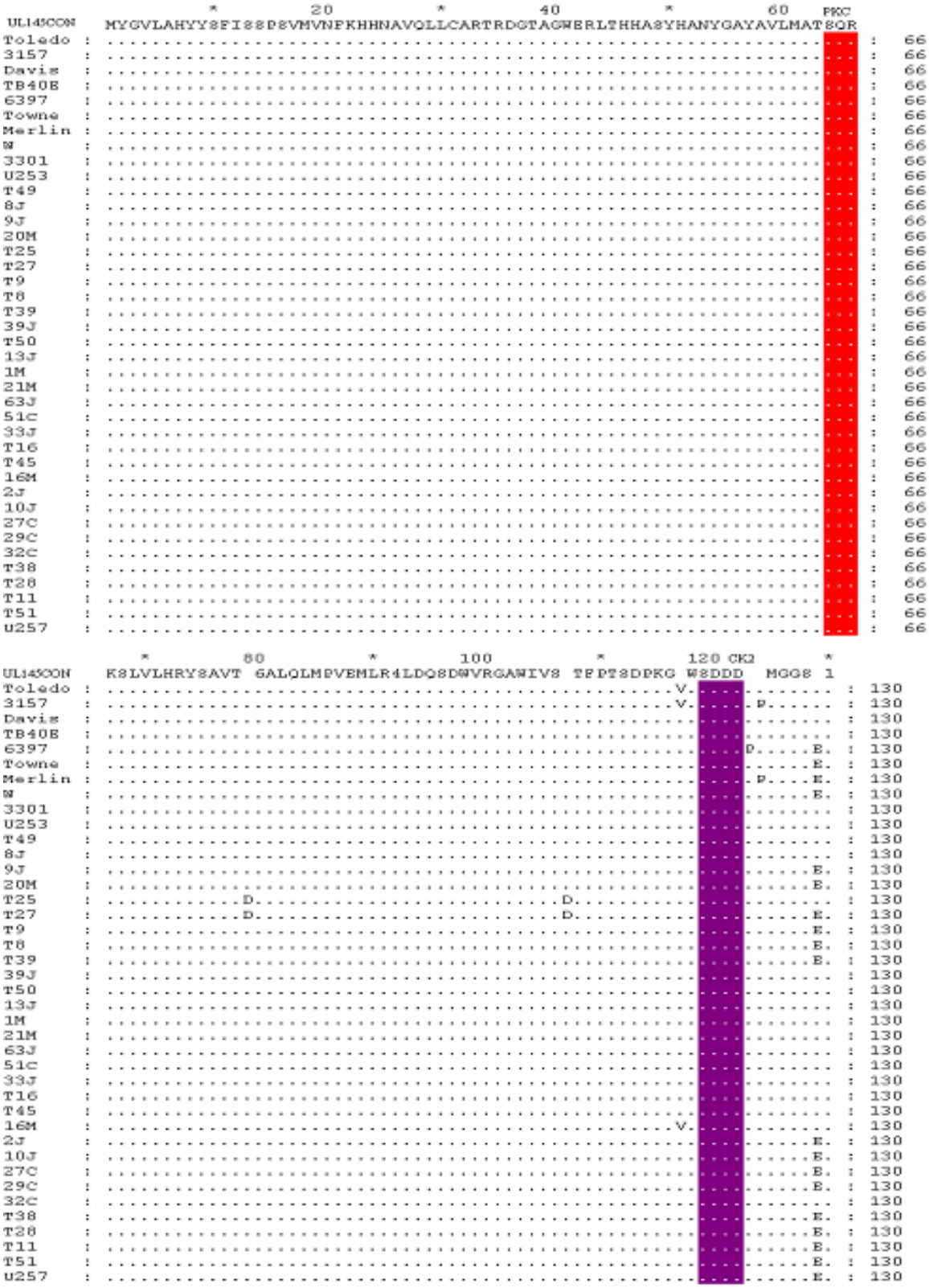

Figure 1. For caption, *see* p. 1115.

9 to 18 (average 12). Compared with that of the Toledo strain, corresponding sequences of *UL145* in the clinical strains have A to G transition at position 90. All clinical strains display C to T transition at position 162 and G to T transversion at 349 except isolate 16M. Clinical strains T49, 8J, 20M, T25, T27, T39 have C to T transition at positions 42, 78, 159; G to A transition at positions 81, 297; A to G transition at position 189, T to C transition at position 333. Almost all the nucleotide substitutions are non-synonymous except those at positions 349 and 387. Among the variant nucleotides, more than 80% were translationally silent (data not shown).

3.3 *Structural analysis of the predicted UL145 protein*

The *UL145* ORF has the potential ability to encode a protein of 130 amino acid residues. The alignment of the deduced amino acid sequences showed a high level of similarity (figure 1). Among all strains, the similarity score varied from 96.6 to 100%, while those of the four previously published strains—6397, W, Merlin and Towne—were more similar to each other (99.3–100%). Each strain had less than 4 amino acid mutations compared with the consensus sequence. Among the strains obtained in this study, consistent variations were found in strains 20M, T27, T9, T8, T39, 2J, 10J, 27C, 29C, T38, T28, T11, T51, U257 at amino acid position 117.

The cysteine residues in the *UL145* putative proteins of the clinical strains were highly conserved, which is important in determining the structure and characteristics of proteins.

Post-translational modification sites of *UL145* predicted proteins were analysed with the Genedoc Software and PROSITE database. The result revealed that the protein kinase C phosphorylation motif (PKC) at residues 65–68 and the putative casein kinase phosphorylation site (CK-II) at residues 121–124 were conserved in *UL145* putative proteins of all the clinical strains and 9 previously published strains (figure 1).

3.4 *Phylogenetic analysis*

Alignment of the *UL145* sequences showed that there was no obvious gene mutation in the *UL145* genes of the 16 low-passage isolates compared with those of the 15 non-passage strains. Among the 19 clinical strains derived from infants with HD, no obvious gene mutations were found between 4 low-passage isolates and 15 non-passage strains. These results indicated that limited passage of HCMV strains in HELF cells could not be enough to cause sequence mutation of the *UL145* gene.

A phylogenetic tree (figure 2) was constructed with the deduced amino acid sequences of the putative *UL145* proteins obtained in this study and those of previously published sequences. Strains from infants with similar clinical manifestations were not clustered into the same branch of the tree. The four strains from Cardiff (6397, Merlin, 3157 and 3301) were not clustered into one group. All of the 15 non-passage clinical strains and 2 non-passage reference strains (3301 and W) were dispersed in different branches. The Cluster dendrogram revealed that the clustering of the amino acid sequences of strains was not correlated with their geographical origin.

4. Discussion

In the HCMV UL/b′ region, the *UL146* gene has been characterized to encode a CXC chemokine that interferes with the recruitment of polymorphonuclear leukocytes at the site of infection (Penfold *et al* 1999; Saedrup *et al* 2002). Moreover, a TNF receptor-like protein encoded by *UL144* has been reported to play a potential role in viral infectivity (Benedict *et al* 1999; Lurain *et al* 1999).

HCMV UL/b′ is a variable gene region. The *UL144* and *UL146* genes have been demonstrated to be highly polymorphic among field isolates from different patients (Hassan-Walker *et al* 2004; Lurain *et al* 1999; Penfold *et al* 1999). We have previously reported that the HCMV *UL144* and *UL146* sequences were highly variable in clinical strains from China. According to their sequence variability, both of the genes were classified into 3 different genotypes (He *et al* 2004; He *et al* 2006). However, little is known about the variability of the *UL145* gene in field isolates. To answer this question, we analysed the *UL145* sequence from 31 clinical strains. Unlike its neighbour genes, *UL144* and *UL146*, the *UL145* gene is relatively conserved in all clinical strains, irrespective of whether the strains come from patients with different manifestations, or from different parts of the world.

The complex nature of the interaction of HCMV with the host immune response implies that more than one variant gene may be involved in the observed difference in seroreactivity, infectivity and pathogenicity. An individual HCMV strain should be defined by the combination of the multiple variant genes that it encodes. Although variable genes could determine the infectivity and pathogenecity of the virus, conserved genes would be more important

Figure 1. Amino acid alignment and post-translational modification motifs of the HCMV UL145 putative protein. Amino acids that differ from the consensus sequence are indicated. The post-translational modification motifs were marked by abbreviations. PKC, protein kinase C phosphorylation motif; CK2, casein kinase II phosphorylation motif.

Figure 2. Phylogenetic analysis based on the UL145 amino acid sequences of the HCMV clinical strains and 9 previously published sequences.

for its survival. High-level conservation has been observed for some important HCMV proteins, such as *UL54* DNA polymerase and *UL97* phosphotransferase (Chou *et al* 1999; Lurain *et al* 2001).

The HCMV infection level of unstimulated endothelial cells is influenced by the basal level of PKC; and stimulation of PKC prior to infection results in an increase of infection by HCMV. HCMV circumvents the NF-κB route in favour

of the PKC-dependent mitogen-activated protein kinase pathway in epithelial cells (Drunen *et al* 1997; Cinatl *et al* 2001). On the other hand, packaging of CK-II into the HCMV virion shows that diverse molecular mechanisms are utilized by HCMV for rapid NF-κB activation in monocytes and endothelial cells infected *in vivo*. The virus–host cell interaction might be a mechanism that promotes HCMV persistence in immune-privileged organs (Nogalski *et al* 2007). Our data show that the putative PKC and CK II sites are highly conserved in *UL145* putative proteins of clinical strains. The relative conservation of *UL145* sequences being retained under selection pressure suggests that it might be essential for virus survival and replication *in vivo*.

In summary, the HCMV *UL145* gene was found to be relatively conserved in all types of clinical strains, irrespective of whether the strains came from different patients, sources of samples and different parts of the world, or from passaged and non-passaged strains.

Acknowledgements

We thank Qing Liu and ShuRong Chen for identifying HCMV-infected infants and providing HCMV isolates, and ShouGuang Jin (Molecular Genetics and Microbiology, University of Florida) for his helpful comments on the manuscript. The research was funded by the National Natural Science Foundation of China (30371492).

References

Bale J F Jr, Petheram S J, Robertson M, Murphy J R and Demmler G 2001 Human cytomegalovirus A sequence and *UL144* variability in strains from infected children; *J. Med.. Virol.* **65** 90–96

Benedict C A, Butrovich K D, Lurain N S, Corbeil J, Rooney I, Schneider P, Tschopp J and Ware C F 1999 Cutting edge: a novel viral TNF receptor superfamily member in virulent strains of human cytomegalovirus; *J. Immunol.* **162** 6967–6970

Brown J M, Kaneshima H and Mocarski E S 1995 Dramatic inter-strain differences in the replication of human cytomegalovirus in SCID-hu mice; *J. Infect. Dis.* **171** 1599–1603

Cha T A, Tom E, Kemble G W, Duke G M, Mocarski E S and Spaete R R 1996 Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains; *J. Virol.* **70** 78–83

Chou S, Lurain N S, Weinberg A, Cai G Y, Sharma P L and Crumpacker C S 1999 Interstrain variation in the human cytomegalovirus DNA polymerase sequence and its effect on genotypic diagnosis of antiviral drug resistance. Adult AIDS Clinical Trials Group CMV Laboratories; Antimicrob. Agents Chemother. 43 1500–1502

Cinatl J Jr, Margraf S, Vogel J U, Scholz M, Cinatl J and Doerr H 2001 Human cytomegalovirus circumvents NF-κB dependence in retinal pigment epithelial cells; *J. Immunol.* **23** 1900–1908

Craig J M, Macauley J C, Weller T H and Wirth P 1957 Isolation of intranuclear inclusion producing agents from infants with illnesses resembling cytomegalic inclusion disease; *Proc. Soc. Exp. Biol. Med.* **94** 4–12

Davison A J, Akter P, Cunningham C, Dolan A, Addison C, Dargan D J, Hassan-Walker A F, Emery V C, Griffiths P D and Wilkinson G W 2003 Homology between the human cytomegalovirus *RL11* gene family and human adenovirus *E3* genes; *J. Gen. Virol.* **84** 657–663

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 and Davison A J 2004 Genetic content of wild-type human cytomegalovirus; *J. Gen. Virol.* **85** 1301–1312

Drunen M, Vossen R, Couwenberg F, Hulsbosch M, Heemskerk J, Dam-Mieras M and Bruggeman C 1997 Activation of protein kinase C enhances the infection of endothelial cells by human cytomegalovirus; *Virus Res*. **48** 207–213

Hahn G, Revello M G, Patrone M, Percivalle E, Campanini G, Sarasini A, Wagner M, Gallina A, Milanesi G, Koszinowski U, Baldanti F and Gerna G 2004 Human cytomegalovirus *UL131-128* genes are indispensable for virus growth in endothelial cells and virus transfer to leukocytes; *J. Virol.* **78** 10023–10033

Hassan-Walker A F, Okwuadi S, Lee L, Griffiths P D and Emery V C 2004 Sequence variability of the alpha-chemokine *UL146* from clinical strains of human cytomegalovirus; *J. Med. Virol.* **74** 573–579

He R, Liu LQ, Lu S M and Ruan Q 2001 Quantitative detection of HCMV-DNA from urine in infants by FQ-PCR; *Chin. J. Pediatr.* **12** 739–742

He R, Ruan Q, Xia C, Liu L Q and Lu S M 2004 Sequence variability of human cytomegalovirus *UL144* open reading frame in low passage clinical isolates; *Chin. Med. Sci. J.* **19** 293–297

He R, Ruan Q, Qi Y, Ma Y M, Huang Y J, Sun Z R and Ji Y H 2006 Sequence variability of human cytomegalovirus *UL146* and UL147 open reading frame in low passage clinical isolates; *Intervirology* **49** 215–223

Ho M 1990 Epidemiology of cytomegalovirus infections; *Rev. Infect. Dis.* **12** S701

Ji Y H, Sun Z R, Ruan Q, Ma Y P, He R, Qi Y and Huang Y J 2006 Polymorphisms of human cytomegalovirus *UL148A*, *UL148B*, *UL148C*, *UL148D* genes in clinical strains; *J. Clin.Virol.* **37** 252–257

Lurain N S, Kapell K S, Huang D D, Short J A, Paintsil J, Winkfield E, Benedict C A, Ware C F and Bremer J W 1999 Human cytomegalovirus *UL144* open reading frame: sequence hypervariability in low passage clinical isolates; *J. Virol.* **73** 10040–10050

Ma Y P, Ruan Q, He R, Qi Y, Sun Z R and Ji Y H 2006 Sequence variability of the human cytomegalovirus UL141 open reading frame in clinical strains; *Arch. Virol.* **151** 827–835

Nogalski M, Podduturi J, DeMeritt I , Milford L and Yurochko A 2007 The human cytomegalovirus virion possesses an activated casein kinase II that allows for the rapid phosphorylation of the inhibitor of NF-κB, Iκβα; *J.Virol.***81** 5305–5314

Penfold M E, Dairaghi D J, Duke G M, Saederup N, Mocarski E S, Kemble G M and Schall T J 1999 Cytomagalovirus encoded a potent alpha chemokine; *Proc. Natl. Acad. Sci. USA* **96** 9839–9844

Plotkin S A, Furukawa T, Zygraich N and Huygelen C 1975 Candidate cytomegalovirus strain for human vaccination; *Infect. Immun.* **12** 521–527

Qi Y, Mao Z Q, Ruan Q, He R, Ma Y P, Sun Z R and Ji Y H 2006 Human cytomegalovirus (HCMV) UL139 open reading frame: sequence variants are clustered into three major genotypes; *J. Med. Virol.* **78** 517–522

Rasmussen L, Geissler A and Winter M 2003 Futer- and intragenic variations complicate the molecular epidermiology of human cytomegalovirus; *J. Infect. Dis.* **187** 809–819

Saedrup N and Mocarski E S 2002 Fatal attraction: cytomegalovirus-encoded chemokine homologs; *Curr. Topics Microbiol. Immunol.* **269** 235–256

Sinzger C, Kahl M, Laib K, Klingel K, Rieger P, Plachter B and Jahn G 2000 Tropism of human cytomegalovirus for endothelial cells is determined by a post-entry step dependent on efficient translocation to the nucleus; *J. Gen. Virol*. **81** 3021–3035

Wirgart B Z, Bryting M, Linde A, Wahren B and Grillner L 1998 Sequence variation within three important cytomegalovirus gene regions in isolates from four different patient populations; *J. Clin. Microbiol.* **36** 3662–3668

Woodroffe S B, Hamilton J and Garnett H M 1997 Comparison of the infectivity of the laboratory strain AD169 and a clinical isolate of human cytomegalovirus to human smooth muscle cells; *J. Virol. Meth.* **63** 181–191

MS received 17 March 2007; accepted 25 June 2007

*e*Publication: 27 July 2007

Corresponding editor: Shahid Jameel