Human Cytomegalovirus UL138 Open Reading Frame Is Highly Conserved in Clinical Strains△
Ying Qi, Rong He, Yan-ping Ma, Zheng-rong Sun, Yao-hua Ji, and Qiang Ruan*
Virus Laboratory, Shengjing Hospital, China Medical University, Shenyang 110004, China

Key words: human cytomegalovirus; UL138; conservation

Objective To investigate the variability of human cytomegalovirus (HCMV) UL138 open reading frame (ORF) in clinical strains.

Methods HCMV UL138 ORF was amplified by polymerase chain reaction (PCR) and PCR amplification products were sequenced directly, and the data were analyzed in 19 clinical strains.

Results UL138 ORF in all 30 clinical strains was amplified successfully. Compared with that of Toledo strain, the nucleotide and amino acid sequence identities of UL138 ORF in all strains were 97.41% to 99.41% and 98.24% to 99.42%, respectively. All of the nucleotide mutations were substitutions. The spatial structure and post-translational modification sites of UL138 encoded proteins were conserved. The result of phylogenetic tree showed that HCMV UL138 sequence variations were not definitely related with different clinical symptoms.

Conclusion HCMV UL138 ORF in clinical strains is high conservation, which might be helpful for UL138 encoded protein to play a role in latent infection of HCMV.

Human cytomegalovirus (HCMV) is a β-herpes virus that causes widespread, lifelong human infection. HCMV infection in healthy individuals is asymptomatic, but HCMV primary infection or reactivation usually produces serious disease in immunocompromised individuals and congenitally infected newborns.1 The viral pathogenesis is poorly understood, but it is likely that both cellular and viral factors contribute to it.2

HCMV is a genetically complicated virus, and its genome consists of 230-235 kb double-stranded DNA and more than 200 predicted open reading frames (ORFs). The UL/b′ region of HCMV DNA containing at least 19 ORFs (UL133-UL151) was found in the Toledo strain (EMBL, AY446871) and in several other low-passaged clinical strains, but it was deleted in laboratory AD169 strain (EMBL, X17403).3 The further study indicated the UL138 ORF, located in the region, might be required for HCMV to establish and/or maintain a latent infection in hematopoietic progenitor cells infected in vitro.4 To determine the variability of the UL138 ORF in clinical HCMV strains, the UL138 ORF was sequenced and analyzed in the present study.
MATERIALS AND METHODS

Samples
The study population consisted of 30 HCMV congenitally infected infants who were identified in Shengjing Hospital of China Medical University during a period between 1988 and 1993 through a screening program. The infants aged less than 14 months with suspected congenital infection presenting different clinical symptoms: jaundice (J, n=18), congenital megacolon (C, n=7), and microcephaly (M, n=5). The strains were recovered from abnormal colon tissue or urine and passaged less than 10 times in human embryonic lung fibroblasts.

PCR amplification
Nested PCR was carried out for the amplification of all samples. The UL138 outer primers (forward: 5′-CGCACTGCTGCCAGAATG-3′, reverse: 5′-GTCCATCGTCCGAACAGC-GC-3′) yielded an 892 bp product, and inner primers (forward: 5′-CTGCCAGAATGGATGGATGCG-3′, reverse: 5′-CATCACCGAGGGAAAGACCG-3′) yielded an 845 bp product.

The reaction mixture of the first round PCR contained 1×buffer, 1.5 mmol/L MgCl₂, 0.2 mmol/L dNTPs, forward and reverse primers 150 ng respectively (outer primers), 0.5 U of Taq polymerase (Promega, USA), template 3.5 µl in a final volume of 50 µl. The conditions for amplification were 95℃ for 4 minutes, followed by 30 cycles at 95℃ for 45 seconds, 54℃ for 1 minute and 72℃ for 1 minute, then followed by a single extension cycle at 72℃ for 10 minutes.

The second round PCR was done with 2 µl of the first round PCR products as templates using the same reaction system and conditions, except for inner primers. In addition, negative control containing no template DNA was included in each batch of PCR test.

Purification and sequencing
PCR products were purified using the PCR Fragment Recovery Kit made by TaKaRa Company (Shiga, Japan). The concentrated PCR products were sequenced directly with BigDye Terminators Cycle Sequencing Kit (ABI Company, USA). Sequencing was carried out on both DNA strands, and a consensus sequence was established in order to eliminate PCR-induced errors.

Sequence analysis
All nucleotide and amino acid sequence analyses were undertaken using the BioEdit 5.0.0, DNASTar 5.1, and Genedoc software (DNASTAR. Inc., WI, USA). Sequences of the UL138 ORF were compared with the previously published data of 9 strains (Table 1). A phylogenetic tree was generated using the neighbour-joining method on the sequence alignment with the original public versions of ClustalW in the DNASTar 5.1 software.

RESULTS

Presence of the UL138 ORF in clinical strains
All clinical strains gave positive amplification. No DNA band was observed in negative control samples. Nineteen selected PCR products that were acquired from infants with three different types of clinical symptoms were sequenced successfully. The GenBank accession numbers of the sequences were listed in Table 1.

Nucleotide and amino acid sequences of UL138 ORF
Sequence analysis revealed that the UL138 ORF of all 19 clinical strains was composed of 510 nucleotides, and identical in size. Compared with that of Toledo strain, the identity in nucleotide ranged from 97.41% to 99.41%. All of the nucleotide mutations were substitutions, most of which were single nucleotide polymorphisms (SNPs).

Table 1. HCMV strains used and their GenBank accession numbers

<table>
<thead>
<tr>
<th>Strain</th>
<th>GenBank accession no.</th>
<th>Location</th>
<th>Passage</th>
<th>Clinical history</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toledo</td>
<td>AY446871</td>
<td>Unknown</td>
<td>Low passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>Merlin</td>
<td>AY446894</td>
<td>Cardiff</td>
<td>Low passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>6397</td>
<td>AY446870</td>
<td>Cardiff</td>
<td>Low passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>Towne</td>
<td>AY446869</td>
<td>Unknown</td>
<td>High passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>Davis</td>
<td>AY446868</td>
<td>Unknown</td>
<td>High passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>3157</td>
<td>AY446867</td>
<td>Cardiff</td>
<td>Low passage</td>
<td>Congenitally infected infant</td>
</tr>
<tr>
<td>TB40/E</td>
<td>AY446866</td>
<td>Unknown</td>
<td>High passage</td>
<td>Bone marrow recipient</td>
</tr>
<tr>
<td>W</td>
<td>AY446865</td>
<td>London</td>
<td>Unpassage</td>
<td>Infected AIDS patient</td>
</tr>
<tr>
<td>3301</td>
<td>AY446864</td>
<td>Cardiff</td>
<td>Unpassage</td>
<td>Congenitally infected infant</td>
</tr>
</tbody>
</table>

HCMV: human cytomegalovirus; J: jaundice; M: microcephaly; C: congenital megacolon; AIDS: Acquired Immunodeficiency Syndrome
which were synonymous ones. Compared to the Toledo sequence, the number of nucleotide mutations within the whole coding sequence of each strain ranged from 3 to 13. The predicted protein of UL138 ORF consisted of 169 amino acids with a calculated molecular mass of 19.3 kD. Compared with that of Toledo strain, the amino acid identity of UL138 ORF of all the clinical strains was 98.24% to 99.42%. The predicted polypeptide of each strain had 3 or fewer amino acid changes compared to the Toledo sequence.

Those foreign strains possessing nucleotide and amino acid variability in the UL138 coding region were concordant with Chinese strains studied here.

Structure of the predicted protein of UL138 ORF
The putative transmembrane domain of UL138 ORF was located within a region between amino acid 5 and 30. The predicted secondary structures of UL138 ORF were similar in all clinical strains (Fig. 1). The α-helices, β-sheets, turns, and coils were conserved in all clinical strains. The post-translational modification motifs of UL138 ORF in all clinical strains were conserved (Fig. 2), which contained protein kinase C phosphorylation site (PKC), tyrosine kinase phosphorylation motif (TKP), sulfation-cAMP-phosphorylation site (SPS), N-myristoylation site (MYR) and casein kinase II phosphorylation site (CKP). The post-translational modification motifs in clinical isolates mutated slightly. For example, Toledo, 20M, Towne, 3157, W and 3301 strains deleted a PKC site due to amino acid substitution from serine to aspartic acid at position 124.
Relationship between UL138 ORF variability and outcomes of HCMV infection

A neighbor-joining phylogenetic tree of UL138 ORF nucleotide consisted of nineteen sequences obtained in this study and nine ones published in GenBank (Fig. 3). The phylogenetic tree showed the strains with the same clinical symptom could not be clustered into one group, but dispersed in different branches.

**DISCUSSION**

In general, low-passage Toledo and clinical strains readily infect human endothelial, epithelial cell, and polymorphonuclear leukocytes.\(^6\) Due to the deletion of UL/b′ region, the high-passage AD169 strain infects and/or replicates less efficiently in these cell types. Moreover, AD169 strain shows attenuated virulence.\(^7\) These facts suggest that the predicted protein products of the 19 ORFs may determine the tropisms and virulence of HCMV strains in vivo. The previous studies showed UL/b′ region of HCMV was a variable gene region, and the variability of UL139,\(^8\) UL144,\(^9\) UL146 and UL147\(^10\) sequences might be related with the characters of HCMV strains.

However, little was known about the variability of UL138 ORF in the same genomic region. This study showed the UL138 ORF was present in all clinical strains. HCMV strains from different geographical regions had similar or identical variants of UL138 ORF. The spatial structure and post-translational modification sites of UL138 encoded proteins were highly conserved compared with those of Toledo. In addition, HCMV UL138 sequence variations were not definitely related with different clinical symptoms. In a word, UL138 ORF was conserved in all types of clinical strains, regardless of their polymorphism.

Although the different tropisms and virulence of HCMV strains could be determined by variable genes,\(^11\) conserved genes would be important in HCMV biology and genetic stability.\(^12\)

It is known to all that latency, which enables HCMV to persist in the cells of infected individuals indefinitely and prevent clearance of the pathogen,\(^13\) is critical importance to HCMV infection. The precise mechanisms governing the establishment of a latent infection are still unknown. For high conservation of UL138 sequences, it might be a target for antiviral agents against latent virus.

In conclusion, identification and characterization of UL138 ORF in clinical strains will allow a further understanding of HCMV infectious cycles.

**REFERENCES**

