Regular paper

Sequence analysis of human cytomegalovirus US28 gene in lowpassage clinical isolates from children and AIDS patients

Rong He¹, Chang Xia², Qiang Ruan^{2^{III}}, Ying Qi², Yan-Ping Ma², Yao-Hua Ji² and Jin-Jin Guo²

¹Clinical Genetics Department, The Affiliated Shengjing Hospital, China Medical University, Shenyang, Liaoning of PR China; ²Virus Laboratory, The Affiliated Shengjing Hospital, China Medical University, Shenyang, Liaoning of PR China.

Human cytomegalovirus (HCMV) is often a dangerous opportunistic pathogen that causes significant morbidity and mortality in newborn children and immunocompromised patients. The different symptoms and tissue tropisms of HCMV infection may result from genetic polymorphism. This study investigated the sequence variability of the HCMV US28 ORF, which shows sequence homology to the G protein-coupled receptor. HCMV isolated from suspected pediatric cases and isolates from AIDS patients were compared in order to examine the possible associations between polymorphisms and pathogenesis. Seventy children with suspected congenital HCMV infection, who suffered from jaundice (47), megacolon (10), and microcephaly (13), and 17 AIDS patients, were studied. Mutation was prevalent among the sequences of US28, with a focus on the two ends of US28. The important functional groups of US28 are highly conserved. An unrooted tree showed that all sequences from suspected congenitally infected infants and AIDS patients were divided into three groups. Comparison showed that most of the sequences (12/17) from pediatric patients were included in the first group (G1), whereas most of the sequences (11/17) from AIDS patients were included in the third group (G3). The specific high mutation sites in US28 from children were located at the C terminus of the protein, whereas those from AIDS patients were located at the N terminus. We demonstrated the existence of polymorphisms among the US28 genes of clinical isolates of HCMV from infants with suspected congenital infection. Comparison of US28 sequences from AIDS patients with those from children showed that both sequences have their own specific high mutation points.

Keywords: chemokine receptors, cytomegalovirus, genetic polymorphism, US28

Received: 28 December, 2010; revised: 23 May, 2011; accepted: 08 June, 2011; available on-line: 16 June, 2011

INTRODUCTION

Human cytomegalovirus (HCMV) causes a wide range of diseases in the general population. However, such infection is not usually associated with overt symptoms in immunocompetent individuals. For newborn children and immunocompromised patients, HCMV is often a dangerous opportunistic pathogen that causes significant morbidity and mortality (Ho, 1990). In fact, HCMV is now the most common viral cause of congenital malformation in developed countries and some under-developed regions. Congenitally acquired HCMV infection can lead to jaundice, hepatitis syndrome, megacolon, and microcephaly (Macris et al., 1981; Demmler, 1991; Tam et al., 1992). HCMV is also associated with exacerbation of vascular disease (Hendrix et al., 1991; Wu et al., 1992; Chen et al., 2003; Miller et al., 2004; Nerheim et al., 2004; Valantine, 2004), particularly after solid organ transplantation.

ORFs with homology to seven-transmembrane spanning receptors have been identified in the genomes of beta and gamma herpes viruses. Many cellular transmembrane spanning receptors have been shown to be G protein-coupled receptors (GPCRs). In the host, GPCRs can signal to a wide variety of pathways to modulate host cell homeostasis, induce chemotaxis and cell proliferation, or alter the expression of downstream regulated genes.

The HCMV genome contains four genes that encode GPCRs, namely US27, US28, UL33 and UL78, among which US28 is the most well described (Beisser et al., 2002). The US28 ORF encodes a functional β -chemokine receptor (US28) that binds and sequesters extracellular chemokines and can modulate host immune responses (Pleskoff et al., 1997; Vieira et al., 1998; Streblow et al., 1999). Expression of the US28 receptor promotes migration of smooth muscle cells. The GPCR family is associated with a G-protein, which leads to an intracellular signal through the G-phospholipase C pathway in a highly constitutive manner (Casarosa et al., 2001). Like many human chemokine receptors, the US28 receptor can operate as a cofactor for HIV cell entry (Pleskoff et al., 1997). Moreover, the US28 receptor can mediate cell fusion, together with other envelope proteins from, e.g., vesicular stomatitis virus (Pleskoff et al., 1998).

In previous studies, variability of the US28 gene has been investigated (Arav-Boger et al., 2002; Rasmussen et al., 2003; Goffard et al., 2006; Wang et al., 2009). However, nearly all HCMV isolates used in those studies were derived from patients who were immunocompromised by AIDS or transplantation. It is unknown whether the genetic variability observed among such isolates

e-mail: ruanq@sj-hospital.org The US28 sequences of the clinical isolates have been assigned GenBank accession No. AF378909 to AF378924 and AJ313049.

Abbreviations: CKP, casein kinase II phosphorylation site; FQ-PCR, fluorescence-quantified PCR; GPCRs, G protein-coupled receptors; HCMV, human cytomegalovirus; HELFs, human embryonic lung fibroblasts; HMA-SSCP, heteroduplex mobility assay/single-stranded conformation polymorphism; LPAS, lipoprotein lipid attachment site; PKC, protein kinase C phosphorylation; RANTES, regulated upon activation normal T cell expressed and secreted

adequately reflects the variability of HCMV isolates that circulate among HCMV-infected children.

The aim of this study was to explore the level of variability in the US28 gene obtained from infants with suspected congenital HCMV infection. The sequence variability of the US28 gene from these infants was compared with that from AIDS patients to examine possible associations between gene subtype and disease.

MATERIALS AND METHODS

Samples and virus. Samples of urine and abnormal colon tissue were collected for HCMV isolation from the Pediatrics Departments, Affiliated Shengjing Hospital, China Medical University in the 1980s. HCMV clinical isolates were isolated at the same time. All virus isolates of low passage were prepared in human embryonic lung fibroblasts (HELFs) cultured in DMEM supplemented with 10% FBS, as described previously (He et al., 2006). The isolates were kept at -70°C until use. Children who were suspected of congenital HCMV infection included those who suffered from jaundice (47), megacolon (10) or microcephaly (13). HCMV infection was confirmed by positive results of virus isolation or fluorescencequantified PCR (FQ-PCR) for HCMV (He et al., 2006). All samples were collected with the permission of the infants' parents and the research had obtained the approval of the Hospital Ethical Committee.

Seventeen sequences from AIDS patients in France were obtained from GenBank (acc. Nos. AJ 406953–69) (Goffard *et al.*, 2006).

PCR amplification. DNA was extracted from cells infected with clinical isolates of HCMV by boiling the culture medium at 100 °C. The extracted DNA was used as templates for PCR amplification of specific regions of the HCMV genome. We designed four pairs of overlapping primers for the amplification of US28, and their sequences are shown in Table 1. They were designed based on the sequence of ADl69 isolate (GenBank accession No. AY174271). The amplification was carried out with a Perkin–Elmer Cetus DNA Thermal Cycler. The conditions for amplification with all primer sets were 94 °C for 5 min followed by 35 cycles of 94 °C for 45 s, 55 °C for 45 s, and 72 °C for 45 s, with a single final extension cycle at 72 °C for 7 min.

Heteroduplex mobility assay/single-stranded conformation polymorphism (HMA-SSCP) analysis. Gel electrophoresis was carried out in 8% polyacrylamide gel with the addition of 5% glycerol to increase the resolution. Equivalent amounts of target and driver DNA from

Table 1. Primers used for PCR to amplify US28 sequences

Primer	Sequence	Product length (bp)
P28 U:	5'TGGTGAACCGCTCATATAGA3';	374
E1 D:	5'GTATTGCATCCACAGAGGTA3';	
P28a U:	5'ACGCGACTCCTTGTGTTTTC3';	435
E2 D:	5'TAAAGTGTGGAATGGCGATG3';	
P28b U:	5'GGTACGGCAGCCAAAAGATG3';	420
E3 D:	5' GGTACGGCAGCCAAAAGATG3';	
P28c U:	5' TCATGCTTGGTGCTTTCGTG3';	413
E4 D:	5'TCCGACGCGAAAAGCTCATG3'	

Cloning and sequencing. PCR products amplified with the primers that included the whole length of the US28 ORF were recovered from a 1% agarose gel slice in 1× TAE buffer, using the Wizard SV Gel and PCR Clean-Up System, according to the manufacturer's instructions (Promega). Purified fragments were eluted in 50 μ l nuclease-free water and cloned into the T-end plasmid (Advantage PCR Cloning Kit, Promega). The recombinant plasmids were extracted with PlasmidPURE kit (Promega).

Cloned PCR products were sequenced using the same primers as for PCR with the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). Sequencing was usually carried out on both DNA strands. The sequencing reactions were performed with a PE Applied Biosystems Geneamp PCR System 2400 at 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 4 min for a total of 30 cycles. The sequencing products were analyzed on an ABI 3700 automated sequencer (Applied Biosystems).

Sequence analysis. Sequence chromatograms were analyzed with the Chromas program (Applied Biosystems). Phylogenetic and functional group analyses were performed with BioEdit (http://www.mbio.ncsu.edu/ BioEdit/bioedit.html) and DNA STAR (Madison, WI, USA), respectively. The ClustalW algorithm was used to align multiple sequences of nucleotides and predicted amino acids. The phylogenetic trees were constructed using neighbor-joining methods with a maximum-likelihood-based distance. For each gene, a single, unrooted tree with 100 bootstrap values was generated. Bootstrap values of more than 60% were used to define nodes of the phylogenetic trees. A phylogenetic tree was displayed using TreeView, and the functional motifs were identified from the PROSITE database.

Nucleotide sequence accession numbers. The US28 sequences of the clinical isolates have been assigned GenBank accession No. AF378909 to AF378924 and AJ313049.

RESULTS

PCR amplification and HMA-SSCP analysis

The results showed that mutation was prevalent among the US28 ORFs. The products of the first and fourth pairs of primers had the greatest number of abnormal mobility shifts.

DNA sequence and phylogenetic analysis of US28

To determine the extent of sequence variability for the entire US28 ORF, the PCR products of 17 selected clinical isolates were cloned and sequenced successfully, including those from patients with jaundice (6), megacolon (6) or microcephaly (5). A comparison of these sequences with those from standard laboratory isolates (AD169: AY174271; Toledo: AY174280; VHLE: AX443409; and Towne: AY174281) and 17 AIDS patients (GenBank accession Nos. AJ406953-69) showed that the US28 ORF was variable. Sequence alignment showed that the mutations were mainly concentrated at the two ends of the gene, although some changes were scattered throughout the remainder of the coding sequence. To analyze the differences between the sequences of clinical isolates from suspected HCMV-infected children and from AIDS patients, the sequences were subjected to phylogenetic analysis to generate an unrooted tree. Phylogenetic analysis showed that US28 sequences from the clinical isolates formed three major groups and four subgroups (Fig. 1), designated G1A (9/34), G1B (9/34), G2 (4/34) and G3 (12/34). Based on this analysis, 18 sequences, including 12 clinical isolates (12/17, 70.1%), from the children suspected of being HCMV-infected formed one of the major groups (G1); four sequences from these children formed the smallest group (G2); and 11 sequences from the AIDS patients (11/17, 64.7%) along with one from the children formed the third group (G3). There were no sequences from AIDS patients in group G2, and only one sequence from pediatric patients in G3.

Amino-acid sequence analysis of the predicted US28 encoded protein

Analysis of amino-acid sequence showed that there were numerous non-synonymous substitutions in the sequences. The majority of mis-sense mutations were located at the two ends of the sequence. To investigate the mutations in the functional groups of the US28 encoded protein, the functional motifs of US28 ORFs were identified from the PROSITE database (Fig. 2). This demonstrated that the important groups of the US28 encoded protein included potential sites of N-glycosylation, protein kinase C (PKC) phosphorylation, casein kinase II phosphorylation, N-myristoylation, and prokaryotic membrane lipoprotein lipid attachment, and

a GPCR signature. Most of these important functional groups were distributed in the first third of the N-end of US28 encoded protein. Almost all the functional groups were highly conserved, with the exception of the clinical isolates 26M and 55J. Isolate 26M missed a PKC and a cAMP/cGMP-dependent protein kinase phosphorylation site, and had two new functional groups of casein kinase II phosphorylation site (CKP) and prokaryotic membrane lipoprotein lipid attachment site (LPAS). Isolate 55J had a new functional group for the PKC site.

Comparison of nucleotide and amino-acid sequences from infants with suspected congenital HCMV infection and AIDS patients

Among the 17 isolates from suspected congenitally HCMV infected infants, 15 contain mutation compared with AD169 isolate, while 13 out of the 17 isolates from AIDS patients have mutation. Most of the sequences (12/17, 70.1%) from children were included in the G1 group, while most of those (11/17, 64.7%) from AIDS patients were included in the G3 group. There were no sequences from AIDS patients in group G2, and only one sequence from pediatric patients in G3.

The US28 gene encodes 354 amino acids. The highest mutation rate of isolates from suspected congenitally HCMV infected infants is 1.69% (6/354) while the highest mutation rate of isolates from AIDS patients is 1.98% (7/354). Comparing the amino-acid sequences of the two groups of patients, we found that the amino acid sequences of US28 in the studied isolates had sites of high mutation at amino acid positions 18, 19 and 25. The high mutation sites in US28 from children were mainly located at the C terminus (AA 310 and 287), whereas those from AIDS patients were mainly located at the N terminus (AA8, 15, 21 and 24) (Fig. 3). The distribution of the mutation sites in the sequences of



the isolates from children and from AIDS patients clearly shows this difference in tendency.

Possible relationship between US28 phylogenetic groups and outcome of congenital HCMV infection

compared We the nucleotide sequence data and group classification according to the clinical source of the HCMV Although isolates. all isolates from children with megacolon were in the G1B and G2 groups only, and most isolates from infants with jaundice were in the G1 group only, no convincing relationships were observed between the presence of HCMV disease and specific US28 group due to the relatively small size of the sample used.

Figure 1. Phylogenetic analysis of US28 nucleotide sequences from 17 infants with suspected congenital HCMV infection and 17 AIDS patients The phylogenetic trees were constructed using neighbor-joining methods with a maximum-likelihood-

The phylogenetic trees were constructed using neighbor-joining methods with a maximum-likelihoodbased distance. For each gene, a single, unrooted tree with 100 bootstrap values was generated. An unrooted tree was displayed by TreeView. J, jaundice; M, microcephaly; C, megacolon; AJ, AIDS patients.



Figure 2. Alignment and functional motifs of amino-acid sequences from infants with suspected congenital HCMV infection 7TM, seven-transmembrane domain; J, jaundice; M: microcephaly; C, megacolon; TOWNE, VHL/E, AD169, laboratory isolates; CKP, casein kinase II phosphorylation site; N-G, N-glycosylation site; cAMP, cGMP, cAMP- and cGMP-dependent protein kinase phosphorylation site; PKC, protein kinase C phosphorylation site; MYS, N-myristoylation site; LPAS, prokaryotic membrane lipoprotein lipid attachment site; GPCR, G-protein coupled receptor signature.

1AT	17.7	• •			•			•	• •	• •		٠	•	•	•		٠		• 1			•			• •			• •		* 1			• •	1	• •		٠	۰.		•			• •								- 1					-					
1911							1	0							- 2	20							3	0						- 4	0							50	3						60)					7	10						1	80	2	
CIGJ	MT	PI	PΠ	°TP	T,	LΕ.	Ľ	T.	L I	819	٦D	e se	DI	R.,	~	νa	12	C)	vı	PЧ	D.	23	LN	IQ.	9F	œ	2	ΓI	.P	ĿŊ	7G	v	υP	т.;	LG	12	T,	GN	1 Pr	ъ		B.	гл	- The	NF	LR.	P.I	0	29	GI	DV	72	P	EN	IL.	~	5	D	ъл	ьP	V¢
93	• • •												- 1	D)	в,		-		. 3	ь.														- 1	r.																										
2014	1 · · ·												- 1	DI	Ξ.		-		- 3	ь.														- 3	г.											-									-	-		-			
103	1 · · ·		-	-		-	-	-	• •		-	-	-	- 3	P -	-	ъ	-			-			-		-	-		-		-	-		- 3	<u>.</u>	-	-		-			-		-		-		-				-			-	-	-	-			
16M									• •				۰.	4.5	<u>.</u>		-	•	1.2				• •								-	•	• •	- 3	F		-				• •					-		•						• •	-	-	•	-		• •	
2214	1.1.1		-	-		-	-	-	• •		-	-	- 1	23	- -	-	-		÷."		-			-		-	-		-		-	-		- 1	P	-	-		-		• •	-		-		-		-				-			-	-		-		• •	
26M	P. 1								• •				•	1.5	g,		-		x,	• •			• •		• •			• •					• •		P			• •			•••		• •		• •	-	• •		• •		• •			• •					• •	• •	
200				•			•	•	• •		•		•	14	к.	• •	-	•		• •	•		• •	•	• •			• •	•	• •		•	•••		E .	•		• •	•	•	•••		• •	•	• •	-	• •	•	• •		• •	•		• •			• •		• •	• •	
300				-			-	-			-	-	-		6.3		-	-	14	F. 7	-			-		-	-		-		-	-				-	-		-			-		-		-		-				-			-	-		-			
320	11.1												-	- 1			-				•		• •								-		•••	11	ř.	•					• •					-						•			-	-		-		÷	- C
322	12.2				2.3			2					- 1	Ď)	È.				1.5	Ē					::							-		2.5	F.					2.3								-					2.3			-					2.1
516	12.2																																	1.1	P																										
53C	12.2																-																	. 1	F.		-									-												-			
55 <i>a</i>	12.2																																		г.							÷.																			
5714				-				-			-	-	-	. 1	D.		-	-			-			-		-	-		-		-	-		- 1	г.	-	-		-			-		-		-		-				-			-	-		-			
693	[1] 1													. 1	ь.		-																	- 1	P.											-									-	-		-			
AJ69	[1] 1																-																	- 3	г.											-															a
AJ53	· · ·													. 3	ь.		-																	- 1	F.											-									-	-		-			
AJ54	· · ·																																	- 3	г.																										
AJSS	· · ·		-	-		-	-	-	• •		-	-	-	- 1	₽.	^	L -	-	- 3	ь.	-			-		-	-		-		-	-		- 3	r.	-	-		-			-		-		-		-				-			-	-	-	-			
AJ56	1 · · ·						•		• •				-	- 3	Ρ.		-	•	1.2				• •								-		• •	- 1	P			• •	-		• •					-								• •	-	-		-			
AJ57	1 I		-	-		-	-	-	• •		-	-	-	- 3	P -		· -	-	- 1	ь.	-			-		-	-		-		-	-		- 3	<u>r</u> -	-	-		-			-		-		-		-				-			-	-		-			
AJSB					-		•							- 6	2.		-	•		• •	•		• •	•									• •	- 1	<u> </u>		-	• •			• •				• •	-	• •				• •			• •	-			-			
AJ39															8.				1.4				• •		• •								• •		E			• •			•••				• •				• •		• •			• •					• •	• •	
AJ 60				•			•	-	• •		-		-	11	61			•			•		• •	•		-	-		-		-	-	• •	11	E.	•	-		-		• •	-		-		-		-				-		• •	-	-	• •	-		• •	
AJ 60														11							•		•••									•	•••	1.5	E .	•					•••			•				•	•••		• •	•								•••	
AJ60	12.2							-			-		-				-				-			-					-			-		1	F.	-	-		-			-		-		-		-				-			-	-		-			
AJ62	12.2				2.3			2					- 1	Ď)	È.				16	ь.		2.2										-		11	È.					2.3													2.3			-					÷.,
A.T63					1.5	2.2		2			Ē	£.,		Ŀ	Ξ.				Ť.															1.1	F .																										
AJT64						Ρ.					E			L	э.				T.																Ρ.																										
AJ65						G											-																	. 1	г.											-									-			-			
		• •			z	40						1	ż	ò		-		-	-	2	9	ė		-		-	3	<u>.</u>	5					żi	Lò			'	• •		22	ò	• •				3:	20			• •	•	1	4	ò					3	50
C163	80	. , BP			Z D	20		ъ	. , 11	o ar	'E	2	2 e DA		ŀe	н			NI	2	9		 72°	v	y . ST	кл	3	-	SIL	M	- 1	L/	AE:		Lo 10		 	81	 R.D		32		н. на	M3	sr.	31	3:	20	 E 2		r.e	s r	282		o T	 5 8	D1	5.A		3	50
C16J	act	BP	B		E E	70 KI		Б.		ber.		20	že LA			H		- 10	NIK		9 9		· · ·	v	s.	ĸ	3		SI.	M	-	Ŀ	E			21.3		81			32	21	на	1413	sr.	31	3:	30	E R	-		sT.	22	-4 5D	T.	58	DI	5A			50
C16J 9J 20M 10J	8C)	BF		R	B	70 KI		н		our.			26			н×		-	NIKK	2	-9 -9				9 T	KI			SIL	H		L.	E			P. 3	G.P.	81			32		HR		8 P.	31	2:	33	EF 70		r.e	s T	22	14 D	÷.		DI E	5A			-5.8
C16J 9J 20M 10J 16M	ac:	SP	B	R		70		. H		o T			Z G			H	-	-	NRR	2 1 1	-9 E		r	v	PT.	RJ			51	H	-	1.4				P. 1	5.8	81	ND		32	2	HR		81	31	3	30	EF 20	-	r.e	sw	22	-4 D	÷.	58	DI	5A			-5.8
C16J 9J 20M 10J 16M 22M	ac:	BP	B	R	H N	70		. H		b T			26 DA			H	-	-	NRR R	2	-9 		/F		ST.	KJ			51.	H<	711	L./	A.E.			P. 1		81	N.D		32	21	HR	0M12	5 F.	31	2:	30	1	994	P. E	sw	22	-4 D			DI				-520
C165 95 20M 105 10M 226M 226M	8C)	BF		N	H S .	70		. H					26			н×		- H.	· NKK · . K ·	2 E ² I	-9 H · · · · · ·		/ P*		9 T	K)				H(1.4	E					81			32		H 28		5 P.	31	30	30				s T		40	T		DI				-5.8
0970111N220	30	BP	B	R R R R R R R R R R R R R R R R R R R		70							26			HX			NER	E 1	-9 H				9 T	K)				HK		L./	E						N.D		32		HR		5 F.	31		20				sw		-4 D			DI	CCCCCC + +			-5
C165 95 105 105 105 105 105 105 200 200 200 200 200 200 200 200 200 2	8C)	BP	. B	RRRRR		70							26			HI		-	NRR . R .R .		-9 H				3710	K)				140		L./	E					81			32		HIR			31	0	20				st		-4D			DI	CCCCCC 2			-5.00
C165 95 105 105 105 105 105 105 105 105 105 10	30	BP		K RRRR R		70		·					2 E A			HIX			· NKK · K · K · · ·	2 H	-9 H		F		37T					HK		L				P. 1					32		H129			31		30				ST.		-4D			DI	C C C C C C C C C C C C C C C C C C C			-5.8
05 05 05 00 05 00 05 00 00 00 00 00 00 0	30			N. RRRRR		70		·		o T				P		H×			· NKK · K · K · · · ·	2 H	-9 H				3 T	KJ				140		L.																30				ST		-4D			DI	C . CCCCC			-5.8
05 05 05 05 05 05 05 05 05 05 05 05 05 0	30	BP		K. RRRRR		70		·								H.			· NKK · K · K · · · · ·		-9 H	Y \				KI			51	HX	710	I.J.	E			P(1												20				ST		-4D			D1	C. C. CCCCC			-58
C165 55 205 105 105 105 105 105 105 105 105 105 1	80	SF.		K RRRR R R		70		·					2 G			H.			· MKK · K · K · · · · · ·	2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			P.			KI			5 IL			L/	E			P. J					32		HIZ		5 P.			223				ST		4		D R	D1	CC C CCCCC 2			-58
0981108MM8000005005	80	5 F	B	K. RRRRR . R. RRR	2	70		·								HN			· NKK · · K · · · · · · ·	2	-9 H	· · · · · · · · · · · · · · · · · · ·	/F		3 W	K)						L/				P. J		81	N.D.		32 81W		HIZ		5 P.		9					ST.		4	T		DI	10000 C C C C C C C C C C C C C C C C C			-58
098112890000520585555555555555555555555555555	30	5 F	B			70		. H								HI		10	- MKK	2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-9 E	· · · · · · · · · · · · · · · · · · ·	/F*	V(K1						1.4	E			24,1		31					HIZ		32		2: (10)	30		5141		str		-4 D	T		DI	0000 0 00000 P			-5.8
115 45 45 45 45 45 45 45 45 45 45 45 45 45	30		B	K		70		- H								HIS			· MKK · . K ·	2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-9 E	¥ \	/F*	V(K1						T	. 15			94.3 							HIZ		5 F.		2: (K)			519(1		ST		-4D			DI	0. 0000 0. 00000 P			-58
00000000000000000000000000000000000000	30		B	K		70		- A								H0			· NKK ·	2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-9 H	· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·	V.	3 W	KJ						T	. 15			91.1 					22 8 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				5 P.		2: (R)			519(1		ST		-4D			D1	20.0000 0.0000 P			-58
C16 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	80		E		2H			·				· · · · · · · · · · · · · · · · · · ·				HX			· NKK · K · K · · · · · · · · · · ·	2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-9 H		· · · · · · · · · · · · · · · · · · ·	V	3 W	KJ				HIC		L.				P. 1											2: Q		· · · · · · · · · · · · · · · · · · ·	5 (1)		51		-4D			D1	200-0000 0 - 00000 - P			-58
10000000000000000000000000000000000000	80		E		2H	-70 KI						· · · · · · · · · · · · · · · · · · ·				HIS			· MRR · . R . R		-9E			V.		KJ				HIC		L															9			5 (1), 1		51		-4			DI	2000 0000 0 0000 P			-58
C15 C25 C25 C25 C25 C25 C25 C25 C25 C25 C2	23-C)					-70 KI						· · · · · · · · · · · · · · · · · · ·				1H1×			· MRR · . R . R		-9E			V.		KI																					2: 00					510		-4	· · · · · · · · · · · · · · · · · · ·		DI				-58
00000000000000000000000000000000000000	80		8	K. RRRRR. N. RRRRR. RRRRR	2. H	-70 KI		- A			E	· · · · · · · · · · · · · · · · · · ·				HIX			. NKK		9E			V (K1																					9					51		-4	· · · · · · · · · · · · · · · · · · ·		D1				-58
005 005 005 005 005 005 005 005 005 005	20		B	K. RRRRR. N. RRRRR. RRRRRQR	21			- A				· · · · · · · · · · · · · · · · · · ·				HIS			· NKK · K · K · · · · · · · · · · · · ·		-9 H			V		K.I																					9					5 T.		-4D	· · · · · · · · · · · · · · · · · · ·		D1	C. CCCCC. CCCC. C. CCCCC. S			-58
000 000 000 000 000 000 000 000 000 00	201		B	K. RRRRR. N. RRRRR. RRRRR	21							· · · · · · · · · · · · · · · · · · ·				HIS			· NKR · . R . K		9H		· · · · · · · · · · · · · · · · · · ·			KJ							A ES														9			3 19.1		5'T		-4D	· · · · · · · · · · · · · · · · · · ·		D1	00.0000.0000.0.0			-58
10 0 0 0 0 0 0 0 0 0 0 0 0 0	30			K. RRERE. H. RERER. RERESORE	21							· · · · · · · · · · · · · · · · · · ·				1410			· NKR · R · K · · · · · · · · · · · · · · ·		9H		· · · · · · · · · · · · · · · · · · ·																								9			5 1941		ST.		-4D	0 T			200-0000-0000-0-0			-58
14 4 15 15 15 15 16 16 16 16 16 16 16 16 16 16 16 16 16	200			K. RERER . N. RERER . RERERERER	22H							· · · · · · · · · · · · · · · · · · ·				1410			· MRR · R · K · · · · · · · · · · · · · · ·		9 H			V																							9			5 PL 8		5 T.		-4D	0 T			2222. 2222. 2222. 22222. 2			-58
10 C0 C0 C0 C0 C0 C0 C0 C0 C0 C	80			K. RREER. R. REER. REERRERER								· · · · · · · · · · · · · · · · · · ·				1400 · · · · · · · · · · · · · · · · · ·			· MRR · R · K · · · · · · · · · · · · · · ·		9 H			V		KI																					9			5 PL 8		5 T		-4D				22226. 666660. 20666. 6 20206. P			-58
14 4 15 15 15 15 15 15 15 15 15 15 15 15 15	25-01			K. RREER. R. REER. REERRERER.								· · · · · · · · · · · · · · · · · · ·				1400 · · · · · · · · · · · · · · · · · ·		-	· MKK · . K · · · · · · · · · · · · · · · ·					V		KI																					9					ST.		-4D							-578
10 0 0 0 0 0 0 0 0 0 0 0 0 0	80			K. RERER . R. RERER . REFERENCE		-70						· · · · · · · · · · · · · · · · · · ·			D	1400			· MKK · . K · · · · · · · · · · · · · · · ·					V		K1																					9 9					ST.		-40			DI	10. 00000 60000 0000 00 00 00000 0			-578
A A A A A A A A A A A A A A A A A	20											· · · · · · · · · · · · · · · · · · ·				140×			· MKK · . K ·		9 E					K1																					9 9					ST.		-40			DI				-500
10 0 0 0 0 0 0 0 0 0 0 0 0 0	80			K. BRRER . K. RRER . REKKORRER . REK								· · · · · · · · · · · · · · · · · · ·			0	1410			· MKK · . K · · · · · · · · · · · · · · · ·		20 E																										Q.					ST.		-40			DI	666. 66666. 666660. 6666. 6 66666. 5			-500

Figure 3. Alignment of US28 amino-acid sequences from 17 infants with suspected congenital HCMV infection and 17 AIDS patients J, jaundice; M, microcephaly; C, megacolon; AJ, AIDS patients.

DISCUSSION

Chemokines are chemoattractants for neutrophils, monocytes, lymphocytes and bone marrow progenitors, as well as other cell types. A characteristic feature of all chemokine receptors is their serpentine transmembranespanning domain structure. Its extracellular portions are involved in chemokine binding, whereas the intracellular portions are involved in cell signaling. GCPs link the binding of an extracellular ligand to processes within the cell by their activation of associated G proteins, whereas the G proteins can activate serine/threonine kinases, phosphatidylinositol 3-kinase and phospholipases. These proteins can stimulate mitogen-activated protein kinase or generate second messenger molecules, which results in activation of PKC and increases intracellular Ca²⁺ levels. Finally, these processes result in amplification of the initial signal that is transduced by the ligand-GCPR interaction into complex cellular processes such as chemotaxis (Gao & Murphy, 1994; Kuhn et al., 1995; Bodaghi et al., 1998; Vischer et al., 2006).

Evidence that HCMV infection might involve various organs and cell types, and the possibility that the virus uses the viral receptor to mediate chemokine trafficking and inflammatory cell migration prompted us to investigate the US28 gene which shows about 33% homology with the cellular β -chemokine receptor CCR-1.

HCMV clinical isolates display genetic polymorphism in several genes (Pignatelli *et al.*, 2004). UL144 ORF, a tumor necrosis factor- α -like receptor gene that belongs to the genomic region that is missing in laboratory isolates, shows significant isolate-specific sequence variability (Benedict *et al.*, 1999; Lurain *et al.*, 1999; Bale *et al.*, 2001; Arav-Boger *et al.*, 2002; Picone *et al.*, 2005). Previous studies have reported sequence divergence also in the US28 gene (Staak *et al.*, 1997; Baldanti *et al.*, 1998; Arav-Boger *et al.*, 2002; Goffard *et al.*, 2006). Here, we report the sequence variability of US28 in clinical isolates from suspected congenitally HCMV-infected infants and AIDS patients.

HMA-SSCP and sequencing showed that there was a wide range of nucleotide variation throughout the US28 gene. The results showed that the mutations were mainly concentrated at the two ends of the gene, although some changes were scattered throughout the remainder of the coding sequence. An unrooted tree showed that all sequences, including isolates from infants with suspected congenital HCMV infection and from AIDS patients, were divided into three groups and four subgroups. Analysis of the US28 amino-acid sequence showed that the mutations were concentrated at the N and C termini and at the ligand binding site of this viral GPCR.

In a previous study, we found that HCMV isolates from patients with microcephaly propagated in HELFs more quickly than other isolates did. This suggests that the sequence variability of HCMV genes underlies the characteristic pathogenic signature ascribed to different isolates in HCMV disease. In the present study, the phylogenetic analysis divided the 34 sequences available into three groups. All isolates from children with megacolon fell into the G1B and G2 groups, and most of the isolates from infants with jaundice fell into the G1 group. This difference suggests that the occurrence of mutations could be associated with pathogenesis.To confirm this, well-controlled studies with a large number of subjects will be required to determine the complex relationships between HCMV isolate type and human disease.

US28 is an early gene and begins to function at the early stage of virus infection (Billstrom *et al.*, 1998; Zipeto *et al.*, 1999). The work of Vieira *et al.* (1998) has demonstrated that infected cells that contain the intact US28 gene can bind with RANTES (regulated upon activation, normal T cell expressed and secreted) and induce Ca^{2+} influx within 24 h after infection, whereas infected cells with a nonfunctional US28 gene lose this ability. This suggests that US28 plays a role as a viral GPCR and mediates infection with HCMV, and affects its tissue tropism.

In the present study, functional group analysis of the US28 protein showed that almost all of the important functional groups were highly conserved. Conservation of the GPCR motif in US28 putative protein implies that this gene region has biological importance for HCMV and indicates that selection pressures favor retention of a specific gene sequence for this motif.

The present study compared the US28 gene sequences of clinical isolates from infants with suspected congenital HCMV infection and those from AIDS patients in France. Phylogenetically, most of the sequences (12/17) from pediatric patients fell within the first group, whereas most of those (11/17) from AIDS patients fell into the third group. The highmutation sites of the predicted amino-acid sequences differed between the suspected HCMV cases and the AIDS cases. Specifically, the high-mutation sites from the pediatric isolates were located at the C terminus, whereas those from the AIDS patients were located at the N terminus. It has been reported that the tissue and organ targets of HCMV infection in AIDS patients differ from those in children. Infection in children occurs mainly in the digestive and nervous systems, whereas in AIDS patients, infection mainly occurs in the lungs and eyes. US28 can bind chemokines and play an important part in the attraction and activation of leukocytes. The migration of leukocytes can transport infected cells to different parts of the body. The intracellular C terminus of the US28 receptor is involved in endocytosis of the protein (Waldhoer et al., 2003). Differences in sequence mutations in clinical isolates from children and AIDS patients could affect the migration of leukocytes and lead to different tissue and organ targets.

Analysis of HCMV isolates shed by young children provides important data for attempting to link genetic polymorphisms and clinical patterns of HCMV disease. In the present study, the isolates from infants with clinical HCMV infection possessed different nucleotide polymorphisms in the US28 coding region compared with HCMV isolates from AIDS patients. The result provides additional material for the study of HCMV pathogenesis. However, we could not preclude the effect of natural selection due to geographic distance and populations.

Acknowledgements

This study was supported by the National Natural Science Fund of China (30672248, 30770109, 30700916, 30801254 and 30901625).

We are grateful to Lanqing Liu and Shengmin LV for providing virus isolates from infants with congenital HCMV infection. Zhunyi Gou provided assistance with the DNA sequence analysis and phylogenetic analysis, which we gratefully acknowledge. The authors declare that they have no competing interests.

REFERENCES

- Arav-Boger R, Willoughby R, Pass R, Zong J, Jang W, Alcendor D, Hayward G (2002) Polymorphisms of the cytomegalovirus (CMV)encoded tumor necrosis factor-alpha and beta-chemokine receptors in congenital CMV disease. J Infect Dis 186: 1057–1064.
- Baldanti F, Sarasini A, Furione M, Gatti M, Comolli G, Revello MG, Gerna G (1998) Coinfection of the immunocompromised but not the immunocompetent host by multiple human cytomegalovirus strains. Arch Virol 143: 1701–1709.
- Bale JF, Petheram SJ, Robertson M, Murph JR, Demmler G (2001)
 Human cytomegalovirus a sequence and UL144 variability in strains from infected children. J Med Virol 65: 90–96.
 Beisser PS, Goh CS, Cohen FE, Michelson S (2002) Viral chemokine
- Beisser PS, Goh CS, Cohen FE, Michelson S (2002) Viral chemokine receptors and chemokines in human cytomegalovirus trafficking and interaction with the immune system. CMV chemokine receptors. *Curr Top Microbiol Immunol* 269: 203–234.
- Benedict ĆA, Butrovich KD, Lurain NS, Corbeil J, Rooney I, Schneider P, Tschopp J, Ware CF (1999) Cutting edge: a novel viral TNF receptor superfamily member in virulent strains of human cytomegalovirus. J Immunol 162: 6967–6970.
- Billstrom MÅ, Johnson GL, Avdi NJ, Worthen GS (1998) Intracellular signaling by the chemokine receptor US28 during human cytomegalovirus infection. J Virol 72: 5535–5544.
- Bodaghi B, Jones TR, Zipeto D, Vita C, Sun L, Laurent L, Arenzana-Seisdedos F, Virelizier JL, Michelson S (1998) Chemokine sequestration by viral chemoreceptors as a novel viral escape strategy: withdrawal of chemokines from the environment of cytomegalovirus-infected cells. J Exp Med 188: 855–866.
 Casarosa P, Bakker RA, Verzijl D, Navis M, Timmerman H, Leurs R,
- Casarosa P, Bakker RA, Verzijl D, Navis M, Timmerman H, Leurs R, Smit MJ (2001) Constitutive signaling of the human cytomegalovirus-encoded chemokine receptor US28. J Biol Chem 276: 1133–1137.
- rus-encoded chemokine receptor US28. J Biol Chem 276: 1133–1137. Chen R, Xiong S, Yang Y, Fu W, Wang Y, Ge J (2003) The relationship between human cytomegalovirus infection and atherosclerosis development. Mol Cell Biochem 249: 91–96.
- Demmler GJ (1991) Infectious Diseases Society of America and Centers for Disease Control. Summary of a workshop on surveillance for congenital cytomegalovirus disease. *Rev Infect Dis* **13:** 315–329.
- Gao JL, Murphy PM (1994) Human cytomegalovirus open reading frame US28 encodes a functional beta chemokine receptor. J Biol Chem 269: 28539–28542.
- Goffard A, Gault E, Rozenberg F, Moret N, Hober D, Dény P (2006) Comparative sequence analysis of US28 gene of human cytomegalovirus strains isolated from HIV-positive patients. *Virus Genes* 33: 175–181.
- He R, Ruan Q, Qi Y, Ma YP, Huang YJ, Sun ZR, Ji YH (2006) Sequence variability of human cytomegalovirus UL146 and UL147 genes in low-passage clinical isolates. *Intervirology* 49: 215–223.
- Hendrix MG, Daemen M, Bruggeman CA (1991) Cytomegalovirus nucleic acid distribution within the human vascular tree. Am J Pathol 138: 563–567.
- Ho M (1990) Epidemiology of cytomegalovirus infections. Rev Infect Dis 12 (Suppl 7): S701–S710.
- Kuhn DE, Beall CJ, Kolattukudy PE (1995) The cytomegalovirus US28 protein binds multiple CC chemokines with high affinity. *Biochem Biophys Res Commun* 211: 325–330.
- Lurain NS, Kapell KS, Huang DD, Short JA, Paintsil J, Winkfield E, Benedict CA, Ware CF, Bremer JW (1999) Human cytomegalovirus

UL144 open reading frame: sequence hypervariability in low-passage clinical isolates. J Virol 73: 10040–10050.

- Macris MP, Nahmias AJ, Bailey PD, Lee FK, Visintine AM, Brann AW (1981) Electron microscopy in the routine screening of newborns with congenital cytomegalovirus infection. J Virol Methods 2: 315–320.
- Miller DM, Espinosa-Heidmann DG, Legra J, Dubovy SR, Sũner IJ, Sedmak DD, Dix RD, Cousins SW (2004) The association of prior cytomegalovirus infection with neovascular age-related macular degeneration. Am J Ophthalmol 138: 323–328.
- Pione O, Costa J, Chaix M, Ville Y, Rouzioux C, Leruez-Ville M (2005) Human cytomegalovirus UL144 gene polymorphisms in congenital infections. J Clin Microbiol 43: 25–29.
- Pignatelli S, Dal Monte P, Rossini G, Landini MP (2004) Genetic polymorphisms among human cytomegalovirus (HCMV) wild-type strains. *Rev Med Virol* 14: 383–410.
- Pleskoff O, Tréboute C, Brelot A, Heveker N, Seman M, Alizon M (1997) Identification of a chemokine receptor encoded by human cytomegalovirus as a cofactor for HIV-1 entry. *Science* 276: 1874– 1878.
- Pleskoff O, Tréboute C, Alizon M (1998) The cytomegalovirus-encoded chemokine receptor US28 can enhance cell-cell fusion mediated by different viral proteins. J Virol 72: 6389–6397.
- Rasmussen L, Geissler A, Winters M (2003) Inter- and intragenic variations complicate the molecular epidemiology of human cytomegalovirus. J Infect Dis 187: 809–819.
- Staak K, Prosch S, Stein J, Priemer C, Ewert R, Docke WD, Kruger DH, Volk HD, Reinke P (1997) Pentoxifylline promotes replication of human cytomegalovirus *in vivo* and *in vitro*. *Blood* 89: 3682–3690.
- Streblow DN, Soderberg-Naucler C, Vieira J, Smith P, Wakabayashi E, Ruchti F, Mattison K, Altschuler Y, Nelson JA (1999) The human cytomegalovirus chemokine receptor US28 mediates vascular smooth muscle cell migration. *Cell* **99:** 511–520.
- Tam PK, Quint WG, van Velzen D (1992) Hirschsprung's disease: a viral etiology? *Pediatr Pathol* 12: 807–810.
- Valantine HA (2004) The role of viruses in cardiac allograft vasculopathy. Am J Transplant 4: 169–177.
- Vieira J, Schall TJ, Corey L, Geballe AP (1998) Functional analysis of the human cytomegalovirus US28 gene by insertion mutagenesis with the green fluorescent protein gene. J Virol 72: 8158–8165.
- Vischer HF, Leurs R, Smit MJ (2006) HCMV-encoded G-protein-coupled receptors as constitutively active modulators of cellular signaling networks. *Trends Pharmacol Sci* 27: 56–63.
- Waldhoer M, Casarosa P, Rosenkilde MM, Smit MJ, Leurs R, Whistler JL, Schwartz TW (2003) The carboxyl terminus of human cytomegalovirus-encoded 7 transmembrane receptor US28 camouflages agonism by mediating constitutive endocytosis. J Biol Chem 278: 19473–19482.
- Wang A, Ren L, Abenes G, Hai R (2009) Genome sequence divergences and functional variations in human cytomegalovirus strains. *FEMS Immunol Med Microbiol* 55: 23–33.
- Wu TC, Hruban RH, Ambinder RF, Pizzorno M, Cameron DE, Baumgartner WA, Reitz BA, Hayward GS, Hutchins GM (1992) Demonstration of cytomegalovirus nucleic acids in the coronary arteries of transplanted hearts. *Am J Pathol* 140: 739–747.
- Zipeto D, Bodaghi B, Laurent L, Virelizier JL, Michelson S (1999) Kinetics of transcription of human cytomegalovirus chemokine receptor US28 in different cell types. J Gen Virol 80 (Pt 3): 543–547.