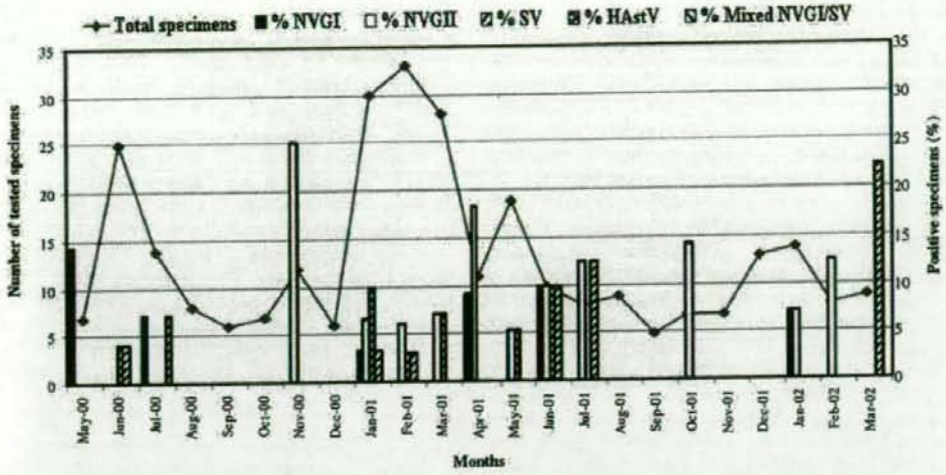



Fig. 1. Monthly distribution of NVGI, NVGII, SV, and HAstV infections among children hospitalized with diarrhea in Chiang Mai, Thailand, from May 2000 to March 2002 (number of tested specimens, -◆-; positive specimens (%), □).

according to the recent SV capsid region classification scheme described by Akihara et al. [2005]. It was observed that SVGI was a more common genogroup (80%) than the SVGII (20%) detected in this study. Eight strains of SVGI were classified further into three genotypes, that is, four strains were GI/1, three were GI/4, and one was GI/5. Moreover, two strains of SVGII were also classified into GII/1 and GII/2 genotypes (Fig. 3).

### Sequence Analysis of Human Astroviruses

Seven HAstV strains detected in the present study, two of each isolate were HAstV-1, HAstV-2, HAstV-5, and one isolate was HAstV-3. The detection rate of the HAstV in children with acute gastroenteritis was rather low.

## DISCUSSION

The present study describes the prevalence of NV and SV infection in children hospitalized with acute gastroenteritis in Chiang Mai city during May 2000–March 2002. The prevalent rates are in good agreement with those reported by Hansman et al. [2004] which also conducted the study in Chiang Mai during July 2000–July 2001. However, the prevalence of NV and SV are somewhat lower when compare to the follow-up study conducted recently in Chiang Mai from March 2002 to December 2004 [Khamrin et al., 2007] as well as to the study conducted in five other regions of Thailand by Guntapong et al. [2004]. The discrepancy of the prevalent rates between our study and others might be due to the difference in the duration and/or geographical area where those studies have been conducted.

Like other studies [Schnagl et al., 2000; Buesa et al., 2002; Oh et al., 2003; Boga et al., 2004; Hansman et al., 2004], the findings showed that NVGI strains are less common (29.2%) than NVGII (70.8%). It should be noted that in 2002 and 2003, NVGI disappeared completely from five other regions of Thailand [Guntapong et al., 2004]. In addition, NVGI was also undetectable in Chiang Mai area during 2002 and 2004 [Khamrin et al., 2007]. For NVGII, GII/4 has been reported as a major cause of global outbreaks and sporadic cases of gastroenteritis [Foley et al., 2001; White et al., 2002; Lau et al., 2004]. This study in Chiang Mai area found that GII/4 circulated as the most predominant genotype (37.5%), which is similar to those reported by Hansman et al. [2004]. However, a study conducted by Guntapong et al. [2004] in five other regions of Thailand during 2002 and 2003 reported a relatively high GII/4 at the incidence of 64.3%. Interestingly, in the following three consecutive years from 2002 to 2004 [Khamrin et al., 2007], GII/4 increased to 62.8% in Chiang Mai region, which similar to the finding of Guntapong et al. [2004]. For SV infection, SV genogroup I (GI) has been reported worldwide as the most predominant strain [Okada et al., 2002; Phan et al., 2004; Akihara et al., 2005; Phan et al., 2005, 2006]. SVGI/1 strains were previously reported as the most predominant genotype, followed by GII/1 strains and one isolate belonging to an intragenogroup recombinant strain in Chiang Mai during 2000 and 2001 [Hansman et al., 2004; Katayama et al., 2004]. In our study, SVGI was also detected at a very high incidence (80%), with GI/1 as the most predominant strain and followed by GI/4, GI/5, GII/1, and GII/2 strains. The study conducted in the other regions of Thailand by Guntapong et al. [2004] during 2002 and 2003 reported a higher detection rate of SV (15.0%). SVGI/1 was the most prevalent genotype, while the other two strains belonged to SVGV and a novel genotype in the SVGII

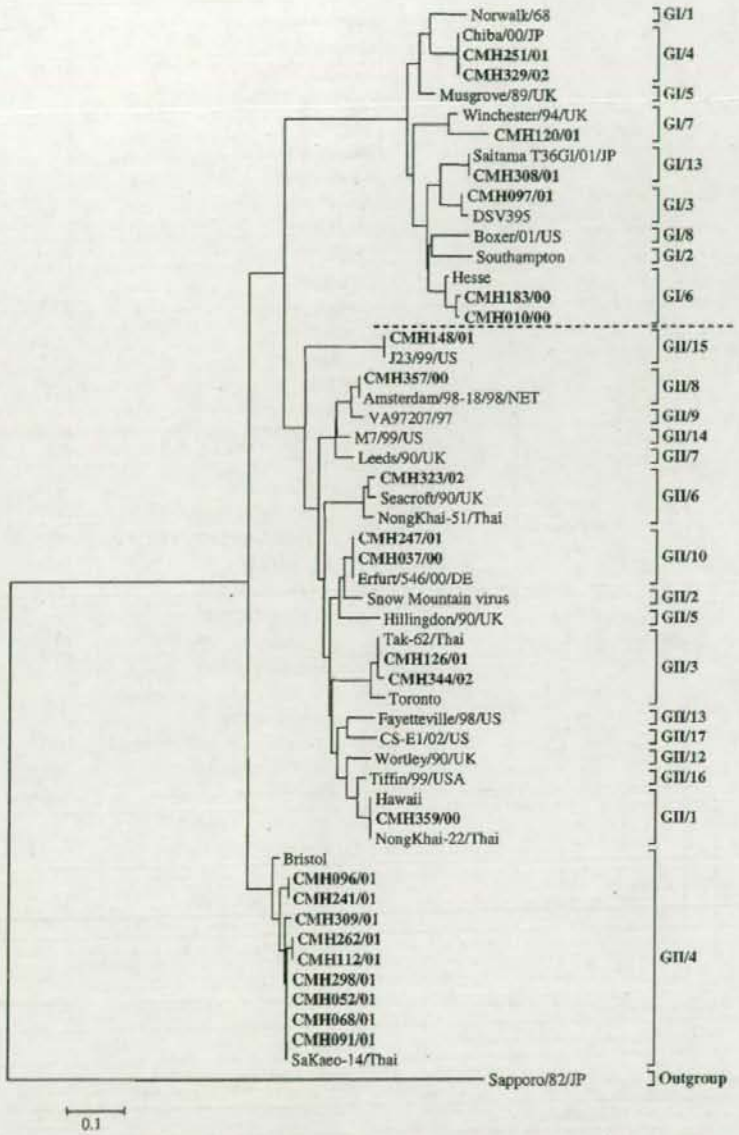



Fig. 2. Phylogenetic analysis of partial capsid deduced amino acid sequences of NVGI and GII strains detected between 2000 and 2002 in Chiang Mai, Thailand. The tree was constructed by multiple alignment of 7 NVGI and 17 NVGII positive sequences (indicated in boldface), 29 reference sequences, and 1 outgroup sequence. In the phylogenetic tree, NVGI strains were classified into eight distinct genotypes from 1 to 8 and NVGII strains were classified into sixteen distinct genotypes from 1 to 10 and 12 to 17 (excepting genotype 11 of porcine NVGII). Sapporo/82/JP was used as an outgroup strain for phylogenetic analysis. Bootstrap values are 1,000 replicates based on neighbor-joining and distance methods. Genotypes or genetic clusters are divided by brackets.

cluster, respectively. Another study in Chiang Mai region from 2002 to 2004 by Khamrin et al. [2007] reported a rather low detection rate of SV (1.2%) in which two strains belonged to SVGI genogroup (SVGI/1, SVGI/2) and other strain belonged to SVGIV genogroup. The data from our study and others [Guntapong et al., 2004; Khamrin et al., 2007] reveals that the detection rate, genogroup and genotype of NV and SV strains circulating in several regions of Thailand vary from time to time. However, the predominant NVGII/4 and SVGI persist in these regions over a number of years.

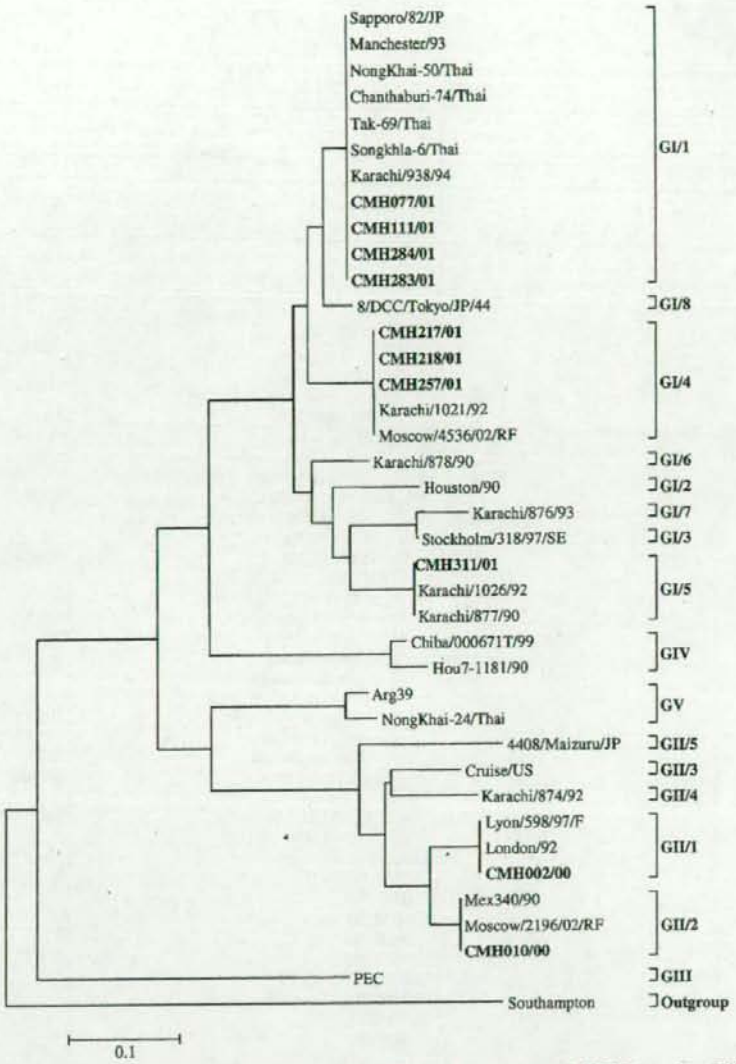



Fig. 3. Phylogenetic analysis of partial capsid deduced amino acid sequences of SV strains detected between 2000 and 2002 in Chiang Mai, Thailand. The tree was constructed by multiple alignment of 10 SV positive sequences (indicated in boldface), 27 reference sequences, and 1 outgroup sequence. In the phylogenetic tree, SV strains were classified into genogroup I, II, IV, and V. SVGI was classified further into eight genotypes and SVGII into five. SVGIII (Porcine enteric calicivirus, PEC) was also analyzed. Southampton was used as an outgroup strain for phylogenetic analysis. Bootstrap values are 1,000 replicates based on neighbor-joining and distance methods. Tree is unrooted. Genotypes or genetic clusters are divided by brackets.

Epidemiological data of HAstV as a causative agent of gastroenteritis in Thailand is rather limited. In 1991, Herrmann et al. first detected HAstV serotype 2 at 8.6% in children hospitalized with diarrhea in Bangkok. Later, Echeverria et al. [1994] reported the detection of HAstV at 14% in children hospitalized with diarrhea in Ratchburi province in the central part of Thailand. In 2004, Sirinavin et al. [2006] reported an outbreak of HAstV with a detection rate of 30.7% in neonates at a nursery in the maternity ward of Ramathibodi Hospital, Bangkok. However, the detection of HAstV infection in the previous studies based on serological assays [Herrmann et al., 1991; Echeverria et al., 1994; Sirinavin et al., 2006]. Our study is the first report that describes the distribution of HAstV genotypes circulating in Chiang Mai city. The data suggest that HAstV infection is less common in children with acute gastroenteritis compared to rotavirus and norovirus infection

in this area. These percentages are similar to those reported from other regions of the world such as Australia [Mustafa et al., 2000], Germany [Oh and Schreier, 2001], and Spain [Guix et al., 2002].

In conclusion, this study describes the genetic diversity of NV, SV, and HAstV genotypes cocirculating in children hospitalized with diarrhea in Chiang Mai, Thailand.

## REFERENCES


Akihara S, Phan TG, Nguyen TA, Yagyu F, Okitsu S, Müller WE, Ushijima H. 2005. Identification of sapovirus infection among Japanese infants in a day care center. J Med Virol 77:595–601.

Bertolotti-Ciarlet A, Crawford SE, Hutson AM, Estes MK. 2003. The 3' end of Norwalk virus mRNA contains determinants that regulate the expression and stability of the viral capsid protein VP1: A novel function for the VP2 protein. J Virol 77:11603–11615.

Boga JA, Melón S, Nicieza I, De Diego I, Villar M, Parra F, De Oña M. 2004. Etiology of sporadic cases of pediatric acute gastroenteritis in asturias, Spain, and genotyping and characterization of norovirus strains involved. J Clin Microbiol 42:2668–2674.

Buesa J, Collado B, López-Andújar P, Abu-Mallouh R, Rodríguez Díaz J, García Díaz A, Prat J, Guix S, Llovet T, Prats G, Bosch A. 2002. Molecular epidemiology of caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. J Clin Microbiol 40:2854–2859.

Echeverria P, Hoge CW, Bodhidatta L, Tungtaem C, Herrmann J, Imlarp S, Tamura K. 1994. Etiology of diarrhea in a rural community in western Thailand: Importance of enteric viruses and enterovirulent Escherichia coli. J Infect Dis 169:916–919.

Farkas T, Zhong WM, Jing Y, Huang PW, Espinosa SM, Martinez N, Morrow AL, Ruiz-Palacios GM, Pickering LK, Jiang X. 2004. Genetic diversity among sapoviruses. Arch Virol 149:1309–1323.

Foley B, O'Mahony J, Hill C, Morgan JG. 2001. Molecular detection and sequencing of "Norwalk-like viruses" in outbreaks and sporadic cases of gastroenteritis in Ireland. J Med Virol 65:388–394.

Guix S, Caballero S, Villena C, Bartolomé R, Latorre C, Rabella N, Simó M, Bosch A, Pintó RM. 2002. Molecular epidemiology of astrovirus infection in Barcelona, Spain. J Clin Microbiol 40:133–139.

Guntapong R, Hansman GS, Oka T, Ogawa S, Kageyama T, Pongsuwanna Y, Katayama K. 2004. Norovirus and sapovirus infection in Thailand. Jpn J Infect Dis 57:276–278.

Hansman GS, Katayama K, Maneekarn N, Peerakome S, Khamrin P, Tonusin S, Okitsu S, Nishio O, Takeda N, Ushijima H. 2004. Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of acute gastroenteritis in Chiang Mai, Thailand. J Clin Microbiol 42:1305–1307.

Herrmann JE, Taylor DN, Echeverria P, Blacklow NR. 1991. Astroviruses as a cause of gastroenteritis in children. N Engl J Med 324:1757–1760.

Jiraphongsa C, Laosiritaworn Y, Ngowabunpat A. 2005. Diarrhea mortality in children 0–5 years old in Thailand, 2001–2004. Bull Dep Med Services 30:43–51.

Katayama K, Miyoshi T, Uchino K, Oka T, Tanaka T, Takeda N, Hansman GS. 2004. Novel recombinant sapovirus. Emerg Infect Dis 10:1874–1876.

Khamrin P, Maneekarn N, Peerakome S, Tonusin S, Malasao R, Mizuguchi M, Okitsu S, Ushijima H. 2007. Genetic diversity of noroviruses and sapoviruses in children hospitalized with acute gastroenteritis in Chiang Mai, Thailand. J Med Virol 79:1921–1926.

Lau CS, Wong DA, Tong LK, Lo JY, Ma AM, Cheng PK, Lim WW. 2004. High rate and changing molecular epidemiology pattern of norovirus infections in sporadic cases and outbreaks of gastroenteritis in Hong Kong. J Med Virol 73:113–117.

Mustafa H, Palombo E, Bishop R. 2000. Epidemiology of astrovirus Infection in young children hospitalized with acute gastroenteritis in Melbourne, Australia, over a period of four consecutive years, 1995 to 1998. J Clin Microbiol 38:1058–1062.

Oh D, Schreier E. 2001. Molecular characterization of human astroviruses in Germany. Arch virol 146:443–455.

Oh DY, Gaedicke G, Schreier E. 2003. Viral agents of acute gastroenteritis in German children: Prevalence and molecular diversity. J Med Virol 71:82–93.

Okada M, Shinozaki K, Ogawa T, Kaiho I. 2002. Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. Arch Virol 147:1445–1451.

Okada M, Ogawa T, Kaiho I, Shinozaki K. 2005. Genetic analysis of noroviruses in Chiba prefecture, Japan, between 1999 and 2004. J Clin Microbiol 43:4391–4401.

Phan TG, Okame M, Nguyen TA, Maneekarn N, Nishio O, Okitsu S, Ushijima H. 2004. Human astrovirus, norovirus (GI, GII), and sapovirus infections in Pakistani children with diarrhea. J Med Virol 73:256–261.

Phan TG, Okame M, Nguyen TA, Nishio O, Okitsu S, Ushijima H. 2005. Genetic diversity of sapovirus in fecal specimens from infants and children with acute gastroenteritis in Pakistan. Arch Virol 150: 371–377.

Phan TG, Trinh QD, Yagyu F, Sugita K, Okitsu S, Müller WE, Ushijima H. 2006. Outbreak of sapovirus infection among infants and children with acute gastroenteritis in Osaka City, Japan during 2004–2005. J Med Virol 78:839–846.

Sakamoto T, Negishi H, Wang QH, Akihara S, Kim B, Nishimura S, Kaneshi K, Nakaya S, Ueda Y, Sugita K, Motohiro T, Nishimura T, Ushijima H. 2000. Molecular epidemiology of astroviruses in Japan from 1995 to 1998 by reverse transcription-polymerase chain reaction with serotype-specific primers (1 to 8). J Med Virol 61: 326–331.

Schnagl RD, Barton N, Patrikis M, Tizzard J, Erlich J, Morey F. 2000. Prevalence and genomic variation of Norwalk-like viruses in central Australia in 1995–1997. Acta Virol 44:265–271.

Schnagl RD, Belfrage K, Farrington R, Hutchinson K, Lewis V, Erlich J, Morey F. 2002. Incidence of human astrovirus in central Australia (1995 to 1998) and comparison of deduced serotypes detected from 1981 to1998. J Clin Microbiol 40:4114–4120.

Sirinavin S, Techasaensiri C, Okascharoen C, Nuntnarumit P, Tonsuttakul S, Pongsuwan Y. 2006. Neonatal astrovirus gastroenteritis during an inborn nursery outbreak. J Hosp Infect 64:196–197.

White PA, Hansman GS, Li A, Dable J, Isaacs M, Ferson M, McIver CJ, Rawlinson WD. 2002. Norwalk-like virus 95/96-US strain is a major cause of gastroenteritis outbreaks in Australia. J Med Virol 68:113–118.

Yan H, Yagyu F, Okitsu S, Nishio O, Ushijima H. 2003. Detection of norovirus (GI, GII), sapovirus and astrovirus in fecal samples using reverse transcription single-round multiplex PCR. J Virol Methods 114:37–44.

Zheng D, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. 2006. Norovirus classification and proposed strain nomenclature. Virology 346:312–323.

*Journal of General Virology* (2008), **89**, 2275–2279 DOI 10.1099/vir.0.83685-0

Short Communication

# Genetic linkage among human cytomegalovirus glycoprotein N (gN) and gO genes, with evidence for recombination from congenitally and post-natally infected Japanese infants

Hainian Yan,[1,2] Shin Koyano,[3] Yuhki Inami,[1] Yumiko Yamamoto,[1] Tatsuo Suzutani,[4] Masashi Mizuguchi,[2] Hiroshi Ushijima,[2] Ichiro Kurane[1] and Naoki Inoue[1]

Correspondence
Naoki Inoue
ninoue@nih.go.jp

[1]Department of Virology I, National Institute of Infectious Diseases, Tokyo, Japan

[2]Department of Developmental Medical Sciences, Institute of International Health, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

[3]Department of Pediatrics, Asahikawa Medical College, Hokkaido, Japan

[4]Department of Microbiology, Fukushima Medical University, Fukushima, Japan

Received 23 December 2007
Accepted 15 May 2008

Investigation of sequence polymorphisms in the glycoprotein N (gN; gp4273), gO (gp4274) and gH (gp4275) genes of human cytomegalovirus (HCMV) strains collected from 63 Japanese children revealed that their gO genotype distribution differed slightly from that of Caucasian populations and that there was a significant linkage between the gN and gO genotypes. Linkage of these genotypes in strains obtained from Caucasian populations has been reported, so our similar findings in Japanese infants are consistent with this, and suggest generality of this linkage. Sequence analysis suggests that recombination between two strains of different linkage groups occurred approximately 200 bp upstream of the 3′-end of the gO gene. Further studies are required to elucidate differences in biological characteristics among the linkage groups and the selective constraints that maintain the linkage.

Human cytomegalovirus (HCMV) infects most people during childhood without clinical symptoms; it is the major viral cause of birth defects and developmental abnormalities. It is also associated with significant morbidity and mortality in immunocompromised individuals. The complete genome of wild-type HCMV strains, such as Merlin, is 236 kb in length, and is predicted to encode 165 genes (Dolan *et al.*, 2004). Genetic characterization of clinical isolates has mainly depended on sequence polymorphisms in the genes that encode viral envelope glycoproteins and cellular homologues (Rasmussen, 1999; Pignatelli *et al.*, 2004). Glycoprotein B (gB), gM, gN, gH, gL and gO are involved in virus entry and egress and are the target molecules recognized by neutralizing antibodies. While extensive sequence variation is found in the gB and gH genes (5–10 %), a greater level is found in the gN and gO genes (40–50 %); the gL and gM genes are highly conserved among clinical strains. To date, the association of a particular genotype with a particular clinical outcome has been controversial (Bale *et al.*, 2000; Barbi *et al.*, 2001; Trincado *et al.*, 2000).

We have recently investigated gB, UL144 and UL149 gene polymorphisms and found that the gB3 genotype was more prevalent in congenitally infected individuals with neurological abnormalities (Yan *et al.*, 2008). Recently, a genetic linkage between the gN and gO genes was reported (Mattick *et al.*, 2004). To learn how common this linkage is and whether the linkage group has any correlation with the clinical outcome of congenital infection, we analysed gN, gO and gH gene polymorphisms.

HCMV strains were collected from 45 urine and 24 dried umbilical cord specimens obtained from 63 Japanese children, consisting of 32 congenitally and 31 post-natally infected children. Although samples of both materials were collected from six infants, specimens from each infant were handled as a single entity, as specimens from the same infant yielded the same sequence. Eleven of the congenital cases were identified previously by Ogawa *et al.* (2007). Six were identified by HCMV-specific IgM in maternal or cord blood specimens, a further six were identified by our HCMV screening programme (Inoue & Koyano, 2008) and

The sequence data determined in this study are available under GenBank accession numbers EU348337–EU348364.

Two supplementary tables are available with the online version of this paper.

the rest were identified by clinical manifestations. All congenital infections were confirmed by the detection of HCMV in urine within 2 weeks of birth or in dried umbilical cord specimens. Twenty-three healthy infants were chosen from >100 volunteers on the basis of the presence of HCMV in urine. HCMV was also collected from eight infants with hepatic damage, pneumonitis or bone marrow transplantation or infants that were born prematurely. Post-natal infection was implied by the absence of HCMV in their cord specimens. Viral DNA was extracted from these specimens as described previously (Ogawa *et al.*, 2007). DNA fragments encoding hypervariable regions of the gN, gO and gH genes were amplified by nested PCR using *Pfu* polymerase (Promega) in 50 μl reaction volumes. Ten to 100 HCMV DNA copies were used as templates for the first-round PCR and 2 μl of these products were used for the second-round PCR. PCR conditions and primers are shown in Supplementary Table S1, available in JGV Online. The PCR products were separated on agarose gels and purified using a DNA extraction kit (QiaEX II, Qiagen). The purified DNA fragments were sequenced with BigDye Terminator Cycle Sequencing kit (Applied Biosystems) using the primers from the second-round of PCR. Sequences were assembled with ATGC version 4.0 (Genetyx, Tokyo) and aligned with Genetyx 7.0. Phylogenic analysis was performed with MEGA version 3.1 (Kumar *et al.*, 2004). Designation of gN, gO and gH genotypes was based on previous publications (Mattick *et al.*, 2004; Stanton *et al.*, 2005; Chou, 1992; Pignatelli *et al.*, 2003).

The total number of available amplicons and the distribution of genotypes of congenitally and post-natally infected cases (including GenBank accesssion numbers) are given in Supplementary Table S2. There was no significant correlation between gH genotype and the incidence or clinical outcome of congenital infection. Since gO and gN have a large number of genotypes, the number of specimens in this study was insufficient to obtain statistically significant results in a clinical context. The presence of gO5 and the absence of gO1c in Japanese children made a slight difference in gO genotype distribution compared with the distribution in Caucasian populations (Mattick *et al.*, 2004). Since the gO5 genotype has only recently been recognized (Stanton *et al.*, 2005), the entire gO sequences of the five gO5 strains were determined. Their gO5 nucleotide sequences were identical to each other and were 99–100 % identical to those of Merlin and 3052. The gO genotypes exhibited a relatively low identity to each other; identity of the consensus gO5 sequence to other gO genotype sequences ranged from 76–81 % and 74–80 % at the nucleotide and the amino acid levels, respectively. Similar results were obtained from the phylogenic analyses for the sequences of the full-length and the middle segment of the gO gene (Fig. 1).

Of the 63 analysed strains, 57 yielded a complete dataset for the gN, gO and gH genotypes (Table 1). gO and/or gH genotypes of the remaining six strains, one from a cord

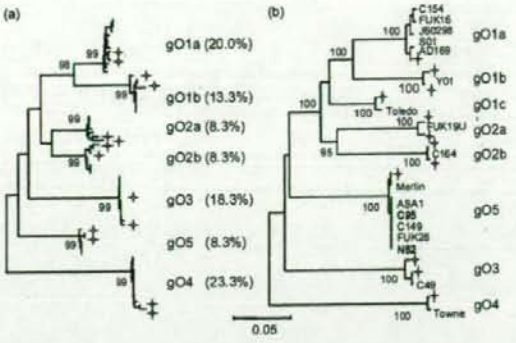



**Fig. 1.** (a) Phylogenic tree generated by the neighbour-joining method based on 60 nucleotide sequences covering the 440 bp middle part of the gO gene. Bootstrap values are indicated at the beginning of each major node. The frequencies of the gO genotypes among the 60 clinical strains are shown. Stars indicate the reference strains. (b) Phylogenic tree based on 13 full-length gO nucleotide sequences. Bar, 0.05 nucleotide substitutions per position.

specimen from a congenital case and five from urine specimens from healthy infants, were not available due to limiting amounts of HCMV DNA. Relationships were identified among the gN, gO and gH genotypes. For example, all gO1b strains link with gN3a and gH1, and all gO5 strains link with gN4c and gH2. Seven linkage groups cover 79 % (45/57) of the strains of these three genes for which sequences are available. If only the linkage between the gN and gO genotypes is considered, the seven groups cover 91 % (52/57) of the strains. Fisher's exact test of the distributions of the matched genotypes between the gN and gO genes yielded a significant association ($P<0.0001$). Thus, our results clearly support the findings from a previous report (Mattick *et al.*, 2004). As indicated in that study, breaking down the gO genotypes into seven or eight genotypes was critical to identifying the linkage that could not be previously identified in strains from US populations (Rasmussen *et al.*, 2002, 2003). Although the relationship between gO/gN and gH genotypes seems to be significant, the small number of gH variations/genotypes limits the value of these statistical analyses. Two gH genotypes were not able to be divided into subgenotypes. The genotypes of gB do not correlate with gN, gO and gH genotypes (data not shown).

In general, genetic polymorphisms and mosaic genotypes in the genome can be explained by the accumulation of spontaneous mutations under selective pressure and/or by homologous recombination, including both inter- and intra-strain recombination. Such recombination events have been reported for the variations within HCMV genes, such as gB (Haberland *et al.*, 1999) and a duplicated pair of virokines (UL146 and UL147) (Arav-Boger *et al.*, 2005). Recombination events have also been observed in other

**Table 1.** Linkage groups among the gN, gO and gH genotypes

Strains were isolated from congenital (C) or post-natal (P) infections.

| Strain | Infection type | Genotype | | | Linkage group |
|---|---|---|---|---|---|
| | | gN | gO | gH | |
| ASA12 | C | 1 | 1a | 1 | 1 |
| ASA59 | C | 1 | 1a | 1 | 1 |
| C106 | C | 1 | 1a | 1 | 1 |
| N42C | C | 1 | 1a | 1 | 1 |
| C102 | P | 1 | 1a | 1 | 1 |
| C140 | P | 1 | 1a | 1 | 1 |
| C141 | P | 1 | 1a | 1 | 1 |
| C177 | C | 3a | 1b | 1 | 2 |
| FUK32 | C | 3a | 1b | 1 | 2 |
| FUK72 | C | 3a | 1b | 1 | 2 |
| FUK82 | C | 3a | 1b | 1 | 2 |
| Y01 | C | 3a | 1b | 1 | 2 |
| C83 | P | 3a | 1b | 1 | 2 |
| J60250 | P | 3a | 1b | 1 | 2 |
| U02 | P | 3a | 1b | 1 | 2 |
| ASA16 | C | 3b | 2a | 1 | 3 |
| FUK19U | C | 3b | 2a | 1 | 3 |
| U01 | P | 3b | 2a | 2 | 3* |
| U06 | P | 3b | 2a | 2 | 3* |
| ASA68 | C | 2 | 2b | 1 | 4 |
| C164 | C | 2 | 2b | 1 | 4 |
| N66 | C | 2 | 2b | 1 | 4 |
| C134 | P | 2 | 2b | 1 | 4 |
| ASA15 | C | 4a | 3 | 2 | 5 |
| FUK03 | C | 4a | 3 | 2 | 5 |
| FUK20 | C | 4a | 3 | 2 | 5 |
| FUK31 | C | 4a | 3 | 2 | 5 |
| C145 | P | 4a | 3 | 2 | 5 |
| J60236 | P | 4a | 3 | 2 | 5 |
| U03 | P | 4a | 3 | 2 | 5 |
| C49 | C | 4a | 3 | 1 | 5* |
| J60249 | P | 4a | 3 | 1 | 5* |
| J60248 | P | 4a | 3 | 1 | 5* |
| ASA19 | C | 4b | 4 | 1 | 6 |
| FUK74 | C | 4b | 4 | 1 | 6 |
| X01 | C | 4b | 4 | 1 | 6 |
| C110 | P | 4b | 4 | 1 | 6 |
| C122 | P | 4b | 4 | 1 | 6 |
| C14 | P | 4b | 4 | 1 | 6 |
| C170 | P | 4b | 4 | 1 | 6 |
| C185 | P | 4b | 4 | 1 | 6 |
| C196 | P | 4b | 4 | 1 | 6 |
| J60223 | P | 4b | 4 | 1 | 6 |
| J60299 | P | 4b | 4 | 1 | 6 |
| N22 | P | 4b | 4 | 1 | 6 |
| C135 | P | 4b | 4 | 2 | 6* |
| U07 | P | 4b | 4 | 2 | 6* |
| ASA01 | C | 4c | 5 | 2 | 7 |
| ASA70 | C | 4c | 5 | 2 | 7 |
| FUK28 | C | 4c | 5 | 2 | 7 |
| N59 | C | 4c | 5 | 2 | 7 |
| C149 | P | 4c | 5 | 2 | 7 |
| C154 | C | 3a | 1a | 1 | † |
| FUK16 | C | 3a | 1a | 1 | † |
| S01 | C | 3a | 1a | 1 | † |
| J60284 | P | 3a | 2b | 1 | † |
| J60298 | P | 4c | 1a | 2 | † |

**Table 1.** cont.

*Strains with a gH genotype that is different to the others.

†Strains that were not classified into the seven groups and could be products of recombination.

herpesviruses, including herpes simplex type 1 (Dutch *et al.*, 1992), varicella-zoster virus (Norberg *et al.*, 2006; Peters *et al.*, 2006) and human herpesvirus 8 (Poole *et al.*, 1999). Recombination depends on various immunological and intracellular constraints because infection of the host, ultimately of a single cell, with two parental strains is required. Concurrent infections with multiple HCMV strains have been observed in immunosuppressed patients, such as transplant recipients and patients with human immunodeficiency virus (Stanton *et al.*, 2005; Puchhammer-Stöckl & Görzer, 2006; Coaquette *et al.*, 2004). It has been demonstrated that pre-existing immunity does not prevent infection with strains of different genotypes (Ishibashi *et al.*, 2007; Boppana *et al.*, 2001). To find evidence indicative of recombination events, the entire gO gene sequences were determined for strains representing each genotype and for those that were not classified into the seven linkage groups.

First, we analysed the gO5 strains. Whilst the gN4c genotype linked with the gO5 genotype in our population, the same gN4c genotype linked with the gO1c genotype in Caucasian populations. Mattick *et al.* (2004) discussed the possibility that gO1c was created by a recombination event. HCMV strains ASA01 and Toledo were chosen as representative strains of gN4c–gO5 and that of gN4c–gO1c, respectively. The similarity between the gN–gO sequences was analysed using the SimPlot program version 3.5 (http://asray.med.som.jhmi.edu/SCRofware/simplot) (Fig. 2a). The identity was >95 % from the gN gene to the 3′-part of the gO gene, but it dropped significantly from 200 bp upstream of the 3′-end of the gO gene. This suggests that the gO5 strains are also the products of recombination. The recombination site could be anywhere within the conserved areas of the gN gene or the 200 bp 3′-end region of the gO gene. If the unidentified counterpart for the recombination has a gN genotype other than gN4c, the transition site is expected to be around 200 bp upstream of the 3′-end of the gO gene. Since gO1c is one of the rarest genotypes and no gO1c strain was identified in this study, further study is required to understand the relationship between the gO1c and gO5 strains.

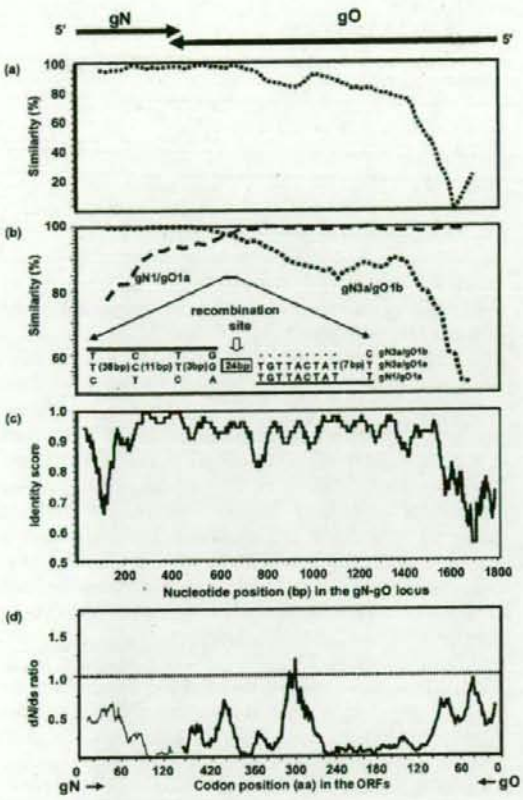



**Fig. 2.** Potential recombination events within the gO gene. The gN and gO genes are shown above the panels. The horizontal axis shows the nucleotide position starting from the 5′-end of the gN open reading frame (a–c) and amino acid position in the gN and gO ORFs (d). (a) SimPlot analysis (window size 200, step size 20) of the similarity (%) between ASA01 (gN4c-gO5) and Toledo (gN4c-gO1c). (b) SimPlot analysis (window size 250, step size 30) of C154 (gN3a-gO1a) with AD169 (gN1/gO1a; dashed line) and Y01 (gN3a-gO1b; dotted line). The divergent sequences around the potential recombination site are shown. (c) SimPlot analysis (window size 60, step size 1) of the mean similarity (identity score) among representative strains (AD169, Y01, FUK19, C164, C49, Towne, ASA1) of the eight linkage groups (seven groups described in this study and gN4c-gO1c). (d) VarPlot analysis (window size 20, step size 1) of the mean values of *dN*/*ds* ratio in the amino acid sequences of the gN and gO genes of the representative strains. *dN*/*ds*=1 is indicated with a dashed line.

Next, we analysed three strains (C154, FUK16 and S01) of the gN3a–gO1a genotype, as they were not classified into the seven linkage groups. Since the gO sequences of these strains were almost identical (99–100 % identity), the gO sequence of C154 was used for further analysis. The sequence of the gN–gO locus of C154 was compared with that of Y01, which represents linkage group 2 (gN3a–gO1b), and that of AD169, which represents linkage group 1 (gN1–gO1a). The sequence similarity of strain C154 with AD169 and Y01 declines around 200 bp upstream of the 3′-end of the gO gene (Fig. 2b). This pattern suggests a potential recombination event between linkage groups 1 and 2 within the gO gene. The potential recombination site is assumed to be in the conserved 24 bp sequence shown in Fig. 2(b). Compared with the surrounding regions, the sequence around the recombination site did not necessarily correspond to the well-conserved regions (Mattick *et al.*, 2004; Pignatelli *et al.*, 2003). To visualize this, sequence similarity was compared for all eight linkage groups across the gN–gO locus (Fig. 2c). It is unlikely that this potential recombination event was due to the presence of two strains in the specimen or to mispriming, since (i) infection with multiple strains was not identified at detectable levels (>25 % in the population) in the raw sequence data of the gN, gO, gH, gB, UL144 and UL149 genes; (ii) three very similar but distinct strains were obtained from individuals from different localities and at different collection times and their DNAs were extracted and analysed in separate tests; and (iii) different primer sets yielded the same genotyping results. It is possible that the recombinant HCMV was generated and circulated naturally.

Since only two cases of recombination were available, we could not tell whether the gO gene contains a hotspot sequence that triggers recombination, similar to those observed in other viruses (Magiorkinis *et al.*, 2003; Kajino *et al.*, 2001; Takeuchi *et al.*, 2008). It was, however, confirmed that no chi site- or V(D)J recombination site-like sequences were present in the gO and gN genes. To obtain insights into the mechanism of the recombination in the gO gene, the non-synonymous distance (*dN*) and synonymous distance (*ds*) of codon-based aligned gN and gO sequences were analysed using the VarPlot program, as a recent study provided evidence of positive selection in the hypervariable gN sequences (Pignatelli *et al.*, 2003). In addition to the gN sequence, the gO sequence showed generally low *dN* and *ds* values, and the *dN*/*ds* ratios were almost all less than 1 (Fig. 2d), indicating that negative pressure tends to maintain the original sequences. Although *dN*/*ds* ratios >1 were observed in limited domains from some genotypes, such as the gO1b, gO2a, gO2b and gO4 sequences, the potential recombination sites do not localize at those positively selected sites, suggesting that positive selective pressure, such as for immune escape, is not providing selection for recombination.

In conclusion, our study demonstrated a significant link between the gN and gO genotypes in Japanese infants, which supports a previous finding in Caucasian populations and suggests generality of the linkage. Whilst we describe a novel homologous recombination event in the gO gene, it will be important to identify additional recombination events in the gO gene in order to explain the mechanisms regulating recombination. Further studies are also required to elucidate differences in biological

characteristics among the linkage groups and to identify the selective constraints that maintain the linkage.

## Acknowledgements

We thank Yasushi Ohkusa, Phil Pellett, Naoki Nozawa and Kenji Fujieda for their intellectual input, and Risa Taniguchi and Hitomi Komura for their technical assistance. This study was supported by a grant (H20-Kodomo-007) from the Ministry of Health and Welfare (N.I.) and a fellowship from the Atsumi International Scholarship Foundation (H.Y.).

## References

**Arav-Boger, R., Zong, J. C. & Foster, C. B. (2006).** Loss of linkage disequilibrium and accelerated protein divergence in duplicated cytomegalovirus chemokine genes. *Virus Genes* **31**, 65–72.

**Bale, J. F., Jr, Murph, J. R., Demmler, G. J., Dawson, J., Miller, J. E. & Petheram, S. J. (2000).** Intrauterine cytomegalovirus infection and glycoprotein B genotypes. *J Infect Dis* **182**, 933–936.

**Barbi, M., Binda, S., Caroppo, S., Primache, V., Didò, P., Guidotti, P., Corbetta, C. & Melotti, D. (2001).** CMV gB genotypes and outcome of vertical transmission: study on dried blood spots of congenitally infected babies. *J Clin Virol* **21**, 75–79.

**Boppana, S. B., Rivera, L. B., Fowler, K. B., Mach, M. & Britt, W. J. (2001).** Intrauterine transmission of cytomegalovirus to infants of women with preconceptional immunity. *N Engl J Med* **344**, 1366–1371.

**Chou, S. (1992).** Molecular epidemiology of envelope glycoprotein H of human cytomegalovirus. *J Infect Dis* **166**, 604–607.

**Coaquette, A., Bourgeois, A., Dirand, C., Varin, A., Chen, W. & Herbein, G. (2004).** Mixed cytomegalovirus glycoprotein B genotypes in immunocompromised patients. *Clin Infect Dis* **39**, 155–161.

**Dolan, A., Cunningham, C., Hector, R. D., Hassan-Walker, A. F., Lee, L., Addison, C., Dargan, D. J., McGeoch, D. J., Gatherer, D. & other authors (2004).** Genetic content of wild-type human cytomegalovirus. *J Gen Virol* **85**, 1301–1312.

**Dutch, R. E., Bruckner, R. C., Mocarski, E. S. & Lehman, I. R. (1992).** Herpes simplex virus type 1 recombination: role of DNA replication and viral *a* sequences. *J Virol* **66**, 277–285.

**Haberland, M., Meyer-König, U. & Hufert, F. T. (1999).** Variation within the glycoprotein B gene of human cytomegalovirus is due to homologous recombination. *J Gen Virol* **80**, 1495–1500.

**Inoue, N. & Koyano, S. (2008).** Evaluation of screening tests for congenital cytomegalovirus infection. *Pediatr Infect Dis J* **27**, 182–184.

**Ishibashi, K., Tokumoto, T., Tanabe, K., Shirakawa, H., Hashimoto, K., Kushida, N., Yanagida, T., Inoue, N., Yamaguchi, O. & other authors (2007).** Association of the outcome of renal transplantation with antibody response to cytomegalovirus strain-specific glycoprotein H epitopes. *Clin Infect Dis* **45**, 60–67.

**Kajino, K., Yamamoto, T., Hayashi, J., Umeda, T., Takahara, T. & Hino, O. (2001).** Recombination hot spot of hepatitis B virus genome binds to members of the HMG domain protein family and the Y box binding protein family; implication of these proteins in genomic instability. *Intervirology* **44**, 311–316.

**Kumar, S., Tamura, K. & Nei, M. (2004).** MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

**Magiorkinis, G., Paraskevis, D., Vandamme, A. M., Magiorkinis, E., Sypsa, V. & Hatzakis, A. (2003).** *In vivo* characteristics of human immunodeficiency virus type 1 intersubtype recombination: determination of hot spots and correlation with sequence similarity. *J Gen Virol* **84**, 2715–2722.

**Mattick, C., Dewin, D., Polley, S., Sevilla-Reyes, E., Pignatelli, S., Rawlinson, W., Wilkinson, G., Dal Monte, P. & Gompels, U. A. (2004).** Linkage of human cytomegalovirus glycoprotein gO variant groups identified from worldwide clinical isolates with gN genotypes, implications for disease associations and evidence for N-terminal sites of positive selection. *Virology* **318**, 582–597.

**Norberg, P., Liljeqvist, J. A., Bergström, T., Sammons, S., Schmid, D. S. & Loparev, V. N. (2006).** Complete-genome phylogenetic approach to varicella-zoster virus evolution: genetic divergence and evidence for recombination. *J Virol* **80**, 9569–9576.

**Ogawa, H., Suzutani, T., Baba, Y., Koyano, S., Nozawa, N., Ishibashi, K., Fujieda, K., Inoue, N. & Omori, K. (2007).** Etiology of severe sensorineural hearing loss in children: independent impact of congenital cytomegalovirus infection and GJB2 mutations. *J Infect Dis* **195**, 782–788.

**Peters, G. A., Tyler, S. D., Grose, C., Severini, A., Gray, M. J., Upton, C. & Tipples, G. A. (2006).** A full-genome phylogenetic analysis of varicella-zoster virus reveals a novel origin of replication-based genotyping scheme and evidence of recombination between major circulating clades. *J Virol* **80**, 9850–9860.

**Pignatelli, S., Dal Monte, P., Rossini, G., Chou, S., Gojobori, T., Hanada, K., Guo, J. J., Rawlinson, W., Britt, W. & other authors (2003).** Human cytomegalovirus glycoprotein N (gpUL73-gN) genomic variants: identification of a novel subgroup, geographical distribution and evidence of positive selective pressure. *J Gen Virol* **84**, 647–655.

**Pignatelli, S., Dal Monte, P., Rossini, G. & Landini, M. P. (2004).** Genetic polymorphisms among human cytomegalovirus (HCMV) wild-type strains. *Rev Med Virol* **14**, 383–410.

**Poole, L. J., Zong, J. C., Ciufo, D. M., Alcendor, D. J., Cannon, J. S., Ambinder, R., Orenstein, J. M., Reitz, M. S. & Hayward, G. S. (1999).** Comparison of genetic variability at multiple loci across the genomes of the major subtypes of Kaposi's sarcoma-associated herpesvirus reveals evidence for recombination and for two distinct types of open reading frame K15 alleles at the right-hand end. *J Virol* **73**, 6646–6660.

**Puchhammer-Stöckl, E. & Görzer, I. (2006).** Cytomegalovirus and Epstein-Barr virus subtypes–the search for clinical significance. *J Clin Virol* **36**, 239–248.

**Rasmussen, L. (1999).** Molecular pathogenesis of human cytomegalovirus infection. *Transpl Infect Dis* **1**, 127–134.

**Rasmussen, L., Geissler, A., Cowan, C., Chase, A. & Winters, M. (2002).** The genes encoding the gCIII complex of human cytomegalovirus exist in highly diverse combinations in clinical isolates. *J Virol* **76**, 10841–10848.

**Rasmussen, L., Geissler, A. & Winters, M. (2003).** Inter- and intragenic variations complicate the molecular epidemiology of human cytomegalovirus. *J Infect Dis* **187**, 809–819.

**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**, 42–46.

**Takeuchi, Y., Myers, R. & Danos, O. (2008).** Recombination and population mosaic of a multifunctional viral gene, adeno-associated virus cap. *PLoS ONE* **3**, e1634.

**Trincado, D. E., Scott, G. M., White, P. A., Hunt, C., Rasmussen, L. & Rawlinson, W. D. (2000).** Human cytomegalovirus strains associated with congenital and perinatal infections. *J Med Virol* **61**, 481–487.

**Yan, H., Koyano, S., Inami, Y., Yamamoto, Y., Suzutani, T., Mizuguchi, M., Ushijima, H., Kurane, I. & Inoue, N. (2008).** Genetic variations in the gB, UL144 and UL149 genes of human cytomegalovirus strains collected from congenitally and post-natally infected Japanese children. *Arch Virol* **153**, 667–674.

Journal of Medical Virology 81:176–182 (2009)

# Molecular Characterization of VP4, VP6, VP7, NSP4, and NSP5/6 Genes Identifies an Unusual G3P[10] Human Rotavirus Strain

**Pattara Khamrin,[1] Niwat Maneekarn,[2] Supatra Peerakome,[2] Rungnapa Malasao,[3] Aksara Thongprachum,[4] Wisoot Chan-it,[4] Masashi Mizuguchi,[4] Shoko Okitsu,[1] and Hiroshi Ushijima[1,4]***

[1]*Aino Health Science Center, Aino University, Tokyo, Japan*
[2]*Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand*
[3]*School of Medical Sciences, Naresuan University Phayao, Phayao, Thailand*
[4]*Department of Developmental Medical Sciences, Institute of International Health, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan*

An unusual strain of human rotavirus G3P[10] (CMH079/05) was detected in a stool sample of a 2-year-old child admitted to the hospital with severe diarrhea in Chiang Mai, Thailand. Analysis of the VP7 gene sequence revealed highest identities with unusual human rotavirus G3 strain CMH222 at 98.7% on the nucleotide and 99.6% on the amino acid levels. Phylogenetic analysis of the VP7 sequence confirmed that the CMH079/05 strain formed a cluster with G3 rotavirus reference strains and showed the closest lineage with the CMH222 strain. Analysis of partial VP4 gene of CMH079/05 revealed highest degree of sequence identities with P[10] rotavirus prototype strain 69M at nucleotide and amino acid levels of 92.9% and 94.6%, respectively. Phylogenetic analysis of the VP4 sequence revealed that CMH079/05 and 69M clustered closely together in a monophyletic branch separated from other rotavirus genotypes. To our knowledge, this is a novel G–P combination of G3 and P[10] genotypes. In addition, analyses of VP6, NSP4, and NSP5/6 genes revealed these uncommon genetic characteristics: (i) the VP6 gene differed from the four other known subgroups; (ii) the NSP4 gene was identified as NSP4 genetic group C, an uncommon group in humans; and (iii) the NSP5/6 gene was most closely related with T152, a G12P[9] rotavirus previously isolated in Thailand. The finding of uncommon G3P[10] rotavirus in this pediatric patient provided additional evidence of the genetic diversity of human group A rotaviruses in Chiang Mai, Thailand. ***J. Med. Virol. 81:176–182, 2009.***



**KEY WORDS:** rotavirus; G3P[10] genotype; VP4; VP6; VP7; NSP4; NSP5/6

## INTRODUCTION

Group A rotaviruses are one of the major causes of severe gastroenteritis in young children, and are associated with 454,000–705,000 deaths annually among children under 5 years of age, predominantly in developing countries [Parashar et al., 2006]. Rotavirus belongs to the *Reoviridae* family, which contains 11 segments of double-stranded RNA genome. The two outer-layer proteins VP7 and VP4 form the basis of the current dual classification system of group A rotavirus into G and P genotypes [Estes and Kapikian, 2007]. To date, at least 15 G (G1–G15) and 27 P (P[1]–P[27]) genotypes of rotaviruses have been identified globally, with various combinations of G and P genotypes [Estes and Kapikian, 2007; Martella et al., 2007; Khamrin et al., 2007a; Steyer et al., 2007a]. The inner capsid protein VP6 bears the subgroup (SG) specificities that allows the classification of group A rotavirus into SG I, SG II, SG (I + II), and SG non-(I + II) based on reactivity with SG specific monoclonal antibodies (MAbs) [Greenberg et al., 1983a,b; Hoshino et al., 1987; Gorziglia et al., 1988; Urasawa et al., 1990; Iturriza-Gomara et al., 2002]. The non-structural glycoprotein, NSP4, plays an important role in rotavirus morphogenesis, pathogenesis, and enterotoxic activity. Sequence analyses of the NSP4 genes revealed the presence of at least six distinct NSP4

Grant sponsors: JSPS Postdoctoral Fellowships, Ministry of Education and Sciences and the Ministry of Health, Labor and Welfare, Japan; Grant sponsor: Faculty of Medicine, Chiang Mai University, Thailand (partial support).

*Correspondence to: Hiroshi Ushijima, MD, PhD, Aino Health Science Center, Aino University, 2-17-3 Shibuya, Shibuya-ku, Tokyo 150-0002, Japan.
E-mail: ushijima-hiroshi@jcom.home.ne.jp

Accepted 15 August 2008
DOI 10.1002/jmv.21336
Published online in Wiley InterScience (www.interscience.wiley.com)

genetic groups among human and animal rotaviruses, termed genetic groups A–F [Horie et al., 1997; Kirkwood and Palombo, 1997; Ciarlet et al., 2000; Mori et al., 2002; Khamrin et al., 2008]. Up to now, the precise role(s) of the NSP5/6 proteins encoded by gene 11 has only been partially characterized. A formal classification system of NSP5/6 genes of rotavirus has not yet been well established.

In recent decades, extensive epidemiological studies of rotavirus have been carried out. Those epidemiological studies have demonstrated that rotavirus genotypes G1, G2, G3, G4, and G9 in combination with P[8] and P[4] genotypes are the most common combinations associated with human rotavirus infection globally [Gentsch et al., 2005; Santos and Hoshino, 2005]. Recently, epidemiological surveillance to monitor the appearance of novel or unusual rotavirus antigenic types has been intensified throughout the world, yielding evidence for the increasing antigenic diversity of group A rotaviruses. Several unusual or animal-like rotavirus strains have been identified sporadically in humans, that is, G3P[3], G3P[9], G5P[6], G6P[14], G10P[14], G11P[25], and G12P[8] [Uchida et al., 2006; Khamrin et al., 2006a, 2007b; Duan et al., 2007; Ghosh et al., 2007; Steyer et al., 2007b]. In addition, the increasing data on the onset of animal-like rotavirus strains in the human population has demonstrated the importance of direct interspecies transmission of animal strains into humans and genetic reassortment between human and animal rotavirus strains. Therefore, information on the distribution of rotavirus G and P genotypes among humans and various animal species is important for understanding rotavirus ecology, and the mechanism by which rotaviruses evolve to cross species barriers, exchange their genes through reassortment event, and accumulation of single-point mutations and/or via genetic rearrangements.

Rotavirus G3 strains have been detected in various host species including humans, monkeys, rabbits, pigs, birds, cats, dogs, horses, mice, cows, and lambs [Estes and Kapikian, 2007]. A search of the literature shows that the P[10] rotavirus genotype is only restrictedly found in humans. Thus far, few P[10] rotavirus strains have been reported. However, each of these P[10] rotavirus strains carried different G–P genotype combinations. The strains 69M and B37 were G8P[10], strain 57M was G4P[10] [Estes and Kapikian, 2007], and two other strains found in Ghana were G9P[10] genotypes [Armah et al., 2003].

The human P[10] rotavirus prototype strain 69M was recovered from an infant with diarrhea in Indonesia [Hasegawa et al., 1984]. This rotavirus prototype strain displays a "super short" RNA pattern as demonstrated by PAGE and RNA–RNA hybridization, and was found to be a G8 genotype by VP7 gene sequence analysis [Matsuno et al., 1985; Green et al., 1989]. In addition, VP4 gene sequence analysis revealed low sequence identity with other rotavirus reference strains, and it was later designated as the P[10] rotavirus genotype [Qian and Green, 1991].

In the present study, P[10] was found in combination with G3 in a human rotavirus strain CMH079/05 which was isolated from a child hospitalized with severe diarrhea. The genetic makeup of this strain has been characterized by analyses of the VP4, VP6, VP7, NSP4, and NSP5/6 genes.

## MATERIALS AND METHODS

### Rotavirus Detection

During an epidemiological surveillance of group A rotavirus infection in Chiang Mai, Thailand, in 2005, a total of 147 specimens were collected from acute gastroenteritis pediatric patients. Of these, 43 (29.3%) were positive for group A rotavirus by reverse transcription-polymerase chain reaction (RT-PCR) and multiplex-PCR. The age of the patients ranged from neonate up to 5 years old.

### RNA Extraction, RT-PCR and Multiplex-PCR for G and P Genotyping

Group A rotavirus G and P genotypes were determined by RT-PCR and then followed by multiplex-PCR using type-specific primers. Viral dsRNA was extracted from 10% fecal supernatant using the QIAamp viral RNA Mini Kit (Qiagen, Hilden, Germany). The extracted dsRNA was denatured in 50% dimethylsulfoxide at 95°C for 5 min. The RT-PCR was carried out according to the methods described by Gouvea et al. [1990] and Gentsch et al. [1992]. For PCR amplification of the VP7 gene, a 1,062 bp fragment was generated using Beg9 (forward) and End9 (reverse) primers. For PCR amplification of the partial VP4 gene, a 876 bp fragment was generated using Con3 as a forward primer and Con2 as a reverse primer. The G genotyping was performed using a pool of different primers specific for G1–G4, G8, and G9. The VP4 gene characterization was performed using a pool of different primers specific for P[4], P[6], and P[8]–P[10]. During this epidemiological survey, an unusual G3P[10] rotavirus strain, CMH079/05, was isolated from a child hospitalized with acute gastroenteritis. In order to determine the genetic backgrounds of this uncommon rotavirus genotype, the VP4, VP6, VP7, NSP4, and NSP5/6 genes were characterized further by nucleotide sequence and phylogenetic analyses.

### Amplification of VP6, NSP4, and NSP5/6 Genes

The full length of the VP6 gene was amplified by VP6-5F and VP6-3R primer pairs [Khamrin et al., 2006a], which were slightly modified from the original VP6 specific primers described by Shen et al. [1994]. The NSP4 full-length gene was amplified by NSP4-1a and NSP4-2b primer pairs [Kudo et al., 2001]. The full-length NSP5/6 gene was amplified with primers GEN-NSP5F and GEN-NSP5R [Matthijnssens et al., 2006].

### Sequence and Phylogenetic Analyses

The PCR amplicons were purified with a Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI) and sequenced in both directions using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an automated sequencer (ABI 3100; Applied Biosystems). The nucleotide and deduced amino acid sequences of VP4, VP6, VP7, NSP4, and NSP5/6 genes were compared with those of reference strains available in the NCBI (National Center for Biotechnology Information) GenBank database using the BLAST (Basic Local Alignment Search Tool) server [Altschul et al., 1990]. Phylogenetic and molecular evolutionary analyses were conducted using MEGA, version 4 [Tamura et al., 2007].

### Nucleotide Sequence Accession Numbers

The nucleotide sequences of VP4, VP6, VP7, NSP4, and NSP5/6 genes of an unusual G3P[10] rotavirus strain have been deposited in GenBank under the accession numbers EU791922, EU791923, EU791924, EU791925, and EU791926, respectively.

## RESULTS

### Analysis of VP7 Gene Sequence

The complete nucleotide and deduced amino acid sequences of the VP7 gene of CMH079/05 strain were determined and compared to those of known G1 to G15 genotypes and also with several other G3 reference strains of human and animal origins. The VP7 sequence of CMH079/05 was most closely related to those of other G3 rotavirus reference strains (81.2–98.7% on nucleotide and 92.5–99.6% on amino acid levels), with the highest identity to an unusual G3 human rotavirus strain, CMH222, which was detected previously in the same surveillance site (Chiang Mai city) in 2001, at 98.7% on the nucleotide and 99.6% on the amino acid levels. The rotaviruses representing other G types exhibited far lesser sequence identities (63.7–82.3% at the nucleotide and 64.6–90.5% at the amino acid levels) with the CMH079/05 strain.

The phylogenetic tree of VP7 nucleotide sequence confirmed that the CMH079/05 strain formed a cluster with G3 rotavirus reference strains and showed closest lineage with CMH222, an uncommon G3 human rotavirus strain (Fig. 1A). In addition, the phylogenetic tree revealed two major lineages of G3 rotaviruses (Fig. 1A). Lineage I consisted mostly of G3 rotaviruses that derived from humans, except for a couple of porcine and bovine rotavirus strains, which were phylogenetically classified in this lineage. In contrast, lineage II consisted almost entirely of G3 strains of animal origins, except for three G3 strains, B4106, CMH079/05, and CMH222, which derived from humans. It was interesting to note that despite isolation from humans, CMH079/05 clustered together with non-human G3 strains. In fact, CMH079/05 located closely in a monophyletic branch with CMH222, which has been reported

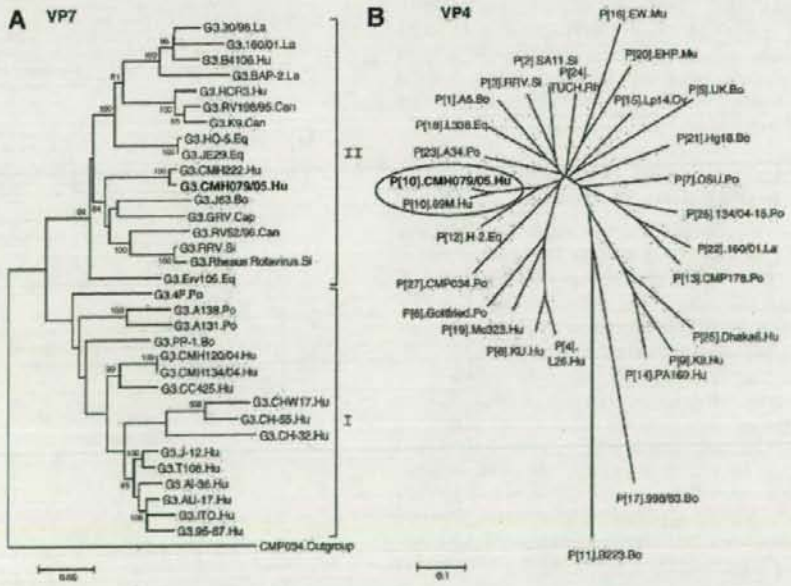



Fig. 1. Phylogenetic analyses of the nucleotide sequences of the VP7 and VP4 genes of CMH079/05 human rotavirus strain. **A**: relationship of VP7 sequence of CMH079/05 with other G3 strains derived from humans and animal species. **B**: comparison of VP4 sequence of CMH079/05 with all the existing P genotypes. The tree was constructed based on the neighbor-joining method using the MEGA 4 program. The scale bars indicate the branch lengths for 5% and 10% nucleotide differences in A,B, respectively.

as an unusual strain of human rotavirus carrying several genes of animal rotavirus genetic background [Khamrin et al., 2006a].

### Analysis of VP4 Gene Sequence

The partial VP4 gene of CMH079/05 rotavirus strain was sequenced and compared with those of the established reference strains of P[1]–P[27] available in the GenBank database. It was observed that partial VP4 sequence of CMH079/05 shared the highest sequence identity with a prototype strain of P[10] rotavirus (strain 69M) at the nucleotide and amino acid levels of 92.9% and 94.6%, respectively. Based on the data of P[10] rotavirus strains previously reported thus far (69M, 57M, B37, and two strains from Ghana), only the VP4 sequence of strain 69M (accession number M60600) had been deposited in the GenBank. Therefore, the partial VP4 sequence of CMH079/05 was the second sequence of the P[10] rotavirus available in the GenBank database.

Comparing the VP4 sequence of CMH079/05 with those of other existing P genotypes revealed that the nucleotide and amino acid sequence identities ranged from 50.8% to 76.4% and 44.7% to 80.8%, respectively. A high degree of nucleotide (92.9%) and amino acid (94.6%) sequence identities of the CMH079/05 with the 69M rotavirus prototype strain suggested that CMH079/05 belonged to genotype P[10]. A phylogenetic tree constructed from the VP4 sequences of all rotavirus genotypes recognized to date from both human and non-human origins (Fig. 1B) revealed relatedness between CMH079/05 and 69M. The CMH079/05 and 69M clustered closely together in a monophyletic branch separated from other rotavirus genotypes. This finding may imply a common evolutionary origin of the VP4 genes of the CMH079/05 and 69M prototype strains.

### Analysis of VP6, NSP4, and NSP5/6 Gene Sequences

Comparative analysis of the nucleotide and deduced amino acid sequences of full-length VP6 with those of four representative established subgroups (SG I, SG II, SG (I + II), and SG non-(I + II)) revealed that the VP6 sequence of the CMH079/05 strain was most closely related with that of human rotavirus strain CMH222, at 84.9% on nucleotide and 97.7% on amino acid levels. A phylogenetic tree constructed from the representative VP6 sequences of four rotavirus SGs recognized to date confirmed that human CMH079/05 strain clustered with CMH222 rotavirus strains, but was related distantly to four other SGs (Fig. 2A). The overall picture from this analysis demonstrates the distinction between the VP6 sequence of CMH079/05 and those of the other representative strains of SG I, SG II, SG (I + II), and SG non-(I + II).

Analysis of the NSP4 sequence revealed that the full-length NSP4 sequence of human rotavirus strain CMH079/05 was most closely related to those of other NSP4 genetic group C reference strains ranging from

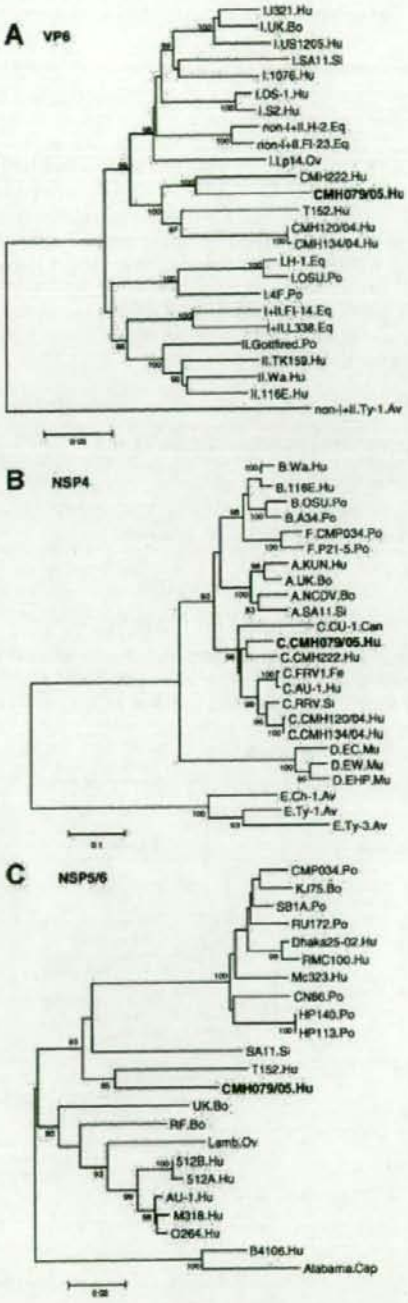



Fig. 2. Phylogenetic analyses of the nucleotide sequences derived from the VP6 (**A**), NSP4 (**B**), and NSP5/6 (**C**) genes of the CMH079/05 human rotavirus strain. The tree was constructed based on the neighbor-joining method using the MEGA 4 program. The scale bars indicate the branch lengths for 5%, 10%, and 2% nucleotide differences in A–C, respectively.

81.6% to 90.3% on the nucleotide level. The highest sequence identity was found with the CMH222 strain at 90.3% and 96.5% on nucleotide and amino acid levels, respectively. A phylogenetic tree constructed from the nucleotide sequence of CMH079/05 and those of other NSP4 genetic group reference strains derived from both human and non-human origins is shown in Figure 2B. It was, again, found that CMH079/05 clustered together with other NSP4 genetic group C rotavirus reference strains, and appeared to be most closely related to the CMH222 strain. Of note, even though rotavirus strains CMH079/05, CMH222, CMH120/04, and CMH134/04 all belong to the same NSP4 genetic group C and were isolated from the same geographical area (Chiang Mai City), they belonged to different lineages. The CMH079/05 and CMH222 clustered closely together in the same branch, which was separated from a branch of the CMH120/04 and CMH134/04 rotavirus strains (Fig. 2B).

The complete nucleotide sequence of the NSP5/6 gene of CMH079/05 was also analyzed. Compared to the other NSP5/6 sequences deposited in the GenBank database, the CMH079/05 showed the highest degree of sequence identity with human rotavirus strain T152, which was isolated previously in Bangkok, Thailand, in 1998 [Pongsuwanna et al., 2002]. Phylogenetic analysis of the NSP5/6 gene also revealed that the CMH079/05 strain clustered in an exclusive branch with the T152 rotavirus strain (Fig. 2C).

## DISCUSSION

The studies of rotavirus infection carried out in Thailand revealed that rotaviruses are the leading etiologic pathogens that causes diarrhea in children, and are responsible for about 27–58% of diarrheal diseases in hospitalized cases [Maneekarn and Ushijima, 2000; Jiraphongsa et al., 2005; Khamrin et al., 2006b]. Epidemiological surveillance of group A rotavirus infection conducted in Chiang Mai, Thailand, during the period of 2000–2004 [Khamrin et al., 2006b, 2007c] revealed that G9P[8] emerged as the most prevalent genotype (91.6%) from 2000 to 2001. It continued to be the most predominant strain in 2002, and then the prevalence rate abruptly decreased to 16.7% and 32.1% in 2003 and 2004, respectively. In addition, G2P[4] reemerged in the epidemic season of 2003, whereas G1P[8] became the most predominant genotype in 2004. During these extensive epidemiological studies, rotavirus strains with unusual G–P combinations were occasionally detected in children hospitalized with acute gastroenteritis; for example, an unusual strain of human rotavirus G3P[3] (CMH222) bearing simian-like VP7 and caprine-like VP4 genes was isolated from a 2-year-old child patient in 2001 [Khamrin et al., 2006a]. Recently, two other isolates (CMH120/04 and CMH134/04) of unusual human rotavirus G3P[9] genotype were detected during an epidemiological survey of human rotavirus infection in Chiang Mai, Thailand in 2004. Genetic analyses of VP4, VP6, VP7, and NSP4 genes of these G3P[9] strains revealed a close genetic relationship with rotavirus strains isolated previously from felines and humans [Khamrin et al., 2007b].

Although epidemiological surveillance and genotype identification of rotavirus infections in human and pig populations in Chiang Mai, Thailand, have been carried out extensively over the past two decades, the P[10] genotype had never been reported previously [Maneekarn and Ushijima, 2000; Maneekarn et al., 2006; Khamrin et al., 2006b, 2007c; Chan-It et al., 2008]. The present study conducted in 2005 described the detection of an unusual strain of G3P[10] (CMH079/05) in a 2-year-old boy admitted to hospital with acute gastroenteritis.

So far, few isolates of human rotavirus P[10] (strains 69M, 57M, B37, and 2 strains from Ghana) have been reported in the literature [Armah et al., 2003; Estes and Kapikian, 2007]. Analysis of the VP4 sequence of CMH079/05 revealed a high degree of sequence identity with strain 69M, a prototype strain of P[10] isolated in Indonesia [Hasegawa et al., 1984], suggesting that it belonged to genotype P[10]. Unfortunately, only the VP4 sequence of strain 69M is available in the GenBank database; none of the VP4 sequences of other P[10] strains could be accessed. Therefore, a comparison of VP4 sequences of our CMH079/05 with those of other P[10] rotaviruses was not possible.

Nevertheless, it was interesting to note that the P[10] rotavirus strains 69M and B37 were found in combination with G8, 57M with G4, and other two strains from Ghana with G9 genotypes. In contrast, CMH079/05, which is a P[10] rotavirus reported in the present study, was found in combination with the G3 genotype. To our knowledge, this is the first combination of G3 with P[10] genotype ever reported in the literature.

The CMH222 has been shown to carry VP6 and VP7 genes homologous to those of simian rotaviruses, while its VP4 and NSP4 genes were homologous to those of caprine rotaviruses [Khamrin et al., 2006a]. Interestingly, several gene segments (VP6, VP7, and NSP4) of our CMH079/05 strain were most closely related to those of CMH222, which was isolated previously from the same epidemiological area in Chiang Mai, Thailand [Khamrin et al., 2006a].

The detection of human rotaviruses carrying several genes of animal rotaviruses indicates an interspecies transmission of rotaviruses between humans and animal species. Additionally, the detections of multiple combinations of G3 with several P genotypes such as G3P[3] [Khamrin et al., 2006a], G3P[9] [Khamrin et al., 2007b], G3P[10] (in the present study), and G3P[19] [Maneekarn et al., 2006] in the same geographical area (Chiang Mai City) suggests that reassortment events among rotaviruses circulating in that area are taken place frequently. Furthermore, analysis of the NSP5/6 sequence of CMH079/05 revealed that it was most closely related with that of T152, a G12P[9] human rotavirus strain isolated in 1998 from an 11-month-old child admitted to hospital in Bangkok, Thailand [Pongsuwanna et al., 2002].

The findings of genetic relationships of CMH079/05 with those of uncommon human rotavirus strains CMH222, T152, and 69M, which were isolated at different times and geographical locations, suggests that these P[10] rotavirus strains may have existed and circulated in the Southeast Asian countries for a while, although with a low prevalence. In addition, the emergence of this human G3P[10] rotavirus strain carrying a genetic background closely related to several animal rotaviruses suggests that multiple interspecies transmissions and genetic reassortment events between human and animal rotaviruses may occur frequently under natural circumstances.

More genetic analyses of complete genome sequences would be helpful to elucidate the provisional evolution of this emerging virus. The increase in the frequency of detection of these uncommon rotavirus genotypes both in humans and several animal species raises questions concerning the sources of virus infection and the way that viruses spread in nature. Continued surveillance studies of rotavirus in both human and animal populations are important for understanding the overall picture of rotavirus distribution and the original source of these uncommon rotaviruses.

## REFERENCES


Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Armah GE, Steele AD, Binka FN, Esona MD, Asmah RH, Anto F, Brown D, Green J, Cutts F, Hall A. 2003. Changing patterns of rotavirus genotypes in Ghana: Emergence of human rotavirus G9 as a major cause of diarrhea in children. J Clin Microbiol 41:2317–2322.

Chan-It W, Khamrin P, Saekhow P, Pantip C, Thongprachum A, Peerakome S, Ushijima H, Maneekarn N. 2008. Multiple combinations of P[13]-like genotype with G3, G4, and G5 in porcine rotaviruses. J Clin Microbiol 46:1169–1173.

Ciarlet M, Linprandi F, Conner ME, Estes MK. 2000. Species specificity and interspecies relatedness of NSP4 genetic groups by comparative NSP4 analyses of animal rotaviruses. Arch Virol 145:371–387.

Duan ZJ, Li DD, Zhang Q, Liu N, Huang CP, Jiang X, Jiang B, Glass R, Steele D, Tang JY, Wang ZS, Fang ZY. 2007. Novel human rotavirus of genotype G5P[6] identified in a stool specimen from a Chinese girl with diarrhea. J Clin Microbiol 45:1614–1617.

Estes MK, Kapikian AZ. 2007. Rotaviruses. In: DM, Knipe Howley PM, Griffin DE, Martin RA, Lamb RA, Martin MA, Roizman B, Straus SE, editors. Fields virology, 5th edition. Philadelphia, PA: Lippincott Williams 1974.

Gentsch JR, Glass RI, Woods P, Gouvea V, Gorziglia M, Flores J, Das BK, Bhan MK. 1992. Identification of group A rotavirus gene 4 types by polymerase chain reaction. J Clin Microbiol 30:1365–1373.

Gentsch JR, Laird AR, Bielfelt B, Griffin DD, Banyai K, Ramachandran M, Jain V, Cunliffe NA, Nakagomi O, Kirkwood CD, Fischer TK, Parashar UD, Bresee JS, Jiang B, Glass RI. 2005. Serotype diversity and reassortment between human and animal rotavirus strains: Implications for rotavirus vaccine programs. J Infect Dis 192:146–159.

Ghosh S, Varghese V, Samajdar S, Sinha M, Naik TN, Kobayashi N. 2007. Evidence for bovine origin of VP4 and VP7 genes of human group A rotavirus G6P[14] and G10P[14] strains. J Clin Microbiol 45:2751–2753.

Gorziglia M, Hoshino Y, Nishikawa K, Maloy WL, Jones RW, Kapikian AZ, Chanock RM. 1988. Comparative sequence analysis of the genomic segment 6 of four rotaviruses each with a different subgroup specificity. J Gen Virol 69:1659–1669.

Gouvea V, Glass RI, Woods P, Taniguchi K, Clark HF, Forrester B, Fang ZY. 1990. Polymerase chain reaction amplification and typing of rotavirus nucleic acid from stool specimens. J Clin Microbiol 28:276–282.

Green KY, Hoshino Y, Ikegami N. 1989. Sequence analysis of the gene encoding the serotype-specific glycoprotein (VP7) of two new human rotavirus serotypes. Virology 168:429–433.

Greenberg H, McAuliffe V, Valdesuso J, Wyatt R, Flores J, Kalica A, Hoshino Y, Singh N. 1983a. Serological analysis of the subgroup protein of rotavirus, using monoclonal antibodies. Infect Immun 39:91–99.

Greenberg HB, Valdesuso J, van Wyke K, Midthun K, Walsh M, McAuliffe V, Wyatt RG, Kalica AR, Flores J, Hoshino Y. 1983b. Production and preliminary characterization of monoclonal antibodies directed at two surface proteins of rhesus rotavirus. J Virol 47:267–275.

Hasegawa A, Inouye S, Matsuno S, Yamaoka K, Eko R, Suharyono W. 1984. Isolation of human rotaviruses with a distinct RNA electrophoretic pattern from Indonesia. Microbiol Immunol 28:719–722.

Horie Y, Masamune O, Nakagomi O. 1997. Three major alleles of rotavirus NSP4 proteins identified by sequence analysis. J Gen Virol 78:2341–2346.

Hoshino Y, Gorziglia M, Valdesuso J, Askaa J, Glass RI, Kapikian AZ. 1987. An equine rotavirus (FI-14 strain) which bears both subgroup I and subgroup II specificities on its VP6. Virology 157:488–496.

Iturriza-Gomara M, Wong C, Blome S, Desselberger U, Gray J. 2002. Molecular characterization of VP6 genes of human rotavirus isolates: Correlation of genogroups with subgroups and evidence of independent segregation. J virol 76:6596–6601.

Jiraphongsa C, Bresee JS, Pongsuwanna Y, Kluabwang P, Poonawagul U, Arporntip P, Kanoksil M, Premsri N, Intusoma U. Rotavirus Surveillance Project Thailand Study Group. 2005. Epidemiology and burden of rotavirus diarrhea in Thailand: Results of sentinel surveillance. J Infect Dis 192:87–93.

Khamrin P, Maneekarn N, Peerakome S, Yagyu F, Okitsu S, Ushijima H. 2006a. Molecular characterization of a rare G3P[3] human rotavirus reassortant strain reveals evidence for multiple human-animal interspecies transmissions. J Med Virol 78:986–994.

Khamrin P, Peerakome S, Wongsawasdi L, Tonusin S, Sornchai P, Maneerat V, Khamwan C, Yagyu F, Okitsu S, Ushijima H, Maneekarn N. 2006b. Emergence of human G9 rotavirus with an exceptionally high frequency in children admitted to hospital with diarrhea in Chiang Mai, Thailand. J Med Virol 78:273–280.

Khamrin P, Maneekarn N, Peerakome S, Chan-It W, Yagyu F, Okitsu S, Ushijima H. 2007a. Novel porcine rotavirus of genotype P[27] shares new phylogenetic lineage with G2 porcine rotavirus strain. Virology 361:243–252.

Khamrin P, Maneekarn N, Peerakome S, Tonusin S, Phan TG, Okitsu S, Ushijima H. 2007b. Molecular characterization of rare G3P[9] rotavirus strains isolated from children hospitalized with acute gastroenteritis. J Med Virol 79:843–851.

Khamrin P, Peerakome S, Tonusin S, Malasao R, Okitsu S, Mizuguchi M, Ushijima H, Maneekarn N. 2007c. Changing pattern of rotavirus G genotype distribution in Chiang Mai, Thailand from 2002 to 2004: Decline of G9 and reemergence of G1 and G2. J Med Virol 79:1775–1782.

Khamrin P, Okitsu S, Ushijima H, Maneekarn N. 2008. Novel nonstructural protein 4 genetic group in rotavirus of porcine origin. Emerg Infect Dis 14:686–688.

Kirkwood C, Palombo EA. 1997. Genetic characterization of the rotavirus nonstructural protein NSP4. Virology 236:258–265.

Kudo S, Zhou Y, Cao XR, Yamanishi S, Nakata S, Ushijima H. 2001. Molecular characterization in the VP7, VP4 and NSP4 genes of human rotavirus serotype 4 (G4) isolated in Japan and Kenya. Microbiol Immunol 45:167–171.

Maneekarn N, Ushijima H. 2000. Epidemiology of rotavirus infection in Thailand. Pediatr Int 42:415–421.

Maneekarn N, Khamrin P, Chan-it W, Peerakome S, Sukchai S, Pringprao K, Ushijima H. 2006. Detection of rare G3P[19] porcine rotavirus strains in Chiang Mai, Thailand, provides evidence for origin of the VP4 genes of Mc323 and Mc345 human rotaviruses. J Clin Microbiol 44:4113–4119.

Martella V, Ciarlet M, Bányai K, Lorusso E, Arista S, Lavazza A, Pezzotti G, Decaro N, Cavalli A, Lucente MS, Corrente M, Elia G, Camero M, Tempesta M, Buonavoglia C. 2007. Identification of group A porcine rotavirus strains bearing a novel VP4 (P) Genotype in Italian swine herds. J Clin Microbiol 45:577–580.

Matsuno S, Hasegawa A, Mukoyama A, Inouye S. 1985. A candidate for a new serotype of human rotavirus. J Virol 54:623–624.

Matthijnssens J, Rahman M, Martella V, Xuelei Y, De Vos S, De Leener K, Ciarlet M, Buonavoglia C, Van Ranst M. 2006. Full genomic

analysis of human rotavirus strain B4106 and lapine rotavirus strain 30/96 provides evidence for interspecies transmission. J Virol 80:3801–3810.

Mori Y, Borgan MA, Ito N, Sugiyama M, Minamoto N. 2002. Diarrhea-inducing activity of avian rotavirus NSP4 glycoproteins, which differ greatly from mammalian rotavirus NSP4 glycoproteins in deduced amino acid sequence, in suckling mice. J Virol 76:5829–5834.

Parashar UD, Gibson CJ, Bresse JS, Glass RI. 2006. Rotavirus and severe childhood diarrhea. Emerg Infect Dis 12:304–306.

Pongsuwanna Y, Guntapong R, Chiwakul M, Tacharoenmuang R, Onvimala N, Wakuda M, Kobayashi N, Taniguchi K. 2002. Detection of a human rotavirus with G12 and P[9] specificity in Thailand. J Clin Microbiol 40:1390–1394.

Qian Y, Green KY. 1991. Human rotavirus strain 69M has a unique VP4 as determined by amino acid sequence analysis. Virology 182:407–412.

Santos N, Hoshino Y. 2005. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. Rev Med Virol 15:29–56.

Shen S, Burke B, Desselberger U. 1994. Rearrangement of the VP6 gene of a group A rotavirus in combination with a point mutation affecting trimer stability. J Virol 68:1682–1688.

Steyer A, Poljsak-Prijatelj M, Barlic-Maganja D, Jamnikar U, Mijovski JZ, Marin J. 2007a. Molecular characterization of a new porcine rotavirus P genotype found in an asymptomatic pig in Slovenia. Virology 359:275–282.

Steyer A, Poljsak-Prijatelj M, Bufon TL, Marcun-Varda N, Marin J. 2007b. Rotavirus genotypes in Slovenia: Unexpected detection of G8P[8] and G12P[8] genotypes. J Med Virol 79:626–632.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599.

Uchida R, Pandey BD, Sherchand JB, Ahmed K, Yokoo M, Nakagomi T, Cuevas LE, Cunliffe NA, Hart CA, Nakagomi O. 2006. Molecular epidemiology of rotavirus diarrhea among children and adults in Nepal: Detection of G12 strains with P[6] or P[8] and a G11P[25] strain. J Clin Microbiol 44:3499–3505.

Urasawa T, Taniguchi K, Kobayashi N, Wakasugi F, Oishi I, Minekawa Y, Oseto M, Ahmed MU, Urasawa S. 1990. Antigenic and genetic analyses of human rotavirus with dual subgroup specificity. J Clin Microbiol 28:2837–2841.

Journal of Medical Virology 81:183–191 (2009)

# Analysis of the VP6 Gene of Human and Porcine Group A Rotavirus Strains With Unusual Subgroup Specificities

**Aksara Thongprachum,[1] Pattara Khamrin,[2] Prayuth Saekhow,[3] Chansom Pantip,[1] Supatra Peerakome,[1] Hiroshi Ushijima,[2] and Niwat Maneekarn[1]***

[1]*Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand*
[2]*Aino Health Science Center, Aino University, Tokyo, Japan*
[3]*The Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai, Thailand*

Full-length VP6 amino acid sequences of human and porcine rotaviruses with subgroup (SG) (I + II) and SG non-(I + II) were analyzed in comparison with those of SG I and SG II. In human rotaviruses, the strains in the same SG shared a very high degree of amino acid identity, ranging from 97.4% to 99.4% for SG I, 95.9% to 100% for SG II, and 99.4% to100% for SG non-(I + II), while viruses in different SGs shared somewhat lower sequence identity at 90.4–93.1%. Conserved amino acids that distinguished the strains of SG I from SG II were observed at 21 positions. The viruses with SG non-(I + II) shared sequence identity with SG II as high as 97.2–99.7%, suggesting that they belonged to genogroup II. Similarly, porcine rotaviruses in the same SG shared 96.4–99.7% for SG I, 98.2–100% for SG II, 97.4–100% for SG (I + II), and 96.2–99.7% for SG non-(I + II), while strains in different SGs shared sequence identity ranging from 91.9% to 94.4%. Interestingly, the strains with SG (I + II) and SG non-(I + II) shared a high degree of sequence identity with SG I, at 96.4–100% and 94.7–99.7% respectively, suggesting that they are related to porcine SG I strains. The conserved amino acids which distinguished SG I from SG II were observed at 13 positions. The strains with SG I, SG (I + II), and SG non-(I + II) showed identical amino acid residues at these positions. Phylogenetic analysis strongly supported the findings of the sequence analysis. ***J. Med. Virol. 81:183–191, 2009.*** © 2008 Wiley-Liss, Inc.

**KEY WORDS:** subgroup; genogroup; VP6; human rotaviruses; porcine rotaviruses

## INTRODUCTION

Rotavirus is the most influential etiologic agent of severe diarrhea in infants and young children, as well as in young animals of many other species. Global mortality associated with rotavirus infection in children has been estimated at 454,000–705,000 deaths annually [Parashar et al., 2006]. Rotavirus has been classified into groups and subgroups based on the antigenic determinants of VP6 protein. Among them, Group A rotavirus is recognized as the most significant group with regard to its highest prevalence and pathogenesis in humans and various animal species [Gouvea et al., 1994; Estes, 2001].

Two monoclonal antibodies (MAbs) that specifically react with subgroup I (SG I) (MAb 255/60) or subgroup II (SG II) (MAb 631/9) rotavirus strains were developed in the early 1980s [Greenberg et al., 1983; Taniguchi et al., 1984] and have been widely used for characterization of human and animal rotavirus strains. The VP6 protein antigens allow the classification of these viruses into four subgroups: SG I, SG II, SG (I + II), and SG non-(I + II), depending on the presence or absence of SG-specific epitopes [Hoshino et al., 1987; Gorziglia et al., 1988; Estes and Cohen, 1989; Urasawa et al., 1990; Iturriza-Gomara et al., 2002]. The unreliability of serological methods for characterization of the VP6 protein into subgroups is a well-recognized problem. This may be the result of an accumulation of point mutations, reassortment, or rearrangement of the rotavirus genome, leading to amino acid changes on a particular epitope recognized by SG-specific MAbs.

Characterization of the VP6 deduced amino acid sequences of four different subgroups of rotavirus

Grant sponsor: Faculty of Medicine, Chiang Mai University (Medical Research fund); Grant sponsor: Graduate School of Chiang Mai University, Thailand (partial support).

*Correspondence to: Niwat Maneekarn, Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai 50200, Thailand. E-mail: nmaneeka@mail.med.cmu.ac.th

Accepted 2 September 2008
DOI 10.1002/jmv.21345
Published online in Wiley InterScience (www.interscience.wiley.com)

strains revealed that the VP6 gene encodes a protein of 397 amino acids in all strains with the exception of SG non-(I + II), which encodes a protein of 399 amino acids with the insertion of a CCACCA motif at position 911 [Gorziglia et al., 1988]. Five regions within the VP6 protein that might contribute to the SG epitopes have been proposed. Region A (amino acids 45, 56) and region C (amino acids 114, 120) might contribute to the SG I epitope while regions B (amino acids 83, 86, 89, 92), D (amino acids 312 or 314, 317 or 319) and E (amino acids 341 or 343, 350 or 352) might contribute to SG II epitope [Gorziglia et al., 1988]. Lopez et al. [1994] have demonstrated that a single amino acid change at position 172 (Met to Ala) or 305 (Asn to Ala) is sufficient to change the SG specificity of human rotavirus Wa from SG II to SG (I + II). These mutations allow the protein to be recognized by the SG I MAb 255/60, while retaining its capacity to interact with the SG II MAb 631/9. In the case of SG II, the mutation of two contiguous amino acids (Ala305 Asn306 to Asn305 Ala306) in the VP6 protein of porcine rotavirus YM (SG I) enables the protein to be recognized efficiently by the SG II MAb 631/9, while losing its capacity to interact with SG I MAb 255/60. Furthermore, it has been proposed that an Ala residue at positions 172 and 305 contribute to determining reactivity to SG I MAb (MAb 255/60), whereas a Glu residue at position 315 contributes to reactivity to SG II MAb (MAb 631/9) [Tang et al., 1997]. Analysis of the VP6 protein, which encompasses the amino acid positions in the regions implicated in the recognition by SG-specific MAbs, revealed that deduced amino acid sequences of human rotavirus strains determined serologically as SG II, SG (I + II), or SG non-(I + II) are indistinguishable from each other [Iturriza-Gomara et al., 2002]. The epitopes recognized by the SG-specific MAbs are thought to be a conformational structure produced by folding of VP6 or interaction between VP6 monomers to form a trimeric structure of the VP6 protein. It has been hypothesized that amino acids at the positions outside the antigenic regions of SG I or SG II may contribute to the reactivity of SG-specific MAbs.

It was, therefore, tempting to analyze the full-length deduced amino acid sequences of VP6 gene of rotavirus strains with SG (I + II) and SG non-(I + II) in comparison with those of SG I and SG II strains to elucidate this hypothesis.

## MATERIALS AND METHODS

### Subjects and Fecal Specimen Collection

The fecal specimens were collected from children hospitalized with diarrhea in four different hospitals and one private clinic in Chiang Mai province between May 2000 and December 2002. A total of 27 out of 429 stool samples with the subgroup specificities of our interest were selected for this study. In addition, fecal specimens were collected from diarrheic piglets during the period of June 2000 to July 2003 from six different farms located in Chiang Mai province. A total of 49 out of 428 stool samples were selected for the present study.

### Screening for Group A Rotavirus by ELISA

The presence of human and porcine Group A rotaviruses in fecal specimens was detected by ELISA using polyclonal antibody against Group A rotavirus as described previously [Hasegawa et al., 1987].

### Subgrouping of Group A Rotavirus by ELISA

Human and porcine Group A rotaviruses were examined for their subgroups using MAbs (Serotec, Oxford, UK) specific for SG I or SG II rotaviruses.

### RNA Extraction and Amplification of VP6 Gene

Rotavirus dsRNA genome was extracted from 20% fecal sample suspension with QIAamp Viral RNA Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. The VP6 gene was reverse transcribed and amplified using the consensus primer pair VP6-F/VP6-R [Shen et al., 1994; Khamrin et al., 2006]. The PCR amplification was performed using cDNA from reverse transcription reaction as the template under the following thermocycling conditions: 35 cycles of 94°C for 1 min, 50°C for 1 min, 72°C for 3 min, and final extension step at 72°C for 10 min in a master cycler.

### VP6 Nucleotide Sequencing

The VP6 full-length PCR products were gel-purified with QIAquick Gel Extraction Kit (QIAGEN). The purified products were sequenced in both directions using the BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA) on an automated sequencer (ABI 3100; Applied Biosystems). The VP6-F or VP6-R was used as a sequencing primer.

### Amino Acid Sequence and Phylogenetic Analyses

Nucleotide sequences of full-length VP6 genes of SG I, SG II, SG (I + II), and SG non-(I + II) rotaviruses were manually assembled and analyzed using ClustalX and BioEdit programs. The sequences were translated into amino acid sequences using GeneDoc version 2 software. The complete sequences were then compared to the reference sequences obtained from GenBank database. The phylogenetic tree was constructed based on the deduced amino sequences by neighbor-joining method using Molecular Evolutionary Genetics Analysis (MEGA 3.1) software.

### Nucleotide Sequence Accession Numbers

The nucleotide sequences of human and porcine rotavirus strains described in the present study have been deposited in GenBank. The accession numbers are given in parentheses. EU372724–EU372750 are human and EU372751–EU372799 are porcine rotaviruses.

Accession numbers of reference strains: S2 (DQ870488), 1076 (D00325), TK159 (AY661888), RV3 (U04741), 116E (U85998), E210 (U36240), Wa (K02086),

X57943 (X57943), OSU (AF317123), 4F (L29184), 4S (L29186), A131 (AF317124), A253 (AF317122), YM (X69487), JL94 (AY538664), Gottfried (POVPVP6), FI-14 (VPXR14), H-2 (VPXR15), B223 (AF317128), EW (U36474), PO-13 (BAA03836).

## RESULTS

### Group and Subgroups of Human and Porcine Rotaviruses

All rotavirus strains of human and porcine origins included in this study belonged to Group A.

Of a total of 27 strains of human Group A rotavirus selected for the present study, 3 belonged to SG I, 21 belonged to SG II, and 3 belonged to SG non-(I + II) as determined by ELISA using MAbs against SG I and SG II rotavirus antigens. The rotavirus strains that were reactive to neither SG I nor SG II MAbs were designated as SG non-(I + II); only 3 such isolates were found in our survey.

Of a total of 49 porcine Group A rotavirus strains included in this study, 4 strains were SG I, 3 were SG II, 24 were SG (I + II), and 18 were SG non-(I + II). The rotavirus strains that were reactive to both SG I and SG II MAbs were designated as SG (I + II) rotavirus. It should be noted that most of the porcine rotavirus strains detected in our survey exhibited SG (I + II) and SG non-(I + II) specificities and the representatives of these strains were included in the present study.

### Analysis of VP6 Deduced Amino Acid Sequences of Human Rotavirus Strains With SG I, SG II, and SG Non-(I + II) Specificities

In order to characterize the SG specificity of human rotavirus strains that exhibited SG non-(I + II) specificity, the full-length VP6 amino acid sequences of rotavirus strains with SG non-(I + II) were compared with those of SG I, SG II, and corresponding reference strains.

The VP6 amino acid sequence identities among rotavirus isolates of the same SG ranged from 97.4% to 99.4% for SG I (CMH171/01, CMH190/01), 95.9% to 100% for SG II (CMH5/00, CMH77/00, CMH8/01, CMH127/01, CMH142/01, CMH202/01, CMH4/02, CMH55/02, CMH95/02), and 99.4% to 100% for SG non-(I + II) (CMH150/01, CMH185/01, CMH186/01).

It was interesting to note that CMH5/00, which was initially assigned as a SG I strain based on reactivity to SG I-specific MAb, showed a very high degree of sequence identity (97.2–99.7%) with SG II strains. It was, therefore, designated as a Genogroup II strain (SG I*) based on VP6 sequence analysis.

Comparison of VP6 amino acid sequences between the viruses of different SGs revealed somewhat lower sequence identities; 90.4–93.1% between SG I and SG II, and 91.4–92.6% between SG I and SG non-(I + II). Surprisingly, SG non-(I + II) shared a high degree of sequence identity with SG II and SG I* as high as 97.2–99.7%, which lay within the range of sequence identity of the virus in the same SG. Therefore, based on VP6 sequence analysis, the viruses with SG non-(I + II) were designated as Genogroup II.

Alignment of the full-length VP6 deduced amino acid sequences of human rotavirus strains (Table I) showed conserved amino acid substitutions at least 21 positions to distinguishing between the strains in SG I and SG II at residues 39, 109 (Ile → Val), 60 (Asn or Ser → Thr), 83 (Asn → Thr), 86, 369 (Asp → Glu), 89, 92 (Val → Ile), 101, 217 (Val → Ala), 115, 120, 348 (Ser → Ala), 151 (Thr → Val), 172 (Ala → Met), 248 (Tyr → Phe or Leu), 305 (Ala → Asn), 310 (Asn → Gln), 315 (Glu → Gln), 339 (Ser → Asn), and 342 (Met → Leu). For the strains that were serologically identified as SG non-(I + II) (CMH150/01, CMH185/01, CMH186/01), the residues at these positions were identical with those of SG II strains over the entire sequence of VP6 (Table I).

Therefore, based on the VP6 amino acid sequence analysis, the rotaviruses with SG non-(I + II) specificity were designated as Genogroup II strains, while the rotaviruses with SG I and SG II were designated as Genogroup I and Genogroup II, respectively.

Phylogenetic analysis of VP6 amino acid sequences of these rotavirus isolates confirmed the data from amino acid sequence alignment; the strains with SG I specificity clustered together within a Genogroup I cluster while the strains with SG II and SG non-(I + II) specificities formed a separate cluster of Genogroup II (Fig. 1). Two genetic lineages, IA and IB, were distinguished within Genogroup I and four genetic lineages, IIA, IIB, IIC, and IID were observed within Genogroup II. However, there was no correlation between serologically determined SG and genetic lineages (Fig. 1).

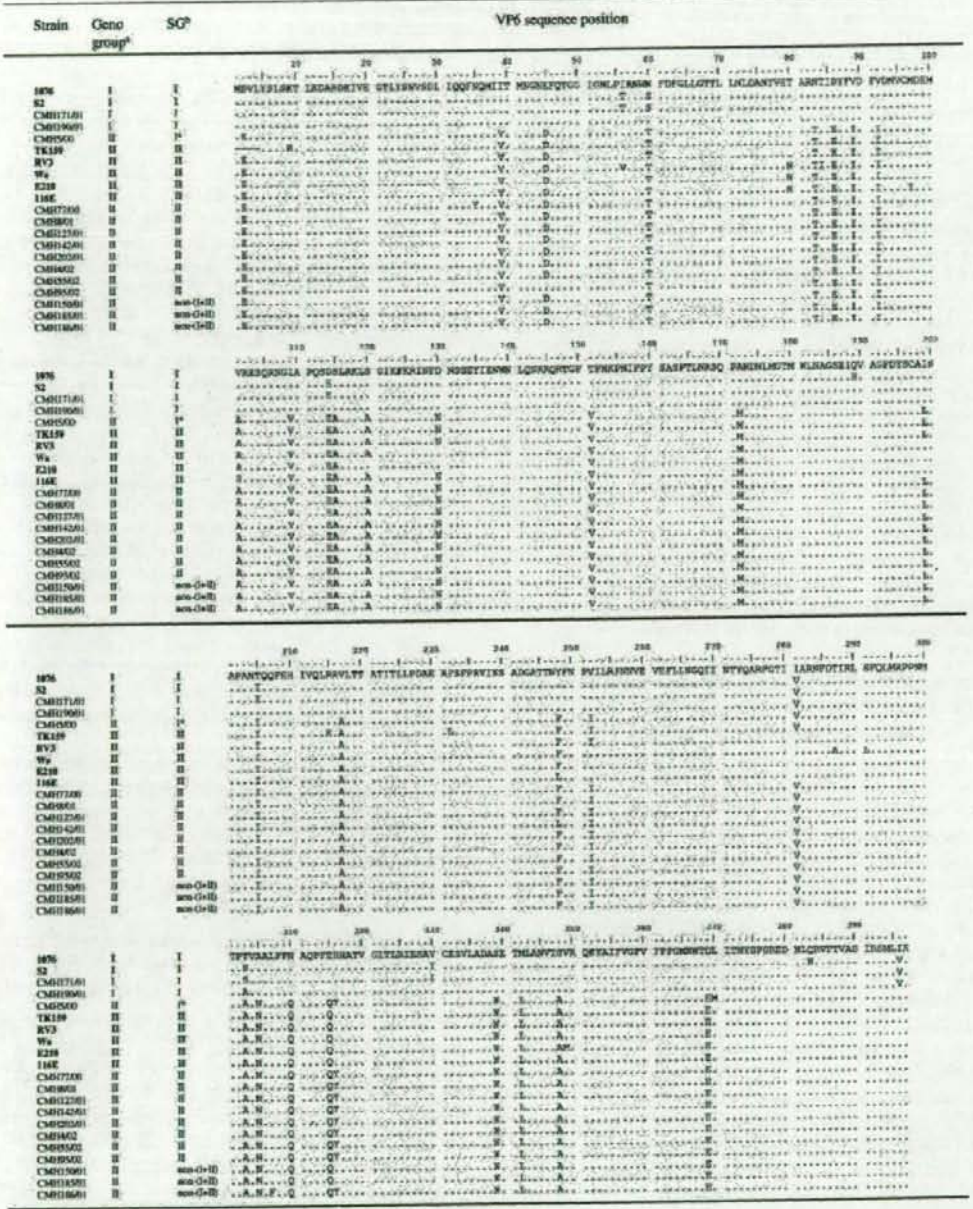
In addition, the phylogenetic tree revealed a clustering of CMH5/00 in a monophyletic branch together with rotavirus strains of SG II and SG non-(I + II) within lineage IID. This phylogenetic data confirmed the amino acid sequence analysis (Table I) which showed that CMH5/00 belonged to Genogroup II.

### Analysis of VP6 Deduced Amino Acid Sequences of Porcine Rotavirus Strains With SG I, SG II, SG (I + II), and SG Non-(I + II) Specificities

The VP6 deduced amino acid sequences of porcine rotavirus strains with SG (I + II) and SG non-(I + II) were compared with those of SG I, SG II, and corresponding reference strains. The VP6 amino sequence identities among rotavirus isolates of the same SG ranged from 96.4% to 99.7% for SG I (CMP39/00, CMP34/01), 98.2% to 100% for SG II (CMP100/01, CMP101/01), 97.4% to 100% for SG (I + II) (CMP52/01, CMP66/01, CMP39/02, CMP54/02, CMP107/02, CMP8/03), and 96.2% to 99.7% for SG non-(I + II) (CMP34/00, CMP29/01, CMP105/01, CMP66/02, CMP12/03).

It is interesting to point out that CMP127/01, which was initially assigned as a SG II strain based on reactivity to SG II-specific MAb, shared a very high degree of sequence identity (95.2–98.7%) with SG I

TABLE I. Alignment of the Deduced Amino Acid Sequences of the VP6 of Human Rotavirus Strains With SG I, SG II, and SG Non-(I + II) Specificities




[a]Genogroup was assigned based on VP6 amino acid sequence analysis.
[b]Subgroup was assigned based on reactivity to SG I- and SG II-specific MAbs.
*CMH5/00 was initially assigned as SG I based on reactivity to SG I-specific MAb and was later assigned as SG I*, a genogroup II based on VP6 amino acid sequence analysis. Reference strains are indicated in boldface.