 10.1111/j.1469-0691.2007.01918.x

Use of sequence analysis of the VP4 gene to classify recent Vietnamese rotavirus isolates

T. A. Nguyen[1,2], L. P. Hoang[2], L. D. Pham[2], K. T. Hoang[2], S. Okitsu[1], M. Mizuguchi[1] and H. Ushijima[1]

[1]Department of Developmental Medical Sciences, Institute of International Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan and [2]Children's Hospital 1, Ho Chi Minh City, Vietnam

ABSTRACT

Twenty-eight strains of P(8), four of P(4) and one of P(19) rotavirus, isolated in Ho Chi Minh City, Vietnam, during 2002–2003, were investigated by sequence analysis of the VP4 gene. Seven of the 28 P(8) rotavirus VP4 sequences clustered in the P(8)-3 lineage, or the rare, so-called OP354-like lineage. Amino-acid sequence comparison revealed that Vietnamese P(8)-3 rotaviruses were generally very similar to Malawian strains, including the prototype OP354 strain. The numerical severity scores of diarrhoeal disease caused by the Vietnamese P(8)-3 rotaviruses were statistically higher than those of diarrhoeal disease caused by rotaviruses in the more common P(8)-2 lineage. Sequence and phylogenetic analysis of the VP4 gene of a Vietnamese G9P(19) rotavirus isolate showed a high degree of homology with the cognate genes of other human and porcine rotaviruses, including the prototype 4F strain.

Keywords Diarrhoeal disease, epidemiology, phylogenetic analysis, rotavirus, sequence analysis, Vietnam

Original Submission: 18 June 2007; **Revised Submission**: 28 August 2007; **Accepted**: 8 October 2007

Clin Microbiol Infect 2008; **14**: 235–241

INTRODUCTION

Group A rotaviruses are members of the family Reoviridae and are a major cause of acute gastroenteritis in infants and young children. Rotaviruses cause *c.* 22% of childhood hospitalisations associated with diarrhoea, and are responsible worldwide for *c.* 611 000 childhood deaths annually [1]. The rotavirus genome consists of 11 segments of double-stranded RNA enclosed in a triple-layered capsid. The outer capsid layer is composed of two proteins: VP7, which defines G types (derived from glycoprotein); and VP4, which defines P types (derived from protease-sensitive protein) [2]. At least 15 different G genotypes and 27 P genotypes have been established, based on sequence analysis of the VP7 and VP4 genes, respectively [2–4]. Epidemiological studies worldwide have revealed that rotaviruses of types G1–G4, P(4) and P(8) are responsible for most infections, and four G–P combinations, G1P(8), G2P(4), G3P(8), and G4P(8), have been linked to 88.5% of the cases of rotavirus diarrhoea among children worldwide. However, these four G–P combinations accounted for <70% of rotavirus infections in South America, Asia and Africa [5].

Since VP7 is one of the two neutralising proteins, extensive studies on the molecular characteristics of the gene encoding this protein have been conducted. However, epidemiological studies of the distribution of P genotypes are also necessary for the development of a rotavirus vaccine, as sequence analyses of the VP4 gene could help to identify novel genotypes or subgenotypes in the human population, including recombinant viruses containing genes of animal origin. Data from epidemiological studies have indicated that P(8) is the predominant type, and that types P(8) and P(4) account for >90% of all cases of rotavirus diarrhoea in most parts of the world [5]. Sequence analysis of available P(8)

Corresponding author and reprint requests: H. Ushijima, Department of Developmental Medical Sciences, Institute of International Health, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan
E-mail: ushijima@m.u-tokyo.ac.jp

rotavirus strains has revealed that they can be divided into three lineages: the P(8)-1 or Wa-like lineage; the P(8)-2 or F45-like lineage; and the P(8)-3 or OP354-like lineage, with the last being a rare subtype comprising only a few rotavirus strains isolated from Malawi [6]. It has also been reported that, besides the dominance of the P(8) type, unusual P genotypes, e.g., P(19), which are normally predominant in animals, have been found increasingly in humans [7]. Further investigation of these strains could provide important insights into the evolution of rotaviruses, including the emergence of new P serotypes or evidence of human–animal recombinant viruses [8,9].

Vietnam is a developing country with a population of 80 million individuals. There are 1.52 million births annually, and a high mortality rate among children aged <5 years (42.2 deaths/1000/year), with 15.4% of deaths being related to diarrhoeal disease [10]. Several epidemiological studies of viruses causing acute gastroenteritis have been conducted, and these have revealed the significant impact of rotavirus on diarrhoeal disease in Vietnam [11–14]. The burden of rotavirus diarrhoea in Vietnam has also been estimated in order to consider the feasibility of a vaccine trial in the near future [15,16]. However, studies concerning the molecular characteristics of rotaviruses, especially the gene encoding the VP4 protein, have been very limited. The first study of the VP4 gene of Vietnamese rotavirus strains reported the presence of P(6) rotaviruses in Khanh Hoa Province, with these viruses probably originating from porcine rotavirus [17]. The present study determined the characteristics of the VP4 gene of rotaviruses belonging to the P(8), P(4) and P(19) genotypes that were detected during surveillance in Ho Chi Minh City in 2002–2003.

MATERIALS AND METHODS

Viruses and patients

A 1-year surveillance study of common viruses causing acute gastroenteritis among children in Ho Chi Minh City was performed and has been described in detail previously [13]. In brief, 1010 faecal specimens were collected from paediatric patients who were admitted to Children's Hospital 1 with a clinical diagnosis of acute gastroenteritis between October 2002 and September 2003. Disease severity was recorded using the 20-point Vesikari score [18]. Group A rotavirus was the most common cause, accounting for 681 (67.4%) of the 1010 cases. Among the 640 samples in which the P genotype could be determined, P(8) was the most frequent (362/640), followed by P(4), P(6) and P(19) (202, five and one cases, respectively). Infections involving more than one P genotype were found in 70 cases. Representative specimens from among the rotavirus-positive samples belonging to each P type were chosen randomly for investigation in the present study.

RT-PCR and nucleotide sequencing of the VP4 gene

A 687-bp fragment of the gene encoding the VP4 protein, including the majority of the hypervariable region VP8* cleavage and a small part of the more conserved region VP5*, was amplified by primers HumCom5 (sense, nucleotides 200–221, 5′-CTCTCGATGGTCCATATCAACC) [19] and Con2 (antisense, nucleotides 887–868, 5′-ATTTCGGACCATTTATA ACC) [20]. PCR products were purified using a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and were sequenced using a Big Dye Terminator Cycle Sequencing Kit v.3.1 and an ABI Prism 310 Genetic Analyser (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instruction. The HumCom5 and Con2 primers were also used for sequencing.

RT-PCR and nucleotide sequencing of the VP6 gene

The full length of the gene encoding the VP6 protein of the G9P(19) strain, VN375/2003, was amplified as described previously [21]. Primers VP6R (antisense, nucleotides 1356–1339, 5′-GGTCACATCCTCTCACTA), derived from primer 1 [21], and VP6_694 (sense, nucleotides 694–711, 5′-CCTTAT TACCAGATGCTG), designed for this study, were used for sequencing. Sequencing was performed as described above.

Sequence and phylogenetic analysis

Similarities of the sequenced strains with other strains were assessed by BLAST searches of partial nucleotide sequences using the default options (DNA DataBank of Japan). Multiple sequence alignments were calculated using the CLUSTALW program, and the phylogenetic trees were constructed using the neighbour-joining method with the MEGA 3.1 software package [22], with the different rotavirus sequences available in GenBank being used for comparisons and as outgroups.

RESULTS

The nucleotide sequences of a 687-bp fragment of the P gene were determined for 28 strains of P(8), four strains of P(4), and one strain of P(19) rotavirus. Results from the BLAST searches confirmed the P genotyping results obtained previously by nested type-specific PCR (data not shown). Table 1 summarises the clinical data of the patients studied. The age distribution of the patients ranged from 2 to 31 months. None showed dehydration, except for one patient, VN964, who was moderately dehydrated. Evaluation of diarrhoea severity, using a 20-point numerical score, showed a range between 7 and 15 points, with a mean score of 10.5 (Table 1).

Table 1. Clinical data and P genotypes of selected Vietnamese P(8) and P(4) rotavirus strains isolated from hospitalised children in Ho Chi Minh City, Vietnam, 2002–2003

Isolate	Date of collection	P type	Age (months)	Dehydration	Vesikari's score	Mean Vesikari score ± SE (number)[a]
VN545	April 2003	P(8)-3	15	None	9	12.6 ± 1.0 (n = 6)
VN546	April 2003	P(8)-3	7	None	12	
VN564	April 2003	P(8)-3	NA	NA	NA	
VN645	May 2003	P(8)-3	11	None	15	
VN827	July 2003	P(8)-3	10	None	15	
VN929	September 2003	P(8)-3	21	None	11	
VN952	September 2003	P(8)-3	9	None	14	
VN6	October 2002	P(8)-2	18	None	7	9.8 ± 0.7 (n = 16)
VN21	October 2002	P(8)-2	14	None	8	
VN23	October 2002	P(8)-2	8	None	8	
VN45	October 2002	P(8)-2	12	NA	NA	
VN318	February 2003	P(8)-2	NA	NA	NA	
VN368	March 2003	P(8)-2	NA	NA	NA	
VN517	April 2003	P(8)-2	15	None	13	
VN532	April 2003	P(8)-2	8	None	12	
VN537	April 2003	P(8)-2	6	None	8	
VN538	April 2003	P(8)-2	4	None	7	
VN543	April 2003	P(8)-2	NA	NA	NA	
VN544	April 2003	P(8)-2	31	None	15	
VN548	April 2003	P(8)-2	NA	NA	NA	
VN550	April 2003	P(8)-2	22	None	12	
VN553	April 2003	P(8)-2	10	None	12	
VN574	April 2003	P(8)-2	24	None	7	
VN621	April 2003	P(8)-2	9	None	8	
VN626	May 2003	P(8)-2	14	None	9	
VN828	July 2003	P(8)-2	30	None	9	
VN851	August 2003	P(8)-2	18	None	8	
VN964	September 2003	P(8)-2	9	Moderate	14	
VN271	February 2003	P(4)	12	None	12	9.5 ± 2.5 (n = 2)
VN322	February 2003	P(4)	NA	NA	NA	
VN580	April 2003	P(4)	4	None	7	
VN594	April 2003	P(4)	NA	NA	NA	

NA, not available or not known.
[a]Mean severity scores ± standard error of diarrhoeal disease caused by rotaviruses in lineages P(8)-3, P(8)-2, and P(4), respectively. Numbers of patients evaluated for the severity scores are also shown.

Sequence analysis of the VP4 gene of P(8) rotaviruses

The phylogenetic tree of the nucleotide sequences demonstrated that the 28 Vietnamese P(8) rotaviruses clustered into two different lineages, lineage P(8)-2 or F45-like (21/28), and lineage P(8)-3 or OP354-like (7/28) (Fig. 1). The results of amino-acid (aa) comparisons, spanning aa 86–263, for seven Vietnamese P(8)-3 rotavirus strains, revealed that they had a high degree of homology with other strains within lineage P(8)-3 (98.3–98.8%), and a lower homology with other P(8) rotaviruses in lineages 1 and 2 (85.9–88.2%) (data not shown). Amino-acid alignments within the same region are shown in Fig. S1 (see Supplementary material), and revealed that the highly conserved proline molecules at residues 224 and 225 were seen in all strains studied. The seven Vietnamese P(8)-3 strains were, in general, identical with other rotaviruses in the OP354-like lineage, within the studied region, and had several unique residues that were characteristic of OP354-like viruses (aa 140-F, aa 149-S, aa 178-K, aa 182-G, aa 184-V, aa 187-G, SD at position 191–192, aa 211-I, aa 221-Y). However, the Vietnamese P(8) rotavirus strains did not have the same conserved residues as some Malawian strains at two positions (aa 194, I instead of T; and aa 258, A instead of T).

In view of the hypothesis that the rare OP354-like viruses might cause more severe diarrhoeal disease than other viruses in common lineages because they could escape protective immunity in communities with common P(8) rotaviruses, the severity of diarrhoea in the patients was investigated using Vesikari's numerical score. The mean score for disease caused by rotaviruses in lineages P(8)-3, P(8)-2 and P(4) was 12.6 ± 1.0 (n = 6), 9.8 ± 0.7 (n = 16) and 9.5 ± 2.5 (n = 2), respectively. The difference between the mean score for the rare P(8)-3 lineage and that for the other lineages was statistically significant ($p < 0.05$).

Sequence analysis of the VP4 gene of P(4) rotaviruses

All four Vietnamese P(4) rotavirus strains in this study were identical with each other, and clustered into a single lineage with other recent P(4) rotaviruses isolated worldwide (Fig. 2).

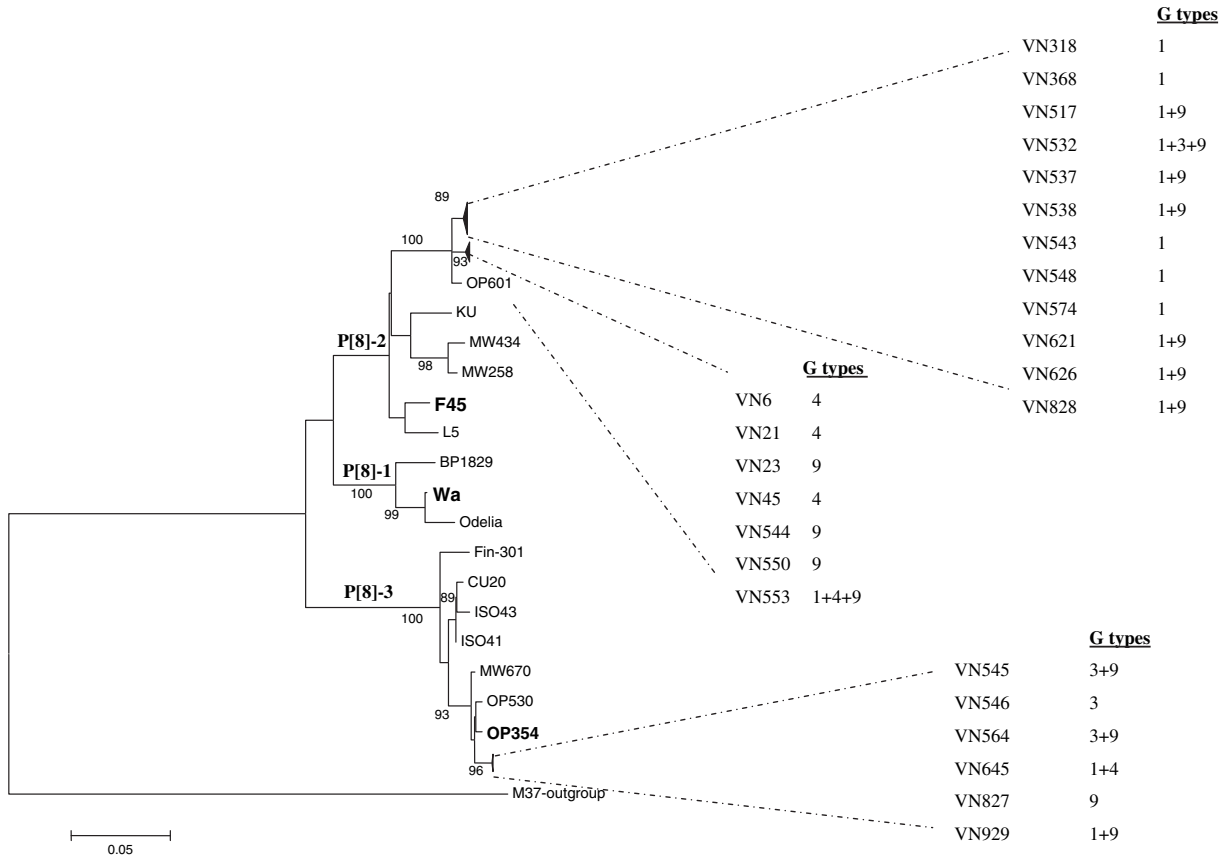

Fig. 1. Phylogenetic tree of the VP8* fragment of 28 Vietnamese P(8) rotavirus strains. Bootstrap values >70% are indicated at the branch nodes. The prototype strains of three lineages are in shown bold type. The associated G types are also shown next to each isolate. Three sub-lineages are indicated. The P(6) rotavirus strain M37 was used as an outgroup.

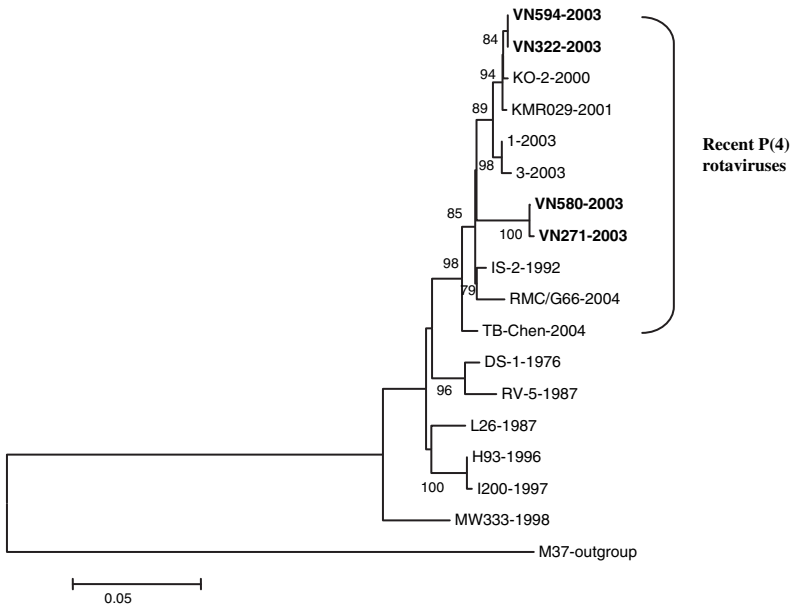

Fig. 2. Phylogenetic tree of the VP8* fragment of four Vietnamese P(4) rotavirus strains. Bootstrap values >70% are indicated at the branch nodes. Year of isolation is also shown. Vietnamese P(4) rotaviruses are shown in bold type. The P(6) rotavirus strain M37 was used as an outgroup.

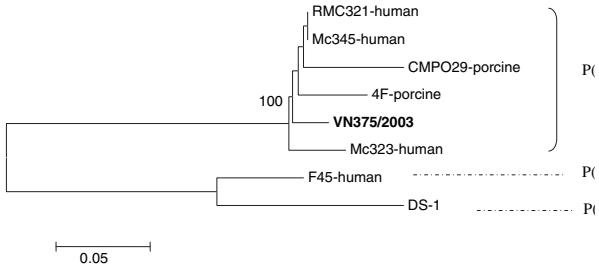

Fig. 3. Phylogenetic tree of the VP8* fragment of Vietnamese and other human and porcine P(19) rotavirus strains. The Vietnamese P(19) isolate is shown in bold type.

Sequence analysis of the VP4 and VP6 genes of a P(19) rotavirus

Nucleotide and amino-acid sequence analysis of the VP4 gene of the VN375/2003 isolate detected during 2002–2003 revealed that this strain had greater identity with P(19) rotaviruses (94.5–96.7% for nucleotides; 94.4–97.2% for amino-acids) than other P genotypes (24.2–68.1% for nucleotides; 44.6–72.2% for amino acids). The phylogenetic tree indicated that the Vietnamese P(19) strain clustered in the same group as other human and several porcine P(19) rotaviruses (Fig. 3). A deduced amino-acid comparison, from aa 88 to aa 273, showed that these P(19) strains were identical. However, the VN375 strain had several unique residues at positions aa 89-I, aa 147-N and aa 195-K. Interestingly, only the Vietnamese strain and the prototype porcine strain 4F had a valine at position 254 (Fig. S2, see Supplementary material).

A 462-bp nucleotide sequence of the VP6 gene of the VN375 strain was also analysed. The results revealed that this gene had closer identity with cognate genes of porcine rotavirus strains (91.1–93.7% for nucleotides; 98.7–99.3% for amino-acids) than with other human strains (82.4–83.7% for nucleotides; 91.5–92.2% for amino-acids).

DISCUSSION

Genotype P(8) has been found to be the most frequent P type among rotaviruses worldwide, with a prevalence of 51% in Africa and 87.8% in Europe [5]. Reports from several surveillance studies in Vietnam have also indicated a high prevalence of P(8) rotavirus (70.7–71.8%) [11,14]. The present study analysed the VP4 gene sequences of these isolates and found that, surprisingly, seven of 28 randomly chosen P(8) strains clustered in the rare P(8)-3, or so-called OP354-like lineage. Since the first report of OP354-like viruses in Malawi during 1998–1999 [6], these rare rotavirus strains have not been isolated elsewhere. A BLAST search of P(8) nucleotide sequences available from GenBank showed that only a few P(8) rotavirus strains had a high identity with the OP354 virus, including some strains from Malawi [6] and India [23], and one each from Thailand and Finland. The detection of these viruses from various countries in different continents suggests that these viruses are widespread around the world. The rather low rate of detection of these rare strains could be explained by large-scale mutation in the nucleotide sequence of the VP4 gene at the primer-binding site of the primer, 1T-1, specifically for P(8), which may make recognition of these rotavirus strains more difficult [6].

Cunliffe *et al.* [6] reported that Malawian P(8) rotaviruses predominantly comprised strains belonging to a distinct G type, in which the G4 type for OP354, the G3 type for MW258, and the G1 type for OP601 were clustered. In the present study, the phylogenetic tree indicated clearly that Vietnamese P(8) rotaviruses belonged to two separate clusters, the OP354 cluster and the OP601 cluster. The strains in the OP601 cluster could be divided further into two variants, in which one variant had a tendency to belong to G types 1 and/or 9, and the second variant had the characteristics of types G9 or G4. However, the concordance between P and G type in the OP354-like viruses was unclear (Fig. 1).

A rotavirus vaccine, which includes a P(8)-1 rotavirus strain as a component, has begun to be used widely in some countries. The finding in the present study that viruses of the P(8)-3 lineage caused, on average, more severe diarrhoeal disease may be important for the development and application of rotavirus vaccines, especially in countries in which OP354-like rotaviruses are prevalent.

The P(19) rotaviruses are known to commonly infect pigs, with the prototype virus being the porcine 4F strain [24]. To our knowledge, there have been only two reports of the isolation of the P(19) rotavirus in humans worldwide, one from Thailand and one from India [7,9]. The detection of this virus in Vietnam emphasises that the P(19)

rotavirus is circulating in Asia. Sequence analysis of the VP6 gene of the Vietnamese G9P(19) strain also showed closest identity with the cognate genes of other porcine rotaviruses, suggesting gene reassortment or a human infection by a virus similar to one that infects animals. Further investigation of this virus might provide interesting data concerning the relationship between human and porcine rotaviruses, especially in Asian countries in which farming is the main activity and close contact between humans and animals could promote inter-species transmission.

ACCESSION NUMBERS

The VP4 nucleotide sequences of the Vietnamese isolates described in this study have been deposited in GenBank under the following accession numbers: EF673752–EF673779 for the P(8) strains; and EF673780–EF673783 for the P(4) strains. The accession numbers of the nucleotide sequences of the VP4 and VP6 genes of the VN375/2003 strain are EF063154 and EU042135, respectively.

ACKNOWLEDGEMENTS

This research was supported by the Ministry of Education, Culture, Sports, Science, and Technology, Japan and the Ministry of Health, Labor, and Welfare, Japan. No information has been provided by the authors concerning the existence or absence of any other conflicting or dual interests.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article online at http://www.blackwell-synergy.com:

Fig. S1. Partial amino-acid sequence alignment (aa 86-263) for selected P(8) rotaviruses. Lineages and sub-lineages are indicated.

Fig. S2. Partial amino-acid sequence (aa 88-273) alignment within selected P(19) rotaviruses. Human and porcine rotaviruses are indicated.

REFERENCES

1. Parashar UD, Gibson CJ, Bresse JS, Glass RI. Rotavirus and severe childhood diarrhea. *Emerg Infect Dis* 2006; **12**: 304–306.
2. Estes MK. Rotaviruses and their replication. In: Knipe DM, Howley PM, Griffin DE *et al.*, eds, *Fields virology*, 4th edn. Philadelphia, PA: Lippincott, Williams and Wilkins, 2001; 1747–1785.
3. Rao CD, Gowda K, Reddy BSY. Sequence analysis of VP4 and VP7 genes of nontypeable strains identifies a new pair of outer capsid proteins representing novel P and G genotypes in bovine rotaviruses. *Virology* 2000; **276**: 104–113.
4. Martella V, Ciarlet M, Banyai K *et al.* Identification of group A porcine rotavirus strains bearing a novel VP4 (P) genotype in Italian swine herds. *J Clin Microbiol* 2007; **45**: 577–580.
5. Santos N, Hoshino Y. Global distribution of rotavirus serotypes/genotypes and its implication for the development and implementation of an effective rotavirus vaccine. *Rev Med Virol* 2005; **15**: 29–56.
6. Cunliffe NA, Gondwe JS, Graham SM *et al.* Rotavirus strain diversity in Blantyre, Malawi, from 1997 to 1999. *J Clin Microbiol* 2001; **39**: 836–843.
7. Urasawa S, Hasegawa A, Urasawa T *et al.* Antigenic and genetic analyses of human rotaviruses in Chiang Mai, Thailand: evidence for a close relationship between human and animal rotaviruses. *J Infect Dis* 1992; **166**: 227–234.
8. Okada J, Urasawa T, Kobayashi N *et al.* New P serotype of group A human rotavirus closely related to that of a porcine rotavirus. *J Med Virol* 2000; **60**: 63–69.
9. Varghese V, Das S, Singh NB *et al.* Molecular characterization of a human rotavirus reveals porcine characteristics in most of the genes including VP6 and NSP4. *Arch Virol* 2004; **149**: 155–172.
10. Canh DG, Do TT. General and diarrhoea related mortality rates in Vietnamese children. *J Prevent Med* 1993; **3** (suppl 12): 48 (in Vietnamese).
11. Doan LT, Okitsu S, Nishio O, Pham DT, Nguyen DH, Ushijima H. Epidemiological features of rotavirus infection among hospitalized children with gastroenteristis in Ho Chi Minh City, Vietnam. *J Med Virol* 2003; **69**: 588–594.
12. Nishio O, Matsui K, Lan DT, Ushijima H, Isomura S. Rotavirus infection among infants with diarrhea in Vietnam. *Pediatr Int* 2000; **42**: 422–424.
13. Nguyen TA, Yagyu F, Okame M *et al.* Diversity of viruses associated with acute gastroenteritis in children hospitalized with diarrhea in Ho Chi Minh City, Vietnam. *J Med Virol* 2007; **79**: 582–590.
14. Van Man N, Luan Le T, Trach DD *et al.* Epidemiological profile and burden of rotavirus diarrhea in Vietnam: 5 years of sentinel hospital surveillance, 1998-2003. *J Infect Dis* 2005; **192** (suppl 1): S127–S132.
15. Anh DD, Thiem VD, Fischer TK *et al.* The burden of rotavirus diarrhea in Khanh Hoa province, Vietnam: baseline assessment for a rotavirus vaccine trial. *Pediatr Infect Dis J* 2006; **25**: 37–40.
16. Fischer TK, Anh DD, Antil L *et al.* Health care costs of diarrheal disease and estimates of the cost-effectiveness of rotavirus vaccination in Vietnam. *J Infect Dis* 2005; **192**: 1720–1726.
17. Ahmed K, Anh DD, Nakagomi O. Rotavirus G5P(6) in child with diarrhea, Vietnam. *Emerg Infect Dis* 2007; **13**: 1232–1235.
18. Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. *Scand J Infect Dis* 1990; **22**: 259–267.
19. Gunasena S, Nakagomi O, Isegawa Y *et al.* Relative frequency of VP4 gene alleles among human rotaviruses recovered over a 10-year period (1982–1991) from Japa-

nese children with diarrhea. *J Clin Microbiol* 1993; **31**: 2195–2197.
20. Gentsch JR, Glass RI, Woods P *et al*. Identification of group A rotavirus gene 4 types by polymerase chain reaction. *J Clin Microbiol* 1992; **30**: 1365–1373.
21. Shen S, Burke B, Desselberger U. Rearrangement of the VP6 gene of a group A rotavirus in combination with a point mutation affecting trimer stability. *J Virol* 1994; **68**: 1682–1688.
22. Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 2004; **5**: 150–163.
23. Samajdar S, Varghese V, Barman P *et al*. Changing pattern of human group A rotaviruses: emergence of G12 as an important pathogen among children in Eastern India. *J Clin Virol* 2006; **36**: 183–188.
24. Burke B, McCrae MA, Desselberger U. Sequence analysis of two porcine rotaviruses differing in growth in vitro and in pathogenicity: distinct VP4 sequences and conservation of NS53, VP6 and VP7 genes. *J Gen Virol* 1994; **75**: 2205–2212.