JOURNAL OF VIROLOGY, Aug. 2008, p. 7336–7345 Vol. 82, No. 15
0022-538X/08/$08.00+0 doi:10.1128/JVI.02371-07

Genetic Analysis of the Capsid Gene of Genotype GII.2 Noroviruses[▽]

Nobuhiro Iritani,[1,2]* Harry Vennema,[1] J. Joukje Siebenga,[1] Roland J. Siezen,[3] Bernadet Renckens,[3] Yoshiyuki Seto,[4] Atsushi Kaida,[2] and Marion Koopmans[1]

Laboratory for Infectious Diseases and Perinatal Screening, National Institute for Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, The Netherlands[1]; Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, 8-34 Tojo-cho, Tennoji-ku, Osaka 543-0026, Japan[2]; Centre for Molecular and Biomolecular Informatics, Radboud University Medical Center, P.O. Box 9101, 6500HB Nijmegen, The Netherlands[3]; and Laboratory of Microbiology, Osaka Prefecture University, 1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan[4]

Received 2 November 2007/Accepted 5 May 2008

Noroviruses (NoVs) are considered to be a major cause of acute nonbacterial gastroenteritis in humans. The NoV genus is genetically diverse, and genotype GII.4 has been most commonly identified worldwide in recent years. In this study we analyzed the complete capsid gene of NoV strains belonging to the less prevalent genotype GII.2. We compared a total of 36 complete capsid sequences of GII.2 sequences obtained from the GenBank (n = 5) and from outbreaks or sporadic cases that occurred in The Netherlands (n = 10) and in Osaka City, Japan (n = 21), between 1976 and 2005. Alignment of all capsid sequences did not show fixation of amino acid substitutions over time as an indication for genetic drift. In contrast, when strains previously recognized as recombinants were excluded from the alignment, genetic drift was observed. Substitutions were found at five informative sites (two in the P1 subdomain and three in the P2 subdomain), segregating strains into five genetic groups (1994 to 1997, 1999 to 2000, 2001 to 2003, 2004, and 2005). Only one amino acid position changed consistently between each group (position 345). Homology modeling of the GII.2 capsid protein showed that the five amino acids were located on the surface of the capsid and close to each other at the interface of two monomers. The data suggest that these changes were induced by selective pressure, driving virus evolution. Remarkably, this was observed only for nonrecombinant genomes, suggesting differences in behavior with recombinant strains.

Noroviruses (NoVs) are an important cause of acute nonbacterial gastroenteritis in adults and children worldwide (13). NoVs are members of the family *Caliciviridae*, having a positive-sense single-stranded RNA genome. Their genome is organized into three open reading frames (ORFs). ORF1 encodes nonstructural proteins including the RNA-dependent RNA polymerase (RdRp), ORF2 encodes a major structural capsid protein including a shell (S) domain and a protruding (P) domain, and ORF3 encodes a minor structural protein (13, 18, 41). The S domain forms the inner part (shell) of the viral capsid, and the P domain forms the arch-like structures that protrude from the virion. The P domain is further divided into P1 and P2 subdomains that correspond to the sides and the top of the arch-like capsomeres, respectively (13, 31).

Based on the genetic analysis of the RdRp and capsid regions, human NoVs can be divided into three genogroups (Gs), GI, GII, and GIV (2, 14, 39), which further segregate into distinct lineages called genotypes (2, 20, 36, 37). Recently, Kageyama et al. (20) proposed that at least 31 genotypes could be distinguished within GI and GII. The GII.4 genotype, which is represented by the Lordsdale/93/UK strain, has been the most commonly identified genotype worldwide in recent years. Genetic characterization of strains belonging to this genotype have shown a sequence of variants that have arisen over time, suggesting that rapid genetic evolution of GII.4 NoVs may in part explain their successful spread and impact on people of all ages (5, 9, 27, 29, 30, 33, 40).

Viruses belonging to other NoV genotypes are found less consistently, causing sporadic outbreaks or temporary epidemics in a limited geographic region or time period (5, 17, 23, 26). As a result, far less is known about the population structure of these variants (28). The genetic analysis of other genotype NoVs excluding GII.4 will improve our understanding of genetic evolution and its relevance for the epidemiology of NoVs.

During the spring of 2004, an epidemic of GII.2 NoV (which is represented by the Melksham/94/UK strain [Melksham]) occurred in Osaka City, Japan. Our previous study of this regional epidemic described the molecular epidemiology of these GII.2 strains (17). Here, we describe the genetic characterization of GII.2 strains from those outbreaks in comparison with viruses detected over a 12-year period in the GenBank, The Netherlands, and Japan.

MATERIALS AND METHODS

GII.2 strains. A total of 36 NoV strains that had been characterized as GII.2 genotype were used for this study (Table 1). The capsid sequence data for five GII.2 strains were obtained from the GenBank. Of these, the Melksham strain and the Chesterfield/434/1997/US strain (11, 29) have been characterized as belonging to the GII.2 genotype on the basis of RdRp as well as capsid regions. The Snow Mountain/76/US strain has been characterized as a recombinant NoV, with a distinct (non-GII.2) RdRp region and a GII.2 capsid region (4, 16). The other two strains from the GenBank (Ina/02/JP and Buds/02/US) were characterized as GII.2 genotype on the basis of the capsid region, but their sequences of the RdRp region were unknown. The capsid sequence data of Buds/02/US strain lacked the first 6 nucleotides (nt) from the 5′ end of the capsid gene. Twenty-one GII.2 strains were obtained from outbreaks or sporadic cases detected in Osaka City, Japan, between April 1996 and March 2005. These were 21

* Corresponding author. Mailing address: Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, 8-34 Tojo-cho, Tennoji-ku, Osaka 543-0026, Japan. Phone: 81 6 6771 3147. Fax: 81 6 6772 0676. E-mail: n-iritani@city.osaka.lg.jp.

▽ Published ahead of print on 14 May 2008.

TABLE 1. GII.2 NoV strains used in this study

Source	Strain (abbreviation)[a]	RdRp sequence type[b]	Accession no.	Note[c]
GenBank	Snow Mountain/76/US (SM)	GII-NA	AY134748	
	Melksham/94/UK (Melksham)	GII.2	X81879	
	Chesterfield/434/97/US (CF434)	GII.2	AY054300	
	Ina/02/JP (Ina)	Unknown	AB195225	
	Buds/02/US (Buds)	Unknown	AY660568	
Osaka City, Japan	OC97049/97/JP (97049)	GII-NA	AB279553	O
	OC01243/01/JP (01243)	GII-NA	AB279554	O
	OC02012/02/JP (02012)	GII.2	AB279555	O
	OC02022/02/JP (02022)	GII-NA	AB279556	O
	OCS020289/02/JP (S020289)	GII.b	AB279570	S
	OC04038/04/JP (04038)[d]	GII.2	AB279557	O*
	OC04042/04/JP (04042)[d]	GII.2	AB279558	O*
	OC04043/04/JP (04043)[d]	GII.2	AB279559	O*
	OCS030697/04/JP (S030697)[d]	GII.2	AB279571	S*
	OC04056-1/04/JP (04056-1)[d]	GII.2	AB279560	O*
	OC04056-2/04/JP (04056-2)[d]	GII.2	AB279561	O*
	OC04059/04/JP (04059)	GII.2	AB279562	O*
	OCS040035/04/JP (S040035)	GII.2	AB279572	S*
	OC04067/04/JP (04067)[d]	GII.2	AB279563	O*
	OC04071/04/JP (04071)	GII.2	AB279564	O*
	OC04073/04/JP (04073)	GII.2	AB279565	O*
	OC04075/04/JP (04075)	GII.2	AB279566	O*
	OC04076/04/JP (04076)[d]	GII.2	AB279567	O*
	OCS040100/04/JP (S040100)	GII.2	AB279573	S*
	OC04169/04/JP (04169)	GII-NA	AB279568	O
	OC05010/05/JP (05010)	GII.b	AB279569	O
The Netherlands	Sensor99-191/99/NL (S99-191)	GII.2	AB281081	S
	OB0037-246/00/NL (OB0037)	GII.2	AB281082	O
	OB0048-318/00/NL (OB0048)	GII.2	AB281083	S
	OB0115-195/01/NL (OB0115)	GII.2	AB281084	O
	EP0125-006/01/NL (EP0125)	GII.2	AB281085	O
	EP0207-001/02/NL (EP0207)	GII.2	AB281086	O
	EP0239-001/02/NL (EP0239)	GII.2	AB281087	O
	OB0371-459/03/NL (OB0371)	GII.2	AB281088	O
	OB0528-158/05NL (OB0528)	GII.2	AB281089	O
	OB0587-470/05/NL (OB0587)	GII.2	AB281090	O

[a] NoV strains are arranged in chronology of detection from top (oldest) to bottom (most recent) for each source.
[b] NA, not assigned.
[c] O, outbreak; S, sporadic case; *, spring epidemic in 2004.
[d] These strains have identical amino acid sequences in the complete capsid gene and only one strain (OC04038/04/JP) has been used for long-term genetic analysis.

of the 23 GII.2 strains identified during a 9-year study period out of a total of 238 outbreaks and 200 sporadic cases of NoV infection. From the genetic analysis across the junction between the RdRp and the capsid regions, 6 of these 21 GII.2 strains have been characterized as recombinants, which have non-GII.2 RdRp regions and GII.2 capsid regions (Fig. 1) (17).

The strains from The Netherlands were collected from a 12-year study period. Between 1994 and 2005, GII.2 NoVs were detected in 13 (1.7%) out of 745 NoV-associated outbreaks and three sporadic cases in The Netherlands. Initially, these GII.2 NoVs were characterized by the comparison of sequences in the RdRp region (Fig. 1). The detection method and criteria for genotyping at the RdRp region have been previously described (8, 36). The complete capsid genes of 10 strains from eight outbreaks and two sporadic cases were amplified by reverse transcription-PCR (RT-PCR) and were used for this analysis.

Amplification and sequencing for the complete capsid gene of GII.2 strains. Viral RNA was extracted from stool suspensions by using a QIAamp viral RNA Mini kit (Qiagen, Hilden, Germany). RT-PCR was carried out with the reaction mixtures and enzymes as previously described (8). RT was performed at 42°C for 2 to 3 h with reverse primer, N235Rex (17), and enzyme was inactivated at 95°C for 5 min. PCR was performed using several pairs of PCR primers (Table 2) with a GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA) under the following conditions: denaturation at 95°C for 1 min; 40 cycles of 95°C for 15 s, 50°C for 30 s, and 72°C for 1 min; and a final cycle of incubation at 72°C for 5 min. When a PCR failed to produce strong products, we performed nested PCR. The amplified fragments were sequenced directly with a Big Dye terminator cycle sequencing kit and ABI 3700 sequencer (Applied Biosystems, Foster City, CA). The nucleotide sequences were determined in both orientations using the PCR primers. DNA sequences were edited using SeqManII (DNAStar Inc., Konstanz, Germany).

Sequence analysis. Nucleotide or amino acid sequence alignments were performed with BioEdit (version 7.052) (15), Clustal X (version 1.81) (35), or MUSCLE (version 3.51) (10). The extraction of the informative sites from nucleotide or amino acid sequence alignments was performed with ProSeq (version 2.91) (12). The rate of change for different domains was compared using chi-square statistics. In this analysis, a site was designated as an informative site when at least two strains had an identical amino acid in the alignment that differed from the other sequences. A phylogenetic tree with 1,000 bootstrap replications was constructed by the neighbor-joining method, and the genetic distances were calculated according to the Kimura two-parameter method (21). We performed additional phylogenetic analysis by the Bayesian method using MrBayes (version 3.1.2) (32). Location of specific domains of the GII.2 NoV capsid gene was done according to Chen et al. (7). For computational predictions of the structure of the GII.2 NoV capsid protein, we used the X-ray crystal structures of the capsid protein of Norwalk/68/US (GI.1 genotype, Protein Data Bank identifier [PDB ID] 1IHM, consisting of a complete trimer) (31) and VA387/98/US (GII.4 genotype, PDB ID 2OBR, consisting of only a monomeric P domain) (6) as templates to build homology models. The sequence alignments for the structure and the three-dimensional (3D) models for GII.2 NoV capsid proteins were made by using the WHAT IF program (38) and the 3D-Jigsaw (3)

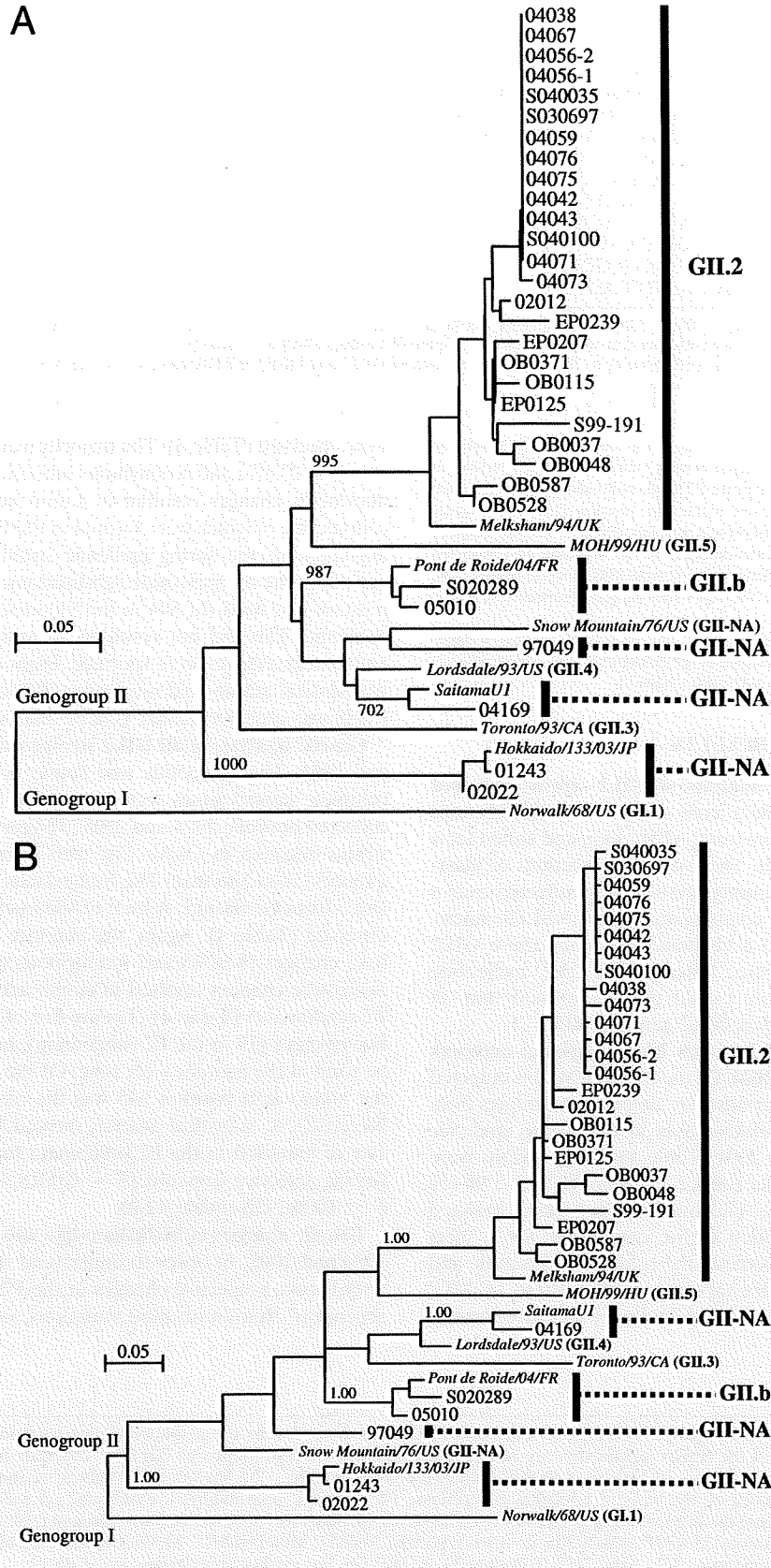

TABLE 2. Primers used to amplify the capsid gene of GII.2 NoV strains

Primer	Sequence (5′ to 3′)[a]	Polarity	Location (nt)	Reference or source
COG2F	CARGARBCNATGTTYAGRTGGATGAG	+	5003[b]	19
G2SKF	CNTGGGAGGGCGATCGCAA	+	5058[b]	22
G2SKR	CCRCCNGCATRHCCRTTRTACAT	−	5401[b]	22
N235Rex	GCWANRAAAGCTCCWGCCAT	−	6273[b]	17
MKcap508-524F	CAGAAAGATGATCCCAA	+	508[c]	This study
MKcap524-508R	TTGGGATCATCTTTCTG	−	524[c]	This study
MKcap662-649R	TCCACTGTTGGTGG	−	662[c]	This study
MKcap1109-1125F	TGGGTCAGATTCAAATT	+	1109[c]	This study
MKcap1125-1109R	AATTTGAATCTGACCCA	−	1125[c]	This study
MKcap1304-1290R	AAGAGCAGGCGCTCC	−	1304[c]	This study
MK35R	CAAAAGCTCCAGCCAT	−	1644[c]	This study

[a] Abbreviations for residues in boldface are as follows: B, not A; H, not G. In addition, N is A, C, G, or T; R is A or G; W is A or T; Y is C or T.
[b] Location of the 5′ end of the primer corresponding to the nucleotide position of Lordsdale/93/UK (X86557).
[c] Location of the 5′ end of the primer corresponding to the nucleotide position of ORF2 and ORF3 of Melksham/94/UK (X81879).

and EsyPred3D (25) servers. WHAT IF could not model residues 342, 344, and 345 based on the Norwalk/68/US capsid protein template (PDB ID 1IHM); as they are present in an inserted loop of the P2 subdomain relative to the template structure, but they are predicted to be in the same position as in the 3D model based on the GII.4 genotype capsid protein template (PDB ID 2OBR). A dimer of the GII.2 NoV capsid protein was modeled by superimposing two predicted monomers onto the trimeric template of the Norwalk/68/US capsid protein. The 3D models were visualized by the YASARA view program (version 6.813, http://www.yasara.org/) (24).

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been deposited in the DDBJ/EMBL/GenBank with the accession numbers AB281081 to AB281090 (Table 1).

RESULTS

Capsid gene sequence analysis of GII.2 strains collected over 30 years. Sequence data from a total of 36 GII.2 strains showed that the capsid gene was 1,629 nt long and coded for a protein of 542 amino acids. There were no deletions or insertions in the capsid gene among these strains collected over a 30-year period. Sequence comparison showed ≥83.6% nucleotide and ≥93.5% amino acid identities among these GII.2 strains. The phylogenetic tree based on the 1,629 nucleotide sequences of the complete capsid gene confirmed that all strains were characterized as GII.2 genotype (Fig. 2).

Genetic analysis of GII.2 strains from a regional outbreak in Japan in the spring of 2004. Of the 21 GII.2 strains detected in Osaka City, Japan, 14 strains were detected in samples from cases in outbreaks or sporadic cases in the spring epidemic between March and May 2004 (Table 1). These strains were closely related to each other (≥99.2% nucleotide and ≥ 99.1% amino acid identities). To find out if these viruses changed genetically during circulation in the community for a short period of time (about 3 months), the complete nucleotide and amino acid sequences of the capsid gene from these 14 GII.2 strains were compared. A total of 45 nucleotide substitutions were observed (Table 3). The majority were third-base position changes (77.8%) and synonymous substitutions (82.2%). These nucleotide changes resulted in amino acid changes in eight codons, five of which were located in the P2 subdomain. In the alignment of the spring epidemic strains over the 3-month period, of the 45 nucleotide substitutions, only one nucleotide position was fixed (nt 594 in the S domain) at the end of this epidemic. This did not result in an amino acid change. All amino acid changes were sporadic. Eight of the 14 strains had an identical amino acid sequences (Table 1). Of these, strain 04038 was used for further genetic analysis.

Genetic analysis of all GII.2 strains collected between 1976 and 2005. The nucleotide and amino acid sequences of the complete capsid genes were compared for 29 GII.2 strains collected between 1975 and 2005 (30 years), excluding 7 GII.2 strains detected in Osaka City with identical amino acid sequences (10 strains from The Netherlands, 14 from Osaka City, and 5 from GenBank). A total of 488 nucleotide changes were observed (Table 4). Again, the majority of these were third-base changes (83.8%) and synonymous (85.9%). In total, 59 nucleotide changes resulted in amino acid changes (34 in the P2 subdomain) (Table 4). Twenty-five of these were informative changes (19 in the P2 subdomain), but none appeared to be fixed in the genome over time. Of the 25 informative sites, the amino acid position 345 was the most variable (Fig. 3). Nevertheless, statistical analysis showed a significantly higher rate of mutation in the P2 subdomain than in P1 and S, suggesting selective pressure ($P = 0.0018$, chi-square 5.63; and $P < 0.0001$, chi-square 12.9).

Genetic analysis of Melksham-like strains detected between 1994 and 2005. In order to understand the apparent discrepancy between selective changes in the P2 subdomain and the absence of fixation of these mutations, we repeated our anal-

FIG. 1. Phylogenetic analysis of partial RdRp gene sequences of GII.2 NoV strains by the neighbor-joining method (A) and the Bayesian method (B). The genotypes at the RdRp regions that are not assigned numbers are represented as GII-NA. The scale indicates the number of substitutions per site. Reference strains of NoV used in this analysis are given in italics. (A) The bootstrap values are indicated on each branch. (B) Tree topology was evaluated on the basis of 1,500,000 generations. The posterior probabilities are indicated on each branch (≥0.95 of the posterior probability means that the branch has high credibility). The GenBank accession numbers of the additional reference strains used in this analysis are as follows: Hokkaido/133/03/JP, AB212306; Lordsdale/93/UK, X86557; MOH/99/HU, AF397256; Norwalk/68/US, M87661; Pont de Roide 673/04/FR, AY682549; Saitama U1, AB039775; Toronto/93/CA, U02030. For strain abbreviations, see Table 1.

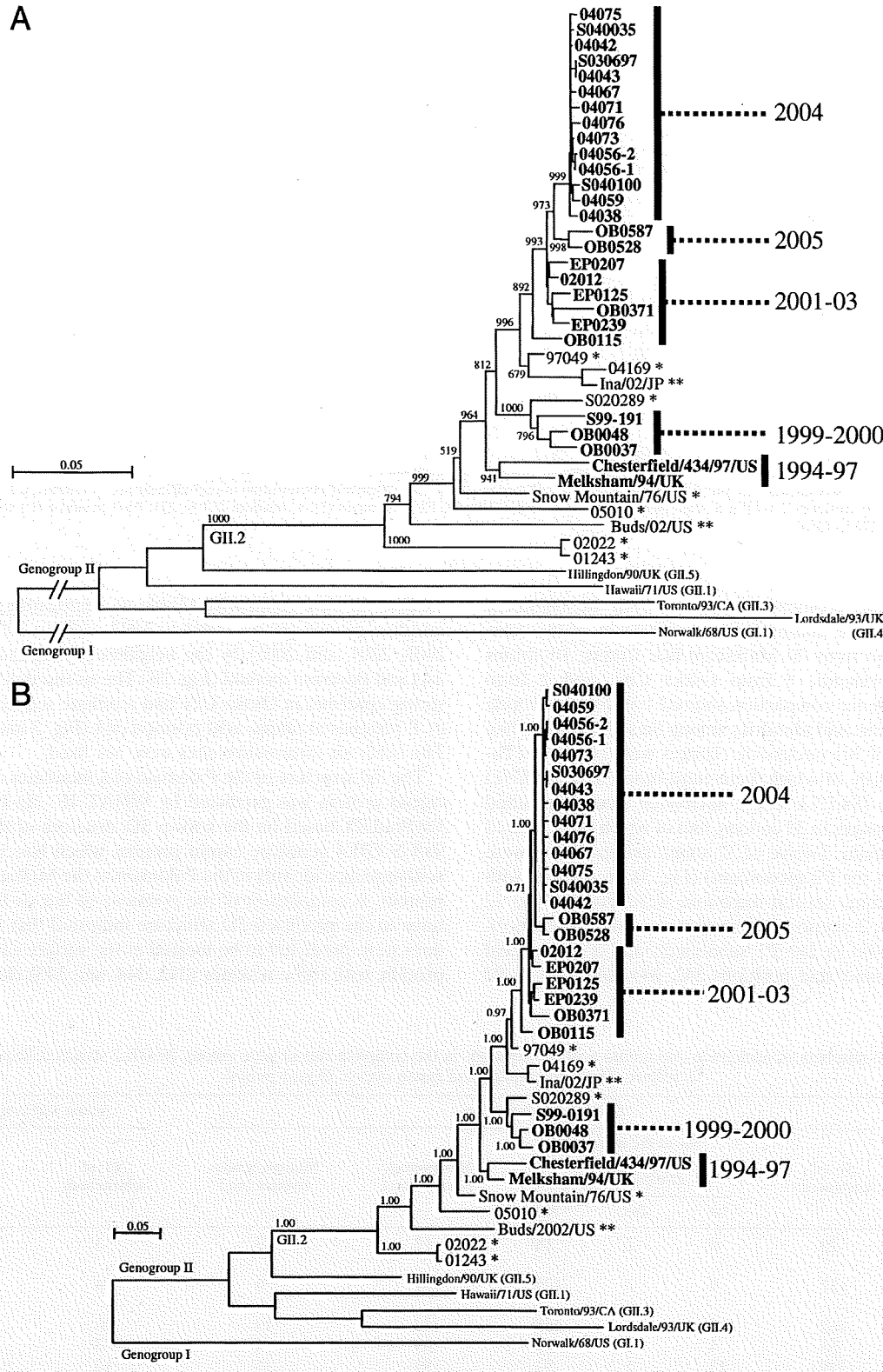

FIG. 2. Phylogenetic analysis of complete capsid gene sequences of GII.2 NoV strains by the neighbor-joining method (A) and the Bayesian method (B). Melksham-like strains, which have a matching (GII.2) RdRp sequence, are shown in boldface. Asterisks indicate the GII.2 strains that have different (*) or unknown (**) genotypes in the RdRp region. The scale indicates the number of substitutions per site. (A) The bootstrap values are indicated on each branch. (B) Tree topology was evaluated on the basis of 500,000 generations. The posterior probabilities are indicated on each branch. The GenBank accession numbers of the additional reference strains used in this analysis are as follows: Hawaii/71/US, U07611; Hillingdon/90/UK, AJ277607; Lordsdale/93/UK, X86557; Norwalk/68/US, M87661; Toronto/93/CA, U02030. For strain abbreviations, see Table 1.

TABLE 3. The numbers of nucleotide and amino acid substitutions in the complete capsid gene among 14 GII.2 strains detected in the spring epidemic of 2004 in Osaka City, Japan

Domain or subdomain[a]	Nucleotide substitutions						Amino acid substitutions	
	No. of substitutions	No. at the indicated position of the changed codon			Synonymous changes (no.)	Nonsynonymous changes (no.)	No. of substitutions	Informative changes (no.)
		1st	2nd	3rd				
N	1	0	0	1	1	0	0	0
S	18	3	2	13	16	2	2	0
P1	7	0	0	7	6	1	1	0
P1-1	4	0	0	4	4	0	0	0
P1-2	3	0	0	3	2	1	1	0
P2	19	3	2	14	14	5	5	0
Total	45	6	4	35	37	8	8	0

[a] Each domain of GII.2 NoV capsid gene was determined according to Chen et al. (7). The region and abbreviation of each domain are as follows: N, 5′ end of ORF2 and N terminal domain (residues 1 to 45); S, shell domain (residues 46 to 216); P1, P1 subdomain (P1-1, residues 217 to 274; P1-2, residues 421 to 542); P2, P2 subdomain (residues 275 to 420).

ysis after removing recombinant genomes from the alignment (Table 5) (nucleotide and amino acid sequences of the complete capsid gene from 20 Melksham-like strains: 10 strains from The Netherlands, 8 from Osaka City, and 2 from GenBank). Sequence comparison showed ≥91.5% nucleotide and ≥97.4% amino acid identities among these Melksham-like strains. A total of 301 nucleotide changes were observed (Table 5), the majority of which were third-base changes (87%) and synonymous (88.7%). These nucleotide changes resulted in amino acid changes in 32 codons, half of which were located in the P2 subdomain. Twelve of 32 amino acid positions were informative (9 in the P2 subdomain) (Fig. 3). In contrast with the previous finding, several mutations were fixed: of the 12 informative sites, 2 amino acid substitutions (amino acid positions 245 and 440) in the P1 subdomain and 3 amino acid substitutions (amino acid positions 342, 344, and 345) in P2 subdomain were cumulative (Table 6), segregating the strains into five genetic groups (1994 to 1997, 1999 to 2000, 2001 to 2003, 2004, and 2005) by the neighbor-joining method (Fig. 2A) and Bayesian method (Fig. 2B).The strains detected in the spring epidemic in Osaka City had a unique sequence, with S or P residues at amino acid position 364 (Fig. 3 and Table 6). The other six informative sites were not fixed.

The 3D structure of the P domain of a monomer of the NoV capsid protein was predicted by WHAT IF, 3D-Jigsaw, and EsyPred3D, based on the known 3D structure of the VA387/98/US GII.4 genotype capsid protein, which has 55% amino acid sequence identity in the P domain to the Melksham capsid protein. A comparison of the positions of the six fixed mutations to the predicted 3D structure indicated that all six residues were predicted to be located at the surface of the capsid protein, with three residues (342, 344, and 345) close to each

TABLE 4. The numbers of nucleotide and amino acid substitutions in the complete capsid gene among 29 GII.2 strains collected in the GenBank, The Netherlands, and Japan over a 30-year period

Domain or subdomain[a]	Nucleotide substitutions						Amino acid substitutions	
	No. of substitutions	No. at the indicated position of the changed codon			Synonymous changes (no.)	Nonsynonymous changes (no.)	No. of substitutions	Informative changes (no.)
		1st	2nd	3rd				
N	20	0	2	18	17	3	3	1
S	137	9	3	125	131	6	6	1
P1	162	17	7	138	143	19	16	4
P1-1	53	4	2	47	49	4	3	2
P1-2	109	13	5	91	94	15	13	2
P2	169	28	13	128	128	41	34	19
Total	488	54	25	409	419	69	59	25

[a] See note to Table 3.

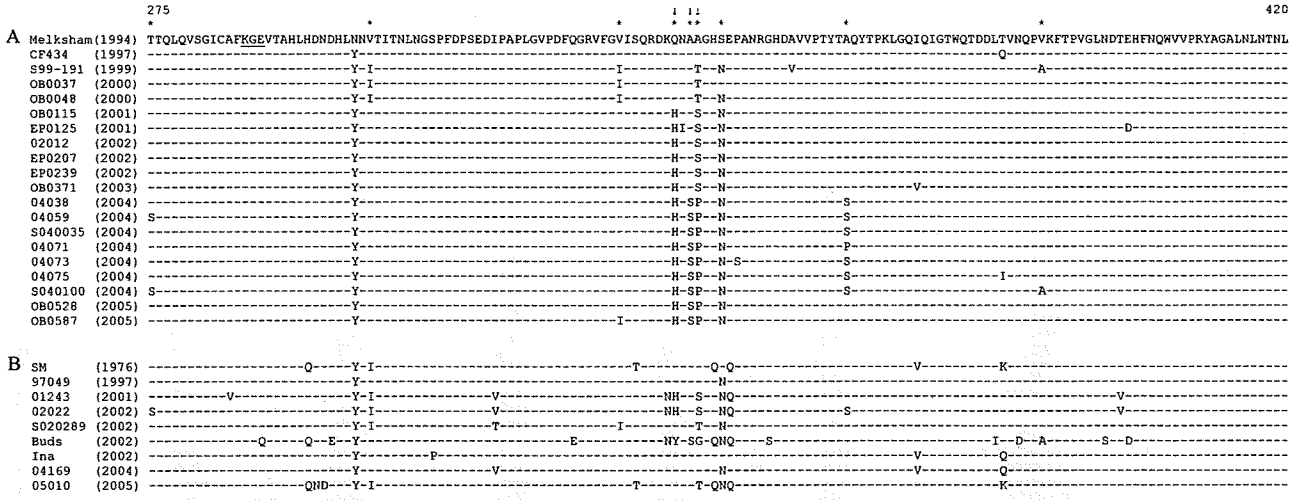

FIG. 3. Amino acid alignment of the P2 subdomains of GII.2 NoV strains showing mutations along the aligned sequences. The upper sequence alignment group (A) includes the Melksham-like strains and the lower group (B)includes the other GII.2 strains, which were recombinant genomes. In each group, sequences are arranged chronologically from top (oldest) to bottom (most recent). The detection years of the strains are indicated in parentheses. The asterisk indicates informative sites among Melksham-like strains. The arrow denotes cumulatively changing amino acid positions (342, 344, and 345) among Melksham-like strains. The numbers above the sequences indicate the sequence position relative to the position in the capsid protein of strain Melksham. Underlined sequences indicate the KGE motif that corresponds to the RGD-like motif of other NoVs and was determined with the amino acid sequence alignment of other NoVs according to Tan et al. (34). For strain abbreviations, see Table 1.

other in the P2 subdomain (Fig. 4). Furthermore, residues 342, 344, 345 of the P2 subdomain and residues 245 and 440 of the P1 subdomain were grouped closely together on the predicted 3D structure of a dimer (3D-modeling by WHAT IF) (Fig. 5).

DISCUSSION

In this study, we analyzed the complete capsid gene of GII.2 NoVs collected over a 30-year period. The collection included a relatively large sample from an epidemic that was observed only in Japan in the spring period of 2004. In this 3-month period, no evolutionary changes were observed, but compared with other GII.2 Melksham-like strains these variants had unique amino acid sequences (S or P) at position 364. One of the recombinant genomes, strain 02022, had the same amino acid residues, suggesting that the presence of this mutation was not causally related to the epidemic pattern. The sequence analysis confirmed that the spring epidemic was an outbreak.

Molecular characterization of the GII.2 capsids over the full study period showed an interesting difference between recombinant genomes and nonrecombinant (Melksham-like) genomes. The Melksham-like strains clearly evolved over time, with accumulation of mutations particularly in the P2 subdomain. In contrast, GII.2 capsids from recombinant genomes did not fit this pattern and had a seemingly erratic pattern of mutations. Melksham-like strains are occasionally observed in molecular surveillance data from The Netherlands and Osaka

TABLE 5. The numbers of nucleotide and amino acid substitutions in the complete capsid gene among 20 Melksham-like strains over a 12-year period

Domain or subdomain[a]	Nucleotide substitutions						Amino acid substitutions	
	No. of substitutions	No. at the indicated position of the changed codon			Synonymous changes (no.)	Nonsynonymous changes (no.)	No. of substitutions	Informative changes (no.)
		1st	2nd	3rd				
N	10	0	2	8	8	2	2	0
S	81	4	3	74	76	5	6	1
P1	98	8	3	87	89	9	8	2
P1-1	34	2	1	31	32	2	2	1
P1-2	64	6	2	56	57	7	6	1
P2	112	13	6	93	94	18	16	9
Total	301	25	14	262	267	34	32	12

[a] See note to Table 3.

TABLE 6. Amino acid substitutions at six positions in the P domain among Melksham-like strains

Strain description			Amino acid and codon at the indicated position of the capsid protein[a]											
			245		342		344		345		364		440	
Name[b]	Country[c]	Detection date (mo/day/yr)	Residue	Codon	Residue	Codon	Residue	Codon	Residue	Codon	Residue	Codon	Residue	Codon
Melksham	UK	1994	P	CCC	Q	CAG	A	GCC	A	GCT	A	GCA	L	CTC
CF434	US	1997	—	CCC	—	CAG	—	GCT	—	GCT	—	GCA	—	CTT
S99-191	NL	1/1999	S	TCT	—	CAA	—	GCC	T	ACT	—	GCA	I	ATC
OB0037	NL	9/2000	S	TCT	—	CAA	—	GCC	T	ACT	—	GCA	I	ATC
OB0048	NL	12/2000	S	TCT	—	CAA	—	GCC	T	ACT	—	GCA	I	ATC
OB0115	NL	2/2001	S	TCT	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
EP0125	NL	11/2001	S	TCC	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
02012	JP	1/2002	S	TCC	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
EP0207	NL	1/2002	S	TCC	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
EP0239	NL	2/2002	S	TCC	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
OB0371	NL	10/2003	S	TCC	H	CAC	—	GCC	S	TCT	—	GCA	I	ATC
04038	JP	3/11/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
04059	JP	4/18–30/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
S040035	JP	4/21/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
04071	JP	5/16–23/2004	S	TCC	H	CAC	S	TCC	P	CCT	P	CCA	I	ATC
04073	JP	5/22/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
04075	JP	5/25/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
S040100	JP	5/31/2004	S	TCC	H	CAC	S	TCC	P	CCT	S	TCA	I	ATC
OB0528	NL	1/2005	S	TCC	H	CAC	S	AGC	P	CCT	—	GCA	I	ATT
OB0587	NL	4/2005	S	TCC	H	CAC	S	AGC	P	CCT	—	GCA	I	ATT

[a] The position number corresponds to the capsid sequence of the Melksham/94/UK strain; a dash indicates sequence identity with this strain. Residues are indicated by their single-letter codes.
[b] For strain abbreviations, see Table 1.
[c] UK, United Kingdom; US, United States; NL, The Netherlands; JP, Japan.

City, and strains from Japan and The Netherlands could be consistently grouped in the same alignment, suggesting continuous widespread circulation and an ability to cause disease. The recombinant genomes with GII.2 capsids, however, occur sporadically and did not show evidence for evolution, suggesting that their circulation is limited (1, 4, 17). The finding that strains from such widely separated geographic regions were similar shows that evolution of GII.2 NoVs is a global phe-

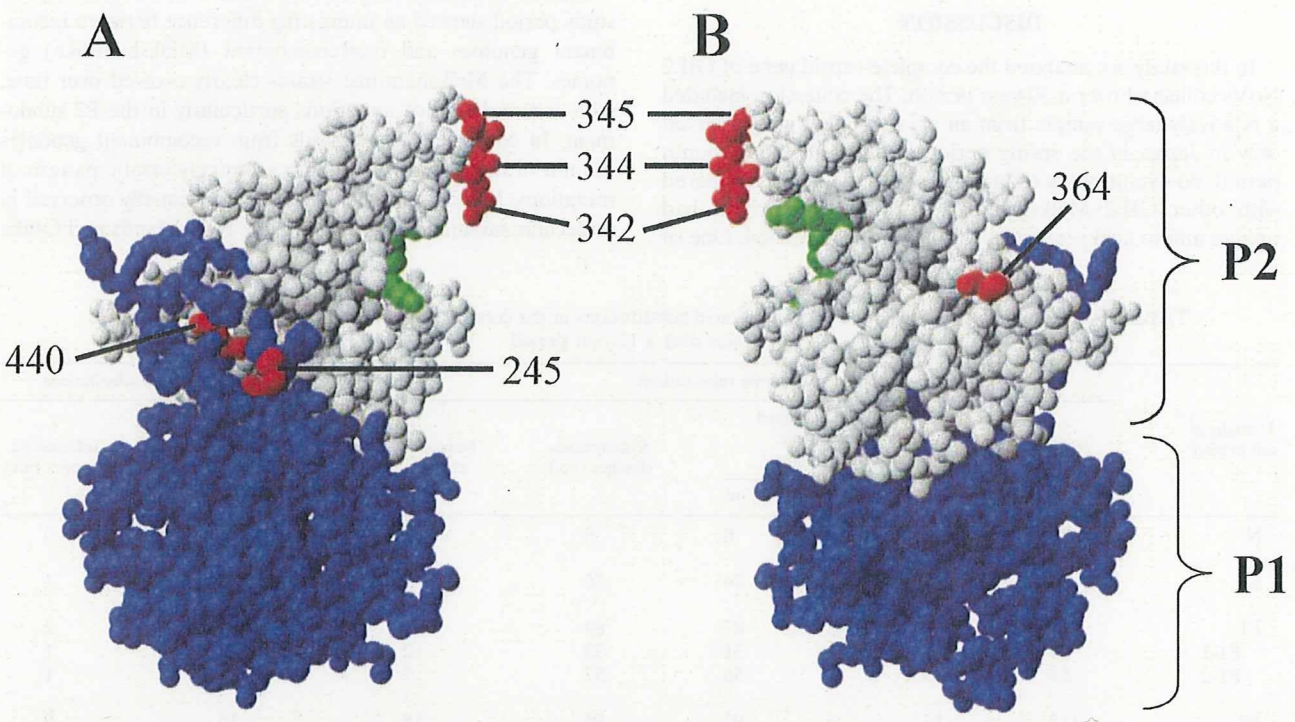

FIG. 4. Location of six fixed amino acid residues (positions 245, 342, 344, 345, 364, and 440, shown in red) on the monomer of the capsid protein. This 3D structure for the monomer P domain of the GII.2 NoV capsid protein was made by WHAT IF. The P1 and P2 subdomains are shown in blue and gray, respectively. The S domain is not shown. The KGE motif in the P2 domain is shown in green. (A) Predicted location of amino acid residues 342, 344, 345, and 364 in the P2 subdomain. (B) A view identical to panel A rotated 180° horizontally showing the location of amino acid residues 245 and 440 in the P1 subdomain.

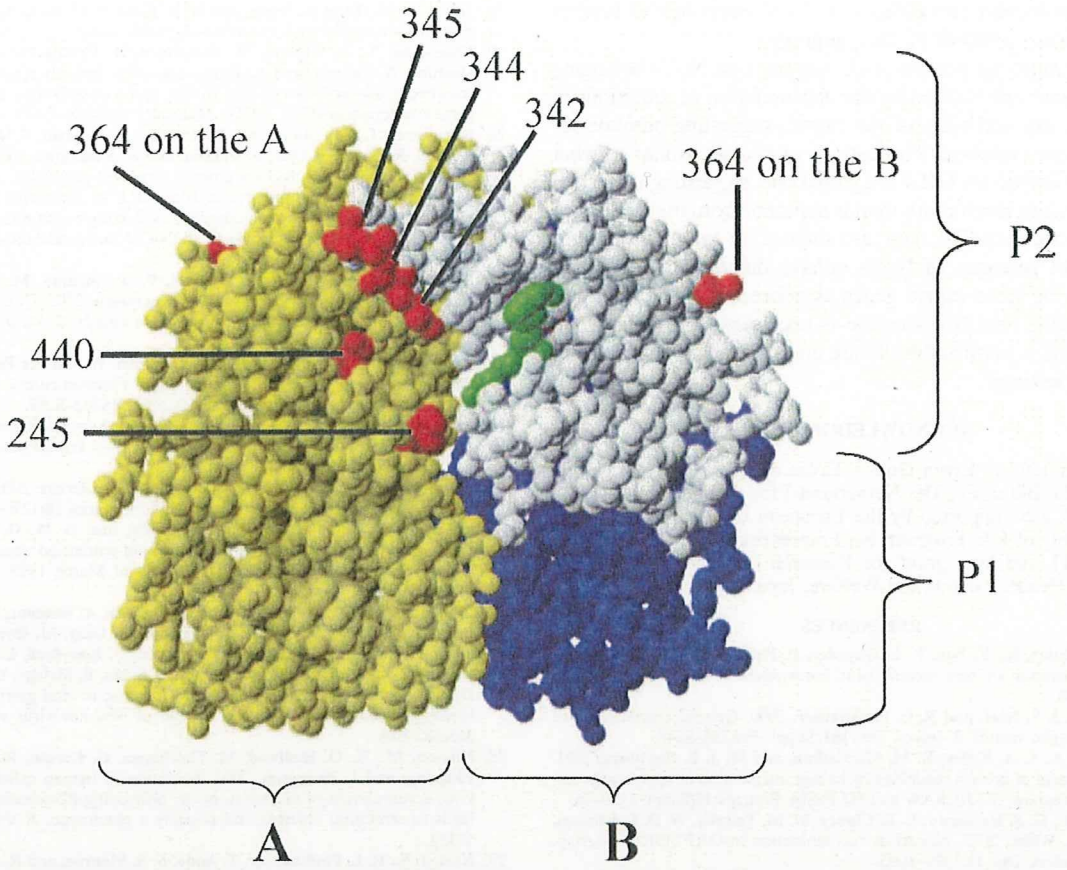

FIG. 5. The 3D structure of a dimer of the GII.2 NoV capsid protein, as predicted by WHAT IF, showing the location of the six informative amino acid residues (red). Monomer A is shown in yellow, while the monomer B is shown in blue (P1) and gray (P2). The S domains are not shown. The KGE motif (green) in the P2 subdomain on monomer B is also shown.

nomenon, similar to what has been described for GII.4 NoVs (5, 27, 33).

The molecular characterization of Melksham-like strains showed accumulation of mutations over a 12-year period, segregating them into five genetic groups. Three of these were supported in phylogenetic analysis with high bootstrap values. Two others (Fig. 2, 1999 to 2000 and 2001 to 2003) were not distinguishable in phylogenetic analysis, possibly because of the small number of isolates in the comparison. The present collection, however, contains all GII.2 strains detected in the two countries in the past 12 years of surveillance. Amino acid position 345 was the most variable in all GII.2 strains. This is suggestive of immune pressure, particularly because the amino acid is predicted to be located on a surface-exposed part of the capsid. The fixed amino acid changes were topologically clustered, judging from the 3D structure prediction, indicating that combined they may form an epitope. Immunity to NoVs has not been studied very systematically, but recent work on GII.4 NoVs has suggested that these viruses evolved under immune pressure (9, 30, 33). The lack of a tissue culture model makes it difficult to corroborate these findings by using cross-neutralization tests with hyperimmune serum directed against specific variants. Evolution of strains in an immunocompromised patient with chronic shedding of NoVs has been demonstrated by Nilsson et al. (28), who suggested that the cumulative amino acid substitutions appearing in the P2 subdomain were immune response driven. To clarify the relations between these amino acid mutations and receptor or antibody binding further immunological studies are needed.

Recently, new variants of the predominant NoV, the GII.4 genotype, appeared in 2002 (27), 2004 to 2005 (5), and 2006 (33) with global distribution. These new GII.4 variants were characterized into different clusters by the phylogenetic analysis based on the capsid gene (5, 9, 30). Okada et al. (30) reported that GII.4 strains had cumulative amino acid changes in the P domain and 5% to 18% amino acid diversity in the P2 subdomain between GII.4 subtypes for 5 years study. Bull et al. (5) also described more than 5% amino acid diversity in the complete capsid gene between past GII.4 strains and variants. Siebenga et al. (33) found 9% fixed amino acid mutations across the capsid gene over a 12-year period, but only five of these consistently delineated subsequent epidemic strains. In contrast, here we found only 2.6% amino acid diversity across the complete capsid gene of Melksham-like strains in a 12-year period, showing a clearly lower rate of change. The difference between these genotypes is intriguing and suggests clear differences in epidemiology that are not easy to explain. This study has been done using a limited number of strains in a rare genotype from a limited number of locations. It is necessary to

continue molecular surveillance of NoV infections to further the evolutional analysis of each genotype.

In conclusion, the present study showed that NoVs belonging to a rare genotype evolved by the accumulation of mutations in the surface-exposed parts of the capsid, suggesting immune response-driven evolution. The pattern of change similar to what has been observed for GII.4 is a global one, suggesting that these viruses circulate much more than is apparent from their presence in surveillance data. The data also showed for the first time that recombinant genomes of NoVs behave differently, even when they share the same capsid genes as nonrecombinant genomes. Understanding how NoVs evolve is necessary for finding more effective ways to control this disease and particularly its impact in healthcare settings.

ACKNOWLEDGMENTS

We are grateful to Erwin Duizer, Erwin de Bruin, and Bas van de Veer (RIVM, Bilthoven, The Netherlands) for supporting our work.

This work was supported by the European Commission, DG Research Quality of Life Program, 6th Framework (EVENT, SP22-CT-2004-502571), and by a grant for Research on Food Safety of the Ministry of Health, Labour and Welfare, Japan.

REFERENCES

1. **Ambert-Balay, K., F. Bon, F. L. Guyader, P. Pothier, and E. Kohli.** 2005. Characterization of new recombinant noroviruses. J. Clin. Microbiol. **43:** 5179–5186.
2. **Ando, T., J. S. Noel, and R. L. Fankhauser.** 2000. Genetic classification of "Norwalk-like viruses. J. Infect. Dis. **181**(Suppl. 2):S336–S348.
3. **Bates, P. A., L. A. Kelley, R. M. MacCallum, and M. J. E. Sternberg.** 2001. Enhancement of protein modelling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins **45**(Suppl. 5):39–46.
4. **Bull, R. A., G. S. Hansman, L. E. Clancy, M. M. Tanaka, W. D. Rawlinson, and P. A. White.** 2005. Norovirus recombination in ORF1/ORF2 overlap. Emerg. Infect. Dis. **11**:1079–1085.
5. **Bull, R. A., T. V. Elise, C. J. McIver, W. D. Rawlinson, and P. A. White.** 2006. Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. J. Clin. Microbiol. **44**:327–333.
6. **Cao, S., Z. Lou, M. Tan, Y. Chen, Y. Liu, Z. Zhang, X. C. Zhang, X. Jiang, X. Li, and Z. Rao.** 2007. Structural basis for the recognition of blood group trisaccharides by norovirus. J. Virol. **81**:5949–5957.
7. **Chen, R., J. D. Neil, J. S. Noel, A. M. Hutson, R. I. Glass, and B. V. Prasad.** 2004. Inter- and intragenus structural variations in caliciviruses and their functional implications. J. Virol. **78**:6469–6479.
8. **de Bruin, E., E. Duizer, H. Vennema, and M. P. G. Koopmans.** 2006. Diagnosis of norovirus outbreaks by commercial ELISA or RT-PCR. J. Virol. Methods **137**:259–264.
9. **Dingle, K. E., and Norovirus Infection Control in Oxfordshire Communities Hospitals.** 2004. Mutation in a Lordsdale norovirus epidemic strain as a potential indicator of transmission routes. J. Clin. Microbiol. **42**:3950–3957.
10. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.
11. **Fankhauser, R. L., J. S. Noel, S. S. Monroe, T. Ando, and R. I. Glass.** 1998. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United states. J. Infect. Dis. **178**:1571–1578.
12. **Filatov, D. A.** 2002. ProSeq: a software for preparation and evolutionary analysis of DNA sequence data sets. Mol. Ecol. Notes **2**:621–624.
13. **Green, K. Y.** 2007. *Caliciviridae*: The noroviruses, p. 949–979. *In* B. N. Fields, D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields virology, 5th ed. Lippincott, Williams & Wilkins, Philadelphia, PA.
14. **Green, S. M., K. E. Dingle, P. R. Lambden, E. O. Caul, C. R. Ashley, and I. N. Clarke.** 1994. Human enteric *Caliciviridae*: a new prevalent small round-structured virus group defined by RNA-dependent RNA polymerase and capsid diversity. J. Gen. Virol. **75**:1883–1888.
15. **Hall, T. A.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. **41**:95–98.
16. **Hardy, M. E., S. F. Kramer, J. J. Treanor, and M. K. Estes.** 1997. Human calicivirus genogroup II capsid sequence diversity revealed by analyses of the prototype Snow Mountain agent. Arch. Virol. **142**:1469–1479.
17. **Iritani, N., A. Kaida, H. Kubo, N. Abe, T. Murakami, H. Venemma, M. Koopmans, N. Takeda, H. Ogura, and Y. Seto.** 21 May 2008. An epidemic of GII.2 genotype noroviruses in the spring of 2004 in Osaka City, Japan. J. Clin. Microbiol. doi:10.1128/JCM.01993-07.
18. **Jiang, X., M. Wang, K. Wang, and M. K. Estes.** 1993. Sequence and genomic organization of Norwalk virus. Virology **195**:51–61.
19. **Kageyama, T., S. Kojima, M. Shinohara, K. Uchida, S. Fukushi, F. B. Hoshino, N. Takeda, and K. Katayama.** 2003. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. J. Clin. Microbiol. **41**:1548–1557.
20. **Kageyama, T., M. Shinohara, K. Uchida, S. Fukushi, F. B. Hoshino, S. Kojima, R. Takai, T. Oka, N. Takeda, and K. Katayama.** 2004. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to norovirus in Japan. J. Clin. Microbiol. **42**:2988–2995.
21. **Kimura, M.** 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.
22. **Kojima, S., T. Kageyama, S. Fukushi, F. B. Hoshino, M. Shinohara, K. Uchida, K. Natori, N. Takeda, and K. Katayama.** 2002. Genogroup-specific PCR primers for detection of Norwalk-like viruses. J. Virol. Methods **100:** 107–114.
23. **Koopmans, M., J. Vinje, M. de Wit, I. Leenen, W. van der Poel, and Y. van Duynhoven.** 2000. Molecular epidemiology of human enteric caliciviruses in The Netherlands. J. Infect. Dis. **181**(Suppl. 2):S262–S269.
24. **Krieger, E., G. Koraimann, and G. Vriend.** 2002. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. Proteins **47**:393–402.
25. **Lambert, C., N. Leonard, X. De Bolle, and E. Depiereux.** 2002. ESyPred3D: prediction of proteins 3D structures. Bioinformatics **18**:1250–1256.
26. **Lewis, D. C., A. Hale, X. Jiang, R. Eglin, and D. W. G. Brown.** 1997. Epidemiology of Mexico virus, a small round-structured virus in Yorkshire, United Kingdom, between January 1992 and March 1995. J. Infect. Dis. **175**:951–954.
27. **Lopman, B., H. Vennema, E. Kohli, P. Pothier, A. Sanchez, A. Negredo, J. Buesa, E. Schreier, M. Reacher, D. Brown, J. Gray, M. Iturriza, C. Gallimore, B. Bottiger, K. Hedlund, M. Torven, C. Bonsdorff, L. Maunula, M. Poljsak-Prijatelj, J. Zimsek, G. Reuter, G. Szucs, B. Melegh, L. Svennson, Y. Duijnhoven, and M. Koopmans.** 2004. Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet **363**:682–688.
28. **Nilsson, M., K. O. Hedlund, M. Thorhagen, G. Larson, K. Johansen, A. Ekspong, and L. Svensson.** 2003. Evolution of human calicivirus RNA in vivo: accumulation of mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a phenotype. J. Virol. **77**:13117–13124.
29. **Noel, J. S., R. L. Fankhauser, T. Ando, S. S. Monroe, and R. I. Glass.** 1999. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J. Infect. Dis. **179**:1334–1344.
30. **Okada, M., T. Ogawa, I. Kaiho, and K. Shinozaki.** 2005. Genetic analysis of noroviruses in Chiba Prefecture, Japan, between 1999 and 2004. J. Clin. Microbiol. **43**:4391–4401.
31. **Prasad, B. V., M. E. Hardy, T. Dokland, J. Bella, M. G. Rossmann, and M. K. Estes.** 1999. X-ray crystallographic structure of the Norwalk virus capsid. Science **286**:287–290.
32. **Ronquist, F., and J. P. Huelsenbeck.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.
33. **Siebenga, J. J., H. Vennema, B. Renckens, E. de Bruin, B. van der Veer, R. J. Siezen, and M. Koopmans.** 2007. Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J. Virol. **81**:9932–9941.
34. **Tan, M., P. Huang, J. Meller, W. Zhong, T. Farkas, and X. Jiang.** 2003. Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. J. Virol. **77:** 12562–12571.
35. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.
36. **Vinjé, J., J. Green, D. C. Lewis, C. I. Gallimore, D. W. Brown, and M. P. Koopmans.** 2000. Genetic polymorphism across regions of the three open reading frames of "Norwalk-like viruses." Arch. Virol. **145**:223–241.
37. **Vinjé, J., R. A. Hamidjaja, and M. D. Sobsey.** 2004. Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses. J. Virol. Methods **116**:109–117.
38. **Vriend, G.** 1990. WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. **8**:52–56.
39. **Wang, J., X. Jiang, H. P. Madore, J. Gray, U. Desselberger, T. Ando, Y. Seto, I. Oishi, J. F. Lew, K. Y. Green, et al.** 1994. Sequence diversity of small, round-structured viruses in the Norwalk virus group. J. Virol. **68**:5982–5990.
40. **Widdowson, M. A., E. H. Cramer, L. Hadley, J. S. Bresse, R. S. Beard, S. N. Bulens, M. Charles, W. Chege, E. Isakbaeva, J. G. Wright, E. Mintz, D. Forney, J. Massey, R. I. Glass, and S. S. Monroe.** 2004. Outbreaks of acute gastroenteritis on cruise ships and on land: identification of a predominant circulating strain of norovirus—United States, 2002. J. Infect. Dis. **190**:27–36.
41. **Xi, J. N., D. Y. Graham, K. N. Wang, and M. K. Estes.** 1990. Norwalk virus genome cloning and characterization. Science **250**:1580–1583.

Arch Virol (2010) 155:111–115
DOI 10.1007/s00705-009-0528-0

BRIEF REPORT

Transition of genotypes associated with norovirus gastroenteritis outbreaks in a limited area of Japan, Hiroshima Prefecture, during eight epidemic seasons

Shinji Fukuda · Shinichi Takao · Naoki Shigemoto · Yukie Tanizawa · Masato Seno

Received: 4 June 2009 / Accepted: 28 September 2009 / Published online: 1 December 2009

Abstract The transition of genotypes implicated in 102 NoV gastroenteritis outbreaks in Hiroshima Prefecture, Japan, during eight epidemic seasons was investigated. Eighteen genotypes were implicated in the outbreaks, with the chronological characteristics as in GII.3, GII.4, GII.5 and GII.12. In GII.4 variants, amino acid changes and positively selected sites were of note and significantly concentrated in the surface-exposed P2 subdomain of the VP1 protein. Notably, variant-specific epitopes at which positively selected sites are located may be significant for distinguishing a new GII.4 variant. The interaction of these genetic changes with developing immunity seems to influence NoV epidemics.

Noroviruses (NoVs), which belong to the family *Caliciviridae*, are classified into five genogroups (GI to GV), of which GI, II, and IV cause acute gastroenteritis in humans [26], and further divided into more than 30 genotypes [12]. In particular, GII.4 has emerged in recent years. Although genetic variants of GII.4 such as <1996, 1996, and 2002 had been identified in the past [25], in 2004 and 2005, a new GII.4 variant, termed Hunter virus, became prevalent [4]. Then, in 2006 and 2007, the variants 2006a and 2006b became epidemic [3, 11, 16, 22]. In Japan, there was more than a fourfold increase in NoV infections compared to the previous year during the 2006/2007 epidemic season, and the increase was attributed to the emergence of the 2006b variant, with a small role played by the 2006a variant [16]. Before 2004, genotypes such as GII.3 and GII.5 as well as GII.4 were associated with outbreaks of gastroenteritis in Japan [19, 24]. Here, we describe the chronological changes of NoV genotypes and the genetic characterization of GII.4 implicated in outbreaks of gastroenteritis in Hiroshima Prefecture, Japan from the 2000/2001 to the 2007/2008 epidemic season.

A total of 121 strains detected from patients of 102 NoV gastroenteritis outbreaks in Hiroshima Prefecture, Japan, between the 2000/2001 and 2007/2008 epidemic seasons were used. Of the 121 strains, 9, 14, 8, 10, 19, 6, 38, and 17 were obtained in the 2000/2001, 2001/2002, 2002/2003, 2003/2004, 2004/2005, 2005/2006, 2006/2007, and 2007/2008 epidemic seasons, respectively.

For phylogenetic analysis, PCR amplicons of 330 and 344 bp for GI and GII, respectively, targeting from the C terminus of the RNA-dependent RNA polymerase gene to the N-terminal/Shell domain region [13, 17] were sequenced directly as described previously [8]. Genotyping was carried out according to the scheme described by Kageyama et al. [12], and the GII.4 variants were typed using the 5′ end of capsid sequences according to the cluster of Motomura et al. [16].

Furthermore, full-length amplicons (1,623 bp) of the VP1 region of 25 GII.4 strains, which were randomly selected in each epidemic season, were made using the primers G2SKF [13] and LVCAPEND [19]. Nucleotide sequencing was conducted after cloning into pDrive cloning vectors using a PCR cloning kit (Qiagen, Valencia, CA) [8]. Three clones were independently analyzed for sequence determination. The following primers were used: M13 FWD(-29) (M&S TechnoSystems, Osaka, Japan), LV5375 (5′-TGGYTATGC AGGTGGYTTTG-3′), LV5862 (5′-CAACCACAAAATG GCAGRTG-3′), and LV6362 (5′-CSGGTGARCARCTTC TTTTCT-3′) for forward sequencing and the anti-sense

S. Fukuda (✉) · S. Takao · N. Shigemoto · Y. Tanizawa · M. Seno
Center for Public Health and Environment,
Hiroshima Prefectural Technology Research Institute,
Minami-machi 1-6-29, Minami-ku, Hiroshima 734-0007, Japan
e-mail: s-fukuda80723@pref.hiroshima.lg.jp

versions of LV5375, LV6862, and LV6862 and M13 reverse (M&S TechnoSystems) for reverse sequencing. The amino acid sequences deduced from the nucleotide sequences using BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) were aligned with the ClustalW program (http://clustalw.ddbj.nig.ac.jp/top-j.html) for comparison of the 25 GII.4 strains and GII.4 VA387 (GenBank accession number AY038600) as a reference. Analysis of codons, such as the estimation of non-synonymous/synonymous substitution rates (dN/dS), was carried out with codeml in the PAML 3.14 package (http://abacus.gene.ucl.ac.uk/software/pamlOLD.html).

Of the 102 NoV gastroenteritis outbreaks, 68 (66.7%) and 34 (33.3%) were caused by person-to-person contact and food-borne transmission, respectively. As shown in Table 1, 18 genotypes were implicated in the outbreaks. Eight genotypes (GI.4, GII.3, GII.4, GII.6, GII.8, GII.12, GII.13, and GII.untypeable) and 15 genotypes (GI.1, GI.2, GI.4, GI.7, GI.8, GI.11, GI.14, GII.1, GII.2, GII.3, GII.4, GII.5, GII,6, GII.12, and GII.14) were identified in person-to-person and food-borne outbreaks, respectively. Between the 2000/2001 and 2002/2003 epidemic seasons, GI.2, GII.5, GII.12, and GII.14 were notable genotypes when at least two epidemic seasons had an identical genotype. Of these, GII.5 and GII.12 were characterized as unique until the 2002/2003 epidemic season [24]. Similarly, between the 2002/2003 and 2004/2005 epidemic seasons, GI.1 and GI.4, which were detected in environmental samples such as bivalves in Japan [17, 23], were notable genotypes. Since the 2004/2005 epidemic season, the major causative genotype has shifted to GII.4 in Hiroshima Prefecture, Japan. GII.3 was associated with outbreaks in the 2000/2001, 2003/2004, 2005/2006, and 2007/2008 epidemic seasons but not detected in two or more consecutive

Table 1 Genotypes implicated in NoV gastroenteritis outbreaks in each epidemic season

Epidemic season	Transmission			
	Person-to-person		Food-borne	
	Genogroup I	Genogroup II	Genogroup I	Genogroup II
2000/2001		GII.4-95/96 (1)	GI.4 (1)	GII.1 (1)
		GII.12 (1)	GI.8 (1)	GII.3 (3)
				GII.14 (1)
2001/2002		GII.4-95/96 (1)	GI.2 (2)	GII.5 (4)
		GII.12 (1)	GI.14 (1)	GII.12 (4)
				GII.14 (1)
2002/2003		GII.8 (1)	GI.2 (2)	GII.5 (2)
			GI.4 (1)	GII.14 (1)
			GI.7 (1)	
2003/2004		GII.4-04/05/JP/CHN (1)	GI.1 (2)	GII.3 (3)
			GI.4 (1)	GII.4-04/05/JP/CHN (1)
			GI.11 (1)	
			GI.14 (1)	
2004/2005	GI.4 (1)	GII.4-04/05/JP/CHN (6)	GI.1 (1)	GII.2 (1)
		GII.4-04/05/AU/NL (2)	GI.8 (1)	GII.4-04/05/JP/CHN (3)
		GII.6 (2)		GII.6 (1)
				GII.12 (1)
2005/2006		GII.3 (1)		GII.4-04/05/JP/CHN (1)
		GII.4-04/05/JP/CHN (3)		
		GII.untypeable (1)		
2006/2007		GII.4-2006b (28)		GII.4-2006b (9)
		GII.13 (1)		
2007/2008	GI.4 (1)	GII.3 (2)	GI.4 (1)	GII.4-2006b (1)
		GII.4-2006b (11)	GI.8 (1)	

The genotyping of NoVs and clustering of GII.4 variants were carried out according to the scheme of Kageyama et al. [12] and Motomura et al. [16], respectively

The GII.4 variant was expressed as a "genotype-variant"

Values in parentheses denote the number of genotypes implicated in the NoV gastroenteritis outbreaks

epidemic seasons. The chronological transition of such genotypes may be caused by the development of immunity in the population, by changing amino acid residues in the VP1 region [10], or by the sustained spread of GII.4 in recent years [3, 11, 16, 22], but there is no good evidence as to why it occurred.

In GII.4, four genetic variants (95/96, 04/05/JP/CHN, 04/05/AU/NL, and 2006b) were identified as the causative agents of gastroenteritis outbreaks involving both transmission modes (Table 1). The 95/96 and 04/05/JP/CHN variants became prevalent in two epidemic seasons (2000/2001 and 2001/2002) and three epidemic seasons (2003/2004, 2004/2005, and 2005/2006), respectively. The 04/05/AU/NL variant was detected only in the 2004/2005 epidemic season. The 04/05/JP/CHN variant was dominant in Hiroshima Prefecture, Japan, compared to the 04/05/AU/NL variant. However, there has been a sudden change to the 2006b variant since the 2006/2007 epidemic season, and the 2006b variant has emerged, with a share of 89.1% (49/55, Table 1), as a leading agent worldwide [3, 11, 16, 22]. The emergence of GII.4 in recent years, in conjunction with frequent asymptomatic infections and a high viral load [2, 9, 18, 20], may be attributed to antigenic drift due to amino acid changes. The amino acid changes in VP1 sequences among the 25 GII.4 strains were found at 47 positions (Fig. 1). Unique changes specific to the 2006b variants were found at 17 positions (residues A_{15}, T_{45}, E_{193}, G_{255}, L_{306}, L_{333}, Y_{352}, A_{356}, P_{357}, E_{372}, H_{378}, G_{393}, S_{407}, N_{412}, P_{414}, M_{475}, and T_{534}) when at least two strains had an identical amino acid sequence. The residues A_{15}, E_{193}, S_{407}, and T_{534}, in addition to L_{306}, Y_{352}, A_{356}, P_{357}, E_{372}, H_{378}, and N_{412} described previously [16], were observed in all 2006b variants analyzed. Twenty-four sites at positions 9, 45, 255, 280, 296, 297, 298, 300, 333, 339, 340, 352, 357, 368, 372, 378, 382, 393, 395, 407, 412, 413, 494, and 539 displayed values of dN/dS ≥ 1 and were judged to be positively selected sites (Fig. 2). Nineteen of 24 sites were located at the Protruding (P) 2 subdomain. The residues at positions 9, 372, 393, 395, and 412 have been detected as positively selected sites in a previous study [14]. Sequential amino acid changes at these sites were 9 ($N/S_{95/96} \rightarrow S_{04/05/JP/CHN}$ or $T_{04/05/AU/NL} \rightarrow N_{2006b}$), 372 ($N_{95/96} \rightarrow D_{04/05/JP/CHN}$ or $S_{04/05/AU/NL} \rightarrow E_{2006b}$), 393 ($N_{95/96} \rightarrow S_{04/05/JP/CHN}$ or $S_{04/05/AU/NL} \rightarrow G/S_{2006b}$), 395 ($N_{95/96} \rightarrow A/T_{04/05/JP/CHN}$ or $T_{04/05/AU/NL} \rightarrow T_{2006b}$), and 412 ($T_{95/96} \rightarrow T_{04/05/JP/CHN}$ or $D_{04/05/AU/NL} \rightarrow N_{2006b}$) (Fig. 1). Although immunity to GII.4 has built up in the population due to the predominance of this variant, amino acid changes have occurred in the P2 subdomain, which contains neutralizing epitopes [5, 6, 15, 21], and the stepwise fixation of amino acid changes in the P2 subdomain is caused by selective pressure due to population immunity, which may have resulted in an emerging variant [14, 25]. In fact, the majority of the amino acid changes were

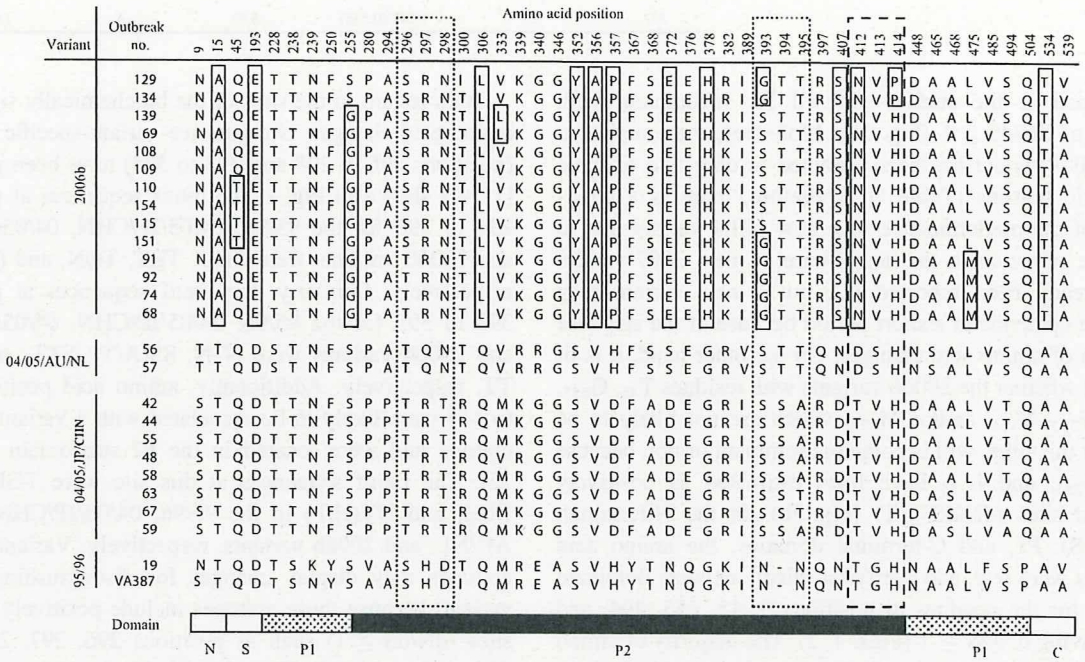

Fig. 1 Amino acid changes in VP1 sequences of genotype GII.4 strains. *Boxes* denote the unique amino acid changes specific to the 2006b variants. *Boxes with dotted lines* denote the sites of putative variant-specific epitopes proposed previously [1]. The *box with the broken line* denotes the additional variant-specific epitope in this study. The GenBank accession number for the VA387 strain is AY038600

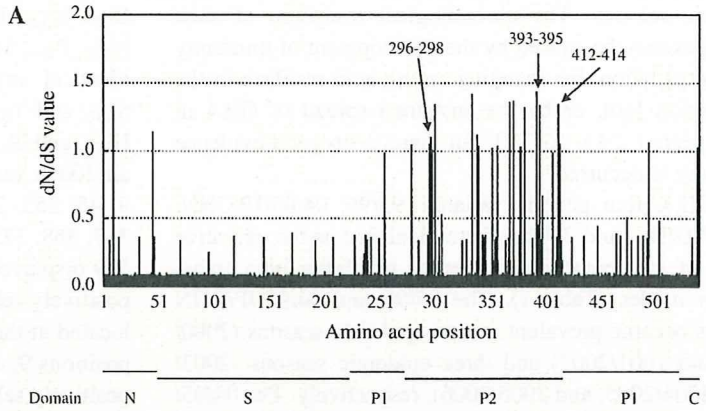

Fig. 2 Bayes-Emperical-Bayes (BEB) positively selected sites in the VP1 region of genotype GII.4 strains. **a** Graphic representation of dN/dS values per site. **b** Sites with a mean value of dN/dS ≥ 1. The dN/dS values were estimated by using codeml with NSsites model 8 in the PAML3.14 package. *Values in parentheses* denote the BEB probabilities

B

Position	Amino acid residue in VA387	Mean value for dN/dS	Position	Amino acid residue in VA387	Mean value for dN/dS
9	S	1.329 (0.796)	357	H	1.034 (0.549)
45	Q	1.143 (0.642)	368	T	1.357 (0.824)
255	S	0.988 (0.512)	372	N	1.365 (0.832)
280	A	1.048 (0.561)	378	G	1.032 (0.548)
296	S	1.050 (0.563)	382	K	1.322 (0.788)
297	H	1.105 (0.608)	393	N	1.431 (0.900)
298	D	1.016 (0.535)	395	N	1.340 (0.807)
300	I	1.348 (0.815)	407	N	0.990 (0.513)
333	M	1.413 (0.879)	412	T	1.342 (0.809)
339	R	1.120 (0.620)	413	G	1.049 (0.562)
340	E	1.037 (0.552)	494	S	1.063 (0.574)
352	S	1.017 (0.535)	539	A	1.027 (0.544)

positioned at the surface-exposed P2 subdomain, with values of dN/dS ≥ 1 (Fig. 2a). Moreover, the emergence of a GII.4 variant has been confined to only two or three epidemic seasons (Table 1). Therefore, there is a possibility of the predominance of a new GII.4 variant due to positive selection in the near future. A new 2007 variant has already been reported [7], but a new variant may become epidemic in a short period because of the stepwise fixation of amino acid changes. On the other hand, it is of interest whether the 2006b variants with residues T_{45}, G_{255}, L_{333}, G_{393}, P_{414}, and/or M_{475}, which are not likely to be fixed at this time, will become predominant or not, because T_{45}, G_{255}, and L_{333} have been estimated as positively selected sites (dN/dS ≥ 1) (Fig. 2b). In the N-terminal, Shell (S), P1, and C-terminal domains, the amino acid changes were few, and purifying selections were dominant except for the residues at positions 9, 45, 255, 494, and 539, having dN/dS ≥ 1 (Figs. 1, 2). The majority of amino acid changes within the S domain occurred during the emergence of the strains gathered in the Grimsby (GenBank accession number AJ004864) cluster, and these became fixed in the population [14].

With regard to the sites of the biochemically significant amino acid changes, two putative variant-specific epitopes (positions 296 to 298 and 393 to 395) have been proposed [1]. As shown in Fig. 1, the motif sequences at positions 296 to 298 for the 95/96, 04/05/JP/CHN, 04/05/AU/NL, and 2006b variants were SHD, TRT, TQN, and (T/S)RN, respectively. Similarly, the motif sequences at positions 393 to 395 for the 95/96, 04/05/JP/CHN, 04/05/AU/NL, and 2006b variants were N–N, SS(A/T), STT, and (G/S) TT, respectively. Additionally, amino acid positions 412 to 414 were likely to be associated with a variant-specific epitope and were located in the P2 subdomain (Figs. 1, 2a). The motif sequences at this site were TGH, TVH, DSH, and NV(H/P) in the 95/96, 04/05/JP/CHN, 04/05/AU/NL, and 2006b variants, respectively. Variant-specific epitopes may display changes for distinguishing a new variant, because these epitopes include positively selected sites (dN/dS ≥ 1) such as positions 296, 297, 298, 393, 395, 412, and 413 (Fig. 2b). Obviously, the amino acid signatures of the three epitopes in the 2006b variant seemed to be related to those of the 04/05/JP/CHN variant rather than other variants, whereas the 2006b variant is

likely a descendant of a strain older than the 2004 variant [25].

An incontestable transition of genotypes and GII.4 variants in NoV gastroenteritis outbreaks was identified in a limited area of Japan, Hiroshima Prefecture. At this time, the main causative agent of NoV gastroenteritis outbreaks is a 2006b variant of GII.4. It is unclear how long the predominance of GII.4 will continue.

The full-length nucleotide VP1 sequences of the 25 GII.4 strains in this study have been deposited in the DDBJ database with the accession numbers AB504306 to AB504330.

References

1. Allen DJ, Gray JJ, Gallimore CI, Xerry J, Iturriza-Gómara M (2008) Analysis of amino acid variation in the P2 domain of the GII-4 norovirus VP1 protein reveals putative variant-specific epitopes. PloS One 3:e1485
2. Bucardo F, Nordgren J, Carlsson B, Paniagua M, Lindgren P-E, Espinoza F, Svensson L (2008) Pediatric norovirus diarrhea in Nicaragua. J Clin Microbiol 46:2573–2580
3. Buesa J, Montava R, Abu-Mallouh R, Fos M, Ribes JM, Bartolomé R, Vanaclocha H, Torner N, Domínguez A (2008) Sequential evolution of genotype GII.4 norovirus variants causing gastroenteritis outbreaks from 2001 to 2006 in eastern Spain. J Med Virol 80:1288–1295
4. Bull RA, Tu ETV, McIver CJ, Rawlinson WD, White PA (2006) Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. J Clin Microbiol 44:327–333
5. Chakravarty S, Huston AM, Estes MK, Prasad BVV (2005) Evolutionary trace residues in noroviruses: importance in receptor binding, antigenicity, virion assembly, and strain diversity. J Virol 79:554–568
6. Chen R, Neill JD, Estes MK, Prasad BVV (2006) X-ray structure of a native calicivirus: structural insights into antigenic diversity and host specificity. Proc Natl Acad Sci USA 103:8048–8053
7. Chhabra P, Dhongade RK, Kalrao VR, Bavdekar AR, Chitambar SD (2009) Epidemiological, clinical, and molecular features of norovirus infections in western India. J Med Virol 81:922–932
8. Fukuda S, Sasaki Y, Takao S, Seno M (2008) Recombinant norovirus implicated in gastroenteritis outbreaks in Hiroshima Prefecture, Japan. J Med Virol 80:921–928
9. Gallimore CI, Cubitt D, du Plessis N, Gray JJ (2004) Asymptomatic and symptomatic excretion of noroviruses during a hospital outbreak of gastroenteritis. J Clin Microbiol 42:2271–2274
10. Iritani N, Vennema H, Siebenga JJ, Siezen RJ, Renckens B, Seto Y, Kaida A, Koopmans M (2008) Genetic analysis of the capsid gene of genotype GII.2 noroviruses. J Virol 82:7336–7345
11. Jin M, Xie H-P, Duan Z-J, Liu N, Zhang Q, Wu B-S, Li H-Y, Cheng W-X, Yang S-H, Yu J-M, Xu Z-Q, Cui S-X, Zhu L, Tan M, Jiang X, Fang Z-Y (2008) Emergence of the GII.4/2006b variant and recombinant noroviruses in China. J Med Virol 80:1997–2004
12. Kageyama T, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Kojima S, Takai R, Oka T, Takeda N, Katayama K (2004) Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to *Norovirus* in Japan. J Clin Microbiol 42:2988–2995
13. Kojima S, Kageyama T, Fukushi S, Hoshino FB, Shinohara M, Uchida K, Natori K, Takeda N, Katayama K (2002) Genogroup-specific PCR primers for detection of Norwalk-like viruses. J Virol Methods 100:107–114
14. Lindesmith LC, Donaldson EF, LoBue AD, Cannon JL, Zheng D-P, Vinje J, Baric RS (2008) Mechanisms of GII.4 norovirus persistence in human populations. PloS Med 5:e31
15. Lochridge VP, Jutila KL, Graff JW, Hardy ME (2005) Epitopes in the P2 domain of norovirus VP1 recognized by monoclonal antibodies that block cell interactions. J Gen Virol 86:2799–2806
16. Motomura K, Oka T, Yokoyama M, Nakamura H, Mori H, Ode H, Hansman GS, Katayama K, Kanda T, Tanaka T, Takeda N, Sato H, The Norovirus Surveillance Group of Japan (2008) Identification of monomorphic and divergent haplotype in the 2006–2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history. J Virol 82:11247–11262
17. Nishida T, Nishio O, Kato M, Chuma T, Kato H, Iwata H, Kimura H (2007) Genotyping and quantitation of noroviruses in oysters from two distinct sea areas in Japan. Microbiol Immunol 51:177–184
18. Noda M, Fukuda S, Nishio O (2008) Statistical analysis of attack rate in norovirus foodborne outbreaks. Int J Food Microbiol 122:216–220
19. 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
20. Osawa K, Oka T, Takeda N, Hansman GS (2007) Norovirus infections in symptomatic and asymptomatic food handlers in Japan. J Clin Microbiol 45:3996–4005
21. Parker TD, Kitamoto N, Tanaka T, Huston AM, Estes MK (2005) Identification of genogroup I and genogroup II broadly reactive epitopes on the norovirus capsid. J Virol 79:7402–7409
22. Ramirez S, Giammanco GM, De Grazia S, Colomba C, Martella V, Arista S (2009) Emerging GII.4 norovirus variants affect children with diarrhea in Palermo, Italy in 2006. J Med Virol 81:139–145
23. Saitoh M, Kimura H, Kozawa K, Nishio O, Shoji A (2007) Detection and phylogenetic analysis of norovirus in *Corbicula fluminea* in a freshwater river in Japan. Microbiol Immunol 51:815–822
24. Seto Y, Iritani N, Kubo H, Kaida A, Murakami T, Haruki K, Nisho O, Ayata M, Ogura H (2005) Genotyping of norovirus strains detected in outbreaks between April 2002, March 2003 in Osaka City, Japan. Microbiol Immunol 49:275–283
25. Siebenga JJ, Vennema H, Renckens B, de Bruin E, van den Veer B, Siezen RJ, Koopmans M (2007) Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J Virol 81:9932–9941
26. Zheng D-P, Ando T, Frankhauser RL, Beard RS, Glass RI, Monroe SS (2006) Norovirus classification and proposed strain nomenclature. Virology 346:312–323

Jpn. J. Infect. Dis., 62, 2009

Laboratory and Epidemiology Communications

Molecular Characterization of Sapoviruses Detected in Sporadic Gastroenteritis Cases in 2007 in Ehime Prefecture, Japan

Yuka Ootsuka, Yasutaka Yamashita, Takako Ichikawa**, Reiko Kondo***, Mitsuaki Oseto***, Kazuhiko Katayama[1], Naokazu Takeda[1]****, and Tomoichiro Oka[1]*

Ehime Prefecture Institute of Public Health and Environmental Science, Ehime 790-0003, and [1]Department of Virology II, National Institute of Infectious Diseases, Tokyo 208-0011, Japan

Communicated by Takaji Wakita

(Accepted April 14, 2009)

Sapovirus (SaV) is an important pathogen of acute gastroenteritis in humans (1-7). The SaV genome is a polyadenylated, single-stranded, positive-sense RNA approximately 7.5 kb long, and it has two or three open reading frames (ORFs). ORF1 encodes nonstructural proteins (i.e., protease, polymerase, etc.) and a major structural (capsid) protein, whereas ORF2 and ORF3 encode a putative protein with an unknown function (8). SaV can be divided into at least five genogroups (GI to GV) based on the capsid protein gene sequences, among which GI, GII, GIV, and GV are known to infect humans, whereas GIII infects porcine species (9). Human SaV strains are noncultivable, and electron microscopy (EM), single-round or nested reverse transcription-polymerase chain reaction (RT-PCR) (10-15), and real-time RT-PCR (16) have been the main methods used for SaV detection. At present, characterization of SaV strains is mainly done with short nucleotide sequences of the capsid or polymerase.

The purpose of this study was to determine the nucleotide sequences of the approximately 2.3-kb 3′ end of the SaV genome detected in patients with sporadic gastroenteritis in 2007 in Ehime Prefecture, Japan.

During the sporadic gastroenteritis surveillance from June to November of 2007 at the Ehime Prefecture Institute of Public Health and Environmental Science, SaV was detected in 6 cases using nested RT-PCR with primers SV-F11 and SV-R1 for the first PCR and SV-F21 and SV-R2 for the second PCR (15). The 6 patients were all children (1 to 7 years old) (Table 1), and none of the adults in their families had gastroenteritis symptoms at the time the specimens were collected (data not shown). The symptoms included diarrhea (83.3%), vomiting (50.0%), abdominal pain (60.6%), fever (higher than 37.0°C) (48.6%), chills (43.1%), and headache (24.8%) (Table 1). These specimens were reexamined for SaV using a recently developed universal nested RT-PCR with the primers SV-F13, SV-F14, SV-R13, and SV-R14 for the first PCR and SV-F22 and SV-R2 for the second PCR (14) and real-time RT-PCR (16). All 6 specimens were positive by these two methods, and the number of cDNA copies per gram of feces ranged between 5.52×10^7 and 2.74×10^9 (Table 1). SaV-like particles were detected in 3 of 6 specimens using EM (Table 1). To better characterize these strains, the approximately 2.3-kb 3′ end of the genome including the entire capsid, ORF2, and untranslated region of these 6 strains was amplified with semi-nested RT-PCR followed by direct sequencing analysis as previously described (3). The SaV nucleotide sequences determined in this study were deposited at DDBJ under the accession nos. AB448761 - AB448766. Nucleotide sequences were aligned with ClustalW version 1.83 (http://clustalw.ddbj.nig.ac.jp/top-j.html). A phylogenetic tree with 1,000 bootstrap replications was constructed by the neighbor-joining method. The genetic distances were calculated by Kimura's two-parameter method (17) and illustrated using Njplot software (http://pbil.univ-lyon1.fr/software/njplot.html) (18). The nucleotide and amino acid sequences were analyzed with GENETYX MAC Software, version 12.2.6 (Genetyx Corp., Tokyo, Japan). BLAST (Basic Local Alignment Search Tool; http://blast.ddbj.nig.ac.jp/top-j.html) was used to find homologous hits.

Ehime-742S/07/JP (AB448762) was 2261 nt in length from the capsid start codon to the genome end (excluding the polyA tail) and was genetically closest to Lyon/30388/98/Fr (AJ251991), detected in 1998 in France, with 97% nucleotide identity over the approximately 2.3-kb fragment. In addition, Ehime-742S was close to Ehime/2K-927/00/JP (AM049931), Ehime/01-1527/01/JP (AM049930), and Ehime/04-311/04/JP (AM049933) detected in Ehime Prefecture in 2000, 2001, and 2004, respectively, with 99% nucleotide identities when the 390-nt capsid gene parts were compared (data not shown). Furthermore, this strain was close to Chiba/001092F/00/JP (AJ412814), Chiba/010591F/01/JP (AJ412823), and Chiba040506/04/JP (AM049928), detected in Chiba Prefecture in 2000, 2001, and 2004, respectively, and Moscow/4548/03/JP (AY538721) and Odessa/7800/05/JP (FJ214051), detected in Russia in 2003 and 2005, with 99% nucleotide identities (data not shown).

Ehime-683S/07/JP (AB448761), -1111S/07/JP (AB448764), -1116S/07/JP (AB448765), and -1121S/07/JP (AB448766) had 2264 nt, and Ehime-1097S/07/JP (AB448763) were 2266 nt in length from the capsid start codon to the genome end (exclude polyA tail), respectively. The Ehime-683S, -1097S, -1111S, -1116S, and -1121S strains had very similar nucleotide sequences, with 99.6 to 99.9% nucleotide identities.

*Corresponding author: Mailing address: Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan. E-mail: oka-t@nih.go.jp

**Present address: Ehime Central Prefectural Hospital, Ehime, Japan

***Retired from Ehime Prefecture Institute of Public Health and Environmental Science

****Present address: Research Collaboration Center on Emerging and Re-emerging Infections, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand.

Table 1. Sporadic gastroenteritis due to sapovirus in Ehime Prefecture in 2007

Specimen	Sex	Age	Onset of illness	Specimen collected date	Symptoms at onset of illness				EM	Nested RT-PCR	Real-time RT PCR (copies/g stool)	Accession no.
					Diarrhea	Vomiting	Abdominal pain	Fever[1)]				
683S	M	2	June 30	July 4	+	+		+	+	+	2.74×10^9	AB448761
742S	M	1	July 15	July 18	+	+			+	+	1.63×10^8	AB448762
1097S	M	7	Nov 13	Nov 16	+		+		-	+	5.52×10^7	AB448763
1111S	M	4	Nov 15	Nov 21		+			-	+	2.12×10^8	AB448764
1116S	F	5	Aug 24	Aug 28	+		+		-	+	3.84×10^8	AB448765
1121S	M	1	Nov 17	Nov 19	+				+	+	1.32×10^9	AB448766

[1)]: higher than 37.0 degrees.
EM, electron microscopy.

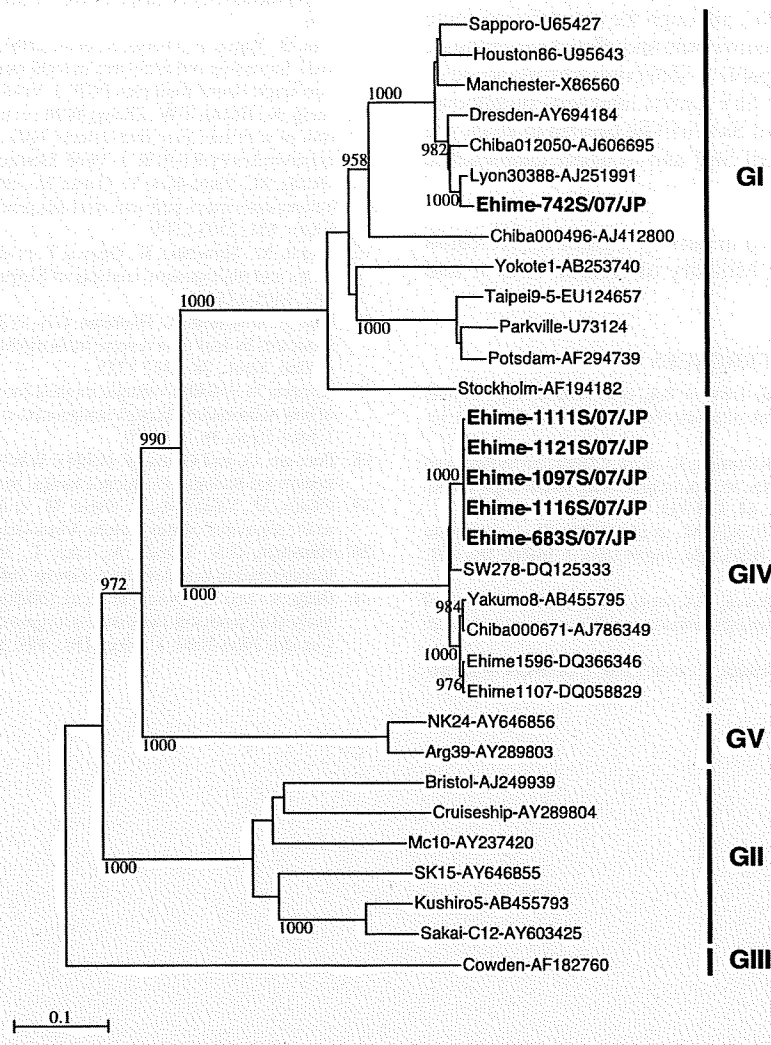

Fig. 1. Phylogenetic tree of SaV based on complete capsid nucleotide sequences. DDBJ accession numbers for Ehime-683S/07/JP (AB448761), -742S/07/JP (AB448762), -1097S/07/JP (AB448763), -1111S/07/JP (AB448764), -1116S/07/JP (AB448765), and -1121S/07/JP (AB448766) are shown in bold letters in the tree. The number on each branch indicates the bootstrap value, where a value of 950 or higher is considered statistically significant for the grouping. The scale represents genetic distances that means nucleotide substitutions per site.

These strains were genetically close to Ehime1596/99/JP (DQ366346) (19) and Ehime1107/02/JP (DQ058829) (20), detected in Ehime in 1999 and 2002, respectively, and to Yakumo8/00/JP (AB455795) detected in Hokkaido in 2000 (3), with 96% nucleotide identities over the approximately 2.3-kb fragment. In addition, the Ehime-683S, -1097S, -1111S, -1116S, and -1121S strains were identical to Yokohama/16/07/JP (AB305049) (2) and close to Osaka19-086/07/JP (AB327281), Saga8151/07/JP (FJ445102), Osaka07-767/08/JP (AB433785), Maizuru8240/08/JP (FJ445108), and Sapporo8411/08/JP (FJ445097), all detected in Japan, with 99% nucleotide identities when the 300- to 400-nt capsid gene parts were compared (data not shown).

Ehime-742S was categorized as GI, whereas Ehime-683S, -1097S, -1111S, -1116S, and -1121S were clustered into GIV based on the complete capsid nucleotide sequences (Fig. 1).

These results indicated that a genetically similar SaV strain belonging to GI likely persisted or circulated between 1998 and 2007 in Japan and Russia, and that strains belonging to GIV likely persisted between 1999 and 2008 in Japan.

In conclusion, this study demonstrated that SaV strains belonging to GIV were predominant among sporadic gastroenteritis cases due to SaV from June to November in 2007 in Ehime. Similar nucleotide sequences were deposited in DDBJ from other prefectures. In addition, the emergence of SaV strains belonging to GIV was reported in 2007 in Canada (21), although no sequence data are currently available for these strains. These results suggest that the SaV strains belonging to GIV may have spread not only in Japan but also throughout the world in 2007, although the transmission route is unknown. Continuous nationwide and global surveillance for SaV using methods capable of detecting a broad range of genetically diverse human SaV strains, and accumulation of the nucleotide sequences of SaV will be needed to study the geographical distribution of SaV and to better characterize SaV strains.

This work was supported in part by a grant for Research on Food Safety from the Ministry of Health, Labour and Welfare of Japan.

REFERENCES

1. Johansson, P.J., Bergentoft, K., Larsson, P.A., et al. (2005): A nosocomial sapovirus-associated outbreak of gastroenteritis in adults. Scand. J. Infect. Dis., 37, 200-204.
2. Usuku, S., Kumazaki, M., Kitamura, K., et al. (2008): An outbreak of food-borne gastroenteritis due to sapovirus among junior high school students. Jpn. J. Infect. Dis., 61, 438-441.
3. Ishida, S., Yoshizumi, S., Miyoshi, M., et al. (2008): Characterization of sapoviruses detected in Hokkaido, Japan. Jpn. J. Infect. Dis., 61, 504-506.
4. Akihara, S., Phan, T.G., Nguyen, T.A., et al. (2005): Identification of sapovirus infection among Japanese infants in a day care center. J. Med. Virol., 77, 595-601.
5. Hansman, G.S., Saito, H., Shibata, C., et al. (2007): Outbreak of gastroenteritis due to sapovirus. J. Clin. Microbiol., 45, 1347-1349.
6. Chiba, S., Nakata, S., Numata-Kinoshita, K., et al. (2000): Sapporo virus: history and recent findings. J. Infect. Dis., 181 (Suppl. 2), S303-308.
7. Wu, F.T., Oka, T., Takeda, N., et al. (2008): Acute gastroenteritis caused by GI/2 sapovirus, Taiwan, 2007. Emerg. Infect. Dis., 14, 1169-1171.
8. Hansman, G.S., Oka, T., Katayama, K., et al. (2007): Human sapoviruses: genetic diversity, recombination, and classification. Rev. Med. Virol., 17, 133-141.
9. Farkas, T., Zhong, W.M., Jing, Y., et al. (2004): Genetic diversity among sapoviruses. Arch. Virol., 149, 1309-1323.
10. Honma, S., Nakata, S., Sakai, Y., et al. (2001): Sensitive detection and differentiation of *Sapporo virus*, a member of the family *Caliciviridae*, by standard and booster nested polymerase chain reaction. J. Med. Virol., 65, 413-417.
11. Vinje, J., Deijl, H., van der Heide, R., et al. (2000): Molecular detection and epidemiology of Sapporo-like viruses. J. Clin. Microbiol., 38, 530-536.
12. Yan, H., Yagyu, F., Okitsu, S., et al. (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.
13. Jiang, X., Huang, P.W., Zhong, W.M., et al. (1999): Design and evaluation of a primer pair that detects both Norwalk- and Sapporo-like caliciviruses by RT-PCR. J. Virol. Methods, 83, 145-154.
14. Okada, M., Yamashita, Y., Oseto, M., et al. (2006): The detection of human sapoviruses with universal and genogroup-specific primers. Arch. Virol., 151, 2503-2509.
15. Okada, M., Shinozaki, K., Ogawa, T., et al. (2002): Molecular epidemiology and phylogenetic analysis of Sapporo-like viruses. Arch. Virol., 147, 1445-1451.
16. Oka, T., Katayama, K., Hansman, G.S., et al. (2006): Detection of human sapovirus by real-time reverse transcription-polymerase chain reaction. J. Med. Virol., 78, 1347-1353.
17. Kimura, M. (1980): A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol., 16, 111-120.
18. Perriere, G. and Gouy, M. (1996): WWW-query: an on-line retrieval system for biological sequence banks. Biochimie, 78, 364-369.
19. Okada, M., Yamashita, Y., Oseto, M., et al. (2006): Genetic variability in the sapovirus capsid protein. Virus Genes, 33, 157-161.
20. Hansman, G.S., Takeda, N., Oka, T., et al. (2005): Intergenogroup recombination in sapoviruses. Emerg. Infect. Dis., 11, 1916-1920.
21. Pang, X., Lee, B., Tyrrell, G., et al. (2009): Epidemiology and genotype analysis of sapovirus associated with gastroenteritis outbreaks in Alberta, Canada: 2004-2007. J. Infect. Dis., 199, 547-551.

Journal of Medical Virology 81:1117–1127 (2009)

Surveillance of Pathogens in Outpatients With Gastroenteritis and Characterization of Sapovirus Strains Between 2002 and 2007 in Kumamoto Prefecture, Japan

Seiya Harada,[1] Mineyuki Okada,[2] Shunsuke Yahiro,[1] Koichi Nishimura,[1] Shigeru Matsuo,[1] Jiro Miyasaka,[1] Ryuichi Nakashima,[1] Yasushi Shimada,[3] Takehiko Ueno,[4] Shigeru Ikezawa,[5] Kuniko Shinozaki,[2] Kazuhiko Katayama,[6] Takaji Wakita,[6] Naokazu Takeda,[6] and Tomoichiro Oka[6]*

[1]*Kumamoto Prefectural Institute of Public Health and Environmental Science, Kumamoto, Japan*
[2]*Chiba Prefectural Institute of Public Health, Chiba, Japan*
[3]*Shimada Childrenś Clinic, Kumamoto, Japan*
[4]*Ueno Pediatric Clinic, Kumamoto, Japan*
[5]*Ikezawa Childrenś Clinic, Kumamoto, Japan*
[6]*Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan*

Infectious acute gastroenteritis is an important public health problem worldwide. A total of 639 stool specimens were tested for the presence of diarrhea pathogens. The specimens were from outpatients with acute gastroenteritis who consulted the pediatric clinic in Kumamoto Prefecture, Japan, from June 2002 to December 2007. Of these, 421 (65.9%) were positive for diarrhea pathogens. Among them were norovirus (NoV) in 260 (61.8%), sapovirus (SaV) in 81 (19.2%), rotavirus in 49 (11.6%), adenovirus in 19 (4.5%), enterovirus in 13 (3.1%), astrovirus in 9 (2.1%), kobuvirus in 1 (0.2%), and bacterial pathogens in 11 (2.6%). Mixed infection (co-infection of viruses) was found in 22 (5.2%) of the 421 pathogen-positive stool samples. NoV was the most prevalent pathogen throughout the study period; however, the SaV detection rate was unexpectedly high and was found to be the secondary pathogen from 2005 to 2007. Genetic analysis of SaV with 81 strains demonstrated that SaV strains belonging to genogroup IV emerged in 2007, and dynamic genogroup changes occurred in a restricted geographic area. This study showed that SaV infection is not as rare as thought previously. ***J. Med. Virol. 81:1117–1127, 2009.*** © 2009 Wiley-Liss, Inc.

KEY WORDS: gastroenteritis; molecular epidemiology; sapovirus; phylogenetic analysis

INTRODUCTION

Infectious acute gastroenteritis is one of the most common diseases among humans [King et al., 2003; Clark and McKendrick, 2004]. Norovirus (NoV) is the dominant cause of viral acute gastroenteritis throughout the world [Buesa et al., 2002; Bon et al., 2005; Okada et al., 2005; Blanton et al., 2006; Ozawa et al., 2007; Svraka et al., 2007; Motomura et al., 2008; Sdiri-Loulizi et al., 2008; Yoon et al., 2008]. The majority of human NoV are divided into two genogroups (GI and GII), which are further subdivided into genotypes [Katayama et al., 2002; Kageyama et al., 2004]. Sapovirus (SaV) is also known as a pathogen of acute gastroenteritis, and SaV is divided into five genogroups (GI to GV), among which GI, GII, GIV, and GV infect humans, whereas SaV GIII infects porcine species [Farkas et al., 2004; Hansman et al., 2007b]. Group A and C rotaviruses (RoV), astrovirus (AstV), enterovirus (EntV), kobuvirus (KV), adenovirus (AdV) 40 and 41, *Campylobacter*, *Escherichia coli* causing diarrhea, *Salmonella*, and *Shigella* are also important pathogens for gastroenteritis in humans [Bon et al., 1999; Phan et al., 2005; Colomba et al., 2006; Hien et al., 2007; Pham et al., 2007; Tcheremenskaia et al., 2007; Phan et al., 2007a; Ambert-Balay et al., 2008]. Polymerase chain reaction (PCR) has been developed and widely applied to molecular epidemiological studies to detect these virus pathogens in clinical specimens [Bon et al., 1999; Chen et al., 2007; Svraka et al., 2007; Sdiri-Loulizi et al., 2008]. However, little epidemiological data

Grant sponsor: Food Safety from the Ministry of Health, Labour, and Welfare of Japan (partial support).

*Correspondence to: Tomoichiro Oka, Department of Virology II, National Institute of Infectious Diseases, Gakuen 4-7-1, Musashi-murayama, Tokyo 208-0011, Japan.
E-mail: oka-t@nih.go.jp

Accepted 13 January 2009
DOI 10.1002/jmv.21454
Published online in Wiley InterScience (www.interscience.wiley.com)

regarding acute gastroenteritis in outpatients had been reported [Bon et al., 1999].

In this study the etiological agent(s) in stool specimens have been explored from outpatients who consulted a pediatric clinic from June 2002 to December 2007 in Kumamoto Prefecture, Japan, and characterized the SaV strains detected during the study period. NoV was the most prevalent pathogen throughout the study period; however, the SaV detection rate was unexpectedly high and was found to be the secondary pathogen between 2005 and 2007 in this area.

MATERIALS AND METHODS

Stool Specimens

Stool specimens were collected from 639 patients, 357 males and 282 females, with acute gastroenteritis who consulted three pediatric clinics from June 2002 to December 2007. Acute gastroenteritis was defined as the exhibition of one or more of the following symptoms: vomiting, abdominal pain, and/or diarrhea. The ages of the patients ranged from 1 month to 60 years old with the following distribution, 0 years old (97 individuals), 1 (135), 2 (86), 3 (61), 4 (58), 5 (48), 6 (34), 7 (30), 8 (26), 9 (22), 10 (7), 11 (14), 12 (3), and >12 (18). The specimens were collected at clinics during the symptomatic period and transfer to the Kumamoto Prefectural Institute of Public Health and Environmental Science, and stored at −80°C until use.

Bacterial Culture

Stool samples were cultured for *Campylobacter*, *E. coli* causing diarrhea, *Salmonella*, and *Shigella* by standard methods [Cheryl et al., 1999; Nachamkin, 1999].

Reverse Transcription (RT) and Polymerase Chain Reaction (PCR)

Nucleic acids were extracted from 140 μl of a 10% (w/v) stool suspension with a QIAamp Viral RNA kit (Qiagen, Tokyo, Japan) according to the manufacturer's protocol. Twenty-four μl of the nucleic acid solution was mixed with 3 μl of 5× ReverTra Ace RT buffer (Toyobo, Osaka, Japan), 0.4 μl DNase I (5 U/μl) (Takara, Shiga, Japan), and 2.6 μl of distilled water, and incubated at 37°C for 30 min and then at 80°C for 10 min. Eight microliters of the solution was mixed with 4.8 μl of 5× ReverTra Ace RT buffer (Toyobo), 6 μl of 2.5 mM dNTPs (Takara), 1.5 μl of ReverTra Ace (100 U/μl) (Toyobo), 1.5 μl of random hexamer primers (Takara), 1.0 μl RNase inhibitor (40 U/μl) (Toyobo), and 7.2 μl of distilled water, and incubated at 30°C for 10 min and then at 42°C for 1 h. The solution was incubated at 99°C for 5 min and stored at −30°C.

Quantitative and genogroup-specific TaqMan-based real-time RT-PCR was performed for NoV as described previously [Kageyama et al., 2003]. RT-multiplex PCR was performed targeting group A and group C RoVs, and AdVs using primers Beg9 and VP7-1 for group A RoV, G8NS1 and G8NA2 for group C RoV, and Ad1 and Ad2 for AdV as previously described [Yan et al., 2004] in a final volume of 25 μl of the reaction mixture containing 1.25 μl of cDNA for RoV, 1.25 μl of DNA without DNase I treatment for AdV, 2.5 μl of 10 × ExTaq DNA polymerase buffer, 2.0 μl of 2.5 mM dNTPs, 1.2 μl of mixed primers (10 μM each), 0.125 μl of ExTaq DNA polymerase (5 U/μl)(Takara), and 16.675 μl of distilled water [Yan et al., 2004]. PCR was performed at 94°C for 3 min, followed by 40 cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 60 sec, and a final extension at 72°C for 5 min, and then the temperature was held at 4°C, generating 395, 352, and 482 bp for group A RoV, group C RoV, and AdV, respectively. RT-multiplex PCR targeting SaV, AstV, KV, and EntV was performed using primers SR80 and JV33 for SaV [Vinje et al., 2000], Mon269 and Mon270 for AstV [Noel et al., 1995], C6261 and C6779 for KV [Yamashita et al., 2000], and EvP4 and OL68-1 for EntV [Olive et al., 1990; Ishiko et al., 2002] in a final volume of 25 μl of the reaction mixture containing 2.5 μl of cDNA, 2.5 μl of 10× ExTaq DNA polymerase buffer, 2.0 μl of 2.5 mM dNTPs, 1.6 μl of mixed primers (3.125 μM each), 0.125 μl of ExTaq DNA polymerase (5 U/μl), and 16.275 μl of distilled water. PCR was performed at 94°C for 3 min, followed by 40 cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 60 sec, and a final extension at 72°C for 5 min, and then the temperature was held at 4°C, generating 320, 449, 519, and 650 bp for SaV, AstV, KV, and EntV, respectively.

RT-PCR for SaV was performed with the primers SaV124F, SaV1F, SaV5F, and SaV1245R, which were designed for recently developed quantitative real-time RT-PCR [Oka et al., 2006], in a final volume of 25 μl of reaction mixture containing 2.5 μl of cDNA, 2.5 μl of 10× ExTaq DNA polymerase buffer, 2.0 μl of 2.5 mM dNTPs, 1.2 μl of mixed primers (10 μM each), 0.125 μl of ExTaq DNA polymerase (5 U/μl), and 16.675 μl of distilled water. PCR was performed at 94°C for 3 min followed by 40 cycles of 94°C for 30 sec, 50°C for 30 sec, 72°C for 60 sec, and a final extension at 72°C for 5 min, and then the temperature was held at 4°C, generating 105 bp.

When the stool specimens were found positive for SaV by either multiplex RT-PCR or RT-PCR, three recently developed methods, namely (i) the universal nested RT-PCR with primers SV-F13, SV-F14, SV-R13, and SV-R14 for the first PCR and SV-F22 and SV-R2 for the second PCR [Okada et al., 2006], (ii) genogroup-specific RT-PCR with primers SV-F13, SV-F14, SV-G1-R, SV-G2-R, SV-G4-R, and SV-G5-R [Okada et al., 2006], and (iii) quantitative real-time RT-PCR with primers SaV124F, SaV1F, SaV5F, and SaV1245R and MGB TaqMan probes SaV124TP and SaV5TP [Oka et al., 2006], were carried out to evaluate the detection rates of these methods. All PCR products were analyzed by 2% agarose gel electrophoresis in the presence of ethidium bromide (0.5 μg/ml), and the gel was visualized under UV light.

DNA Sequencing and Phylogenetic Analysis of Sapovirus

The PCR products generated by universal nested RT-PCR or genogroup-specific RT-PCR were purified by