THE UNIVERSITY OF TOKYO
GRADUATE SCHOOL OF MEDICINE

MOLECULAR EPIDEMIOLOGY OF VIRUSES CAUSING ACUTE GASTROENTERITIS IN JAPANESE CHILDREN

日本の急性胃腸炎患児における下痢症ウイルスの分子疫学的研究

CHAN-IT WISOOT

THESIS ADVISOR: PROFESSOR MASASHI MIZUGUCHI

TOKYO 2011

TABLE OF CONTENT

ABSTRACT ····· **3**

ABBREVIATIONS ····· **6**

LIST OF TABLES ····· **8**

LIST OF FIGURES ····· **9**

LIST OF PUBLICATIONS ····· **11**

INTRODUCTION ····· **12**

AIMS OF THE STUDY ····· **24**

MATERIALS AND METHODS ····· **25**

Fecal specimens

Extraction of viral genome

Reverse transcription

Virus detection by multiplex PCR

Group A rotavirus G-genotyping

Group A rotavirus P-genotyping

Norovirus and sapovirus genotyping

Amplification of polymerase and capid genes of noroviruses

Amplification of polymerase and capid genes of sapoviruses

Electrophoresis

Nucleotide sequencing and phylogenetic analyses

Protein structure modeling

RESULTS, DISCUSSION, AND CONCLUSION

Chapter I: Diversity of diarrhea virus infections and detection and genetic characterization of noroviruses……...35

Chapter II: Detection and genetic characterization of rotaviruses……52

Chapter III: Detection and genetic characterization of sapoviruses……...67

Chapter IV: Emergence of a new variant of GII.6 norovirus in Shizuoka..79

SUMMARY……91

FUTURE WORKS……..93

ACKNOWLEDGEMENTS ····· **94**

REFERENCES ····· **95**

APPENDIX ····· **105**

ABSTRACT

Acute gastroenteritis remains one of the most common illnesses of children worldwide. Viral pathogens, such as rotavirus, norovirus and sapovirus, are the most frequent causative agent for the most severe disease in children. The accumulated data revealed that the predominant genotype of diarrhea viruses in each epidemic season changed over time. Thus, the aim of this study was to follow up the molecular epidemiology of diarrhea viruses in Japanese children with acute gastroenteritis.

The study was carried out on 1,195 fecal specimens collected from children of acute gastroenteritis in six different locations of Japan: Sapporo, Saga, Tokyo, Osaka, Maizuru, and Shizuoka. The specimens were divided into two groups: (1) 1,018 collected from Sapporo, Saga, Tokyo, Osaka, and Maizuru during 2007-2009 and (2) 187 from Shizuoka during 2008-2009. The presence of diarrhea viruses was examined by multiplex PCR, semi-nested PCR, and DNA sequencing.

Diarrhea viruses were detected in 54.5% (652 of 1,195). Norovirus was recognized as the causative agent of illness in 26.6% of the children with acute gastroenteritis in five areas (Sapporo, Saga, Tokyo, Osaka, Maizuru) during 2007-2009. GII.4/2006b variant accounted for 100% of the detected GII.4 strains. However, other genotypes, such as GII.1, GII.2, GII.3, GII.6, GII.7, GII.12, and GII.14, were also detected. Moreover, GII.6/GII.14 recombinant strains emerged, for the first time in Japan, and became the second leading strain (11.9%) after the GII.4/2006b variant in 2007-2008. Homology modeling of the GII.14 capsid protein showed several amino acid changes in the surface exposed P domain, suggesting that these changes may be induced by selective pressure, driving virus evolution.

Group A rotavirus was detected in 15.5% (156 out of 1,008). The infection in 2007-2008 (19.3%) was higher than those in 2008-2009 (12.1%). G1P[8] was the

most prevalent (62.8%), followed by G3P[8] (21.8%), G9P[8] (14.7%), and G2P[4] (0.7%). The number of G3P[8] strains increased threefold from the former season (2006-2007) from 7.3% to 21.8%, whereas G2P[4] decreased from 11.4% to 0.7%. In the phylogenetic analysis, G3 rotaviruses were closely related to "the new variant G3" 5091 strain, which previously emerged in Japan and China. Furthermore, nucleotide sequence analysis of 33 P[8]-nontypeable strains revealed 5 nucleotide mismatches at the primer binding site. Based on previously reported (2003-2007) and current (2007-2009) data of rotavirus surveillance in the five areas of Japan, in Sapporo, Osaka, and Maizuru, G1P[8] and G3P[8] were detected at high frequencies, ranging from 47.2-57.7% and 31.7-47.4%, respectively. In Tokyo, G1P[8] (47.4%) was the predominant strain, followed by G9P[8] (20.6%), whereas in Saga, G3P[8] (38.9%) and G9P[8] (36.1%) were identified as the most dominant types. None of G9P[8] was detected in Sapporo. This study highlights the genetic diversity and the significance of rotavirus diarrhea in Japan.

Sapovirus was detected in 3.4% (37 of 1,008). GIV genogroup emerged as the predominant strain in 2007-2008. The change of sapovirus distribution continued in 2008-2009 with the appearance of GI/1 and the disappearance of the GIV strains. The study also showed that the GIV strains had their polymerase gene of genogroup II, suggesting that the GIV strains isolated in 2007-2008 were intergenogroup recombinants (GII/GIV). This study is the first to report the emergence of this GII/GIV recombination in Japan.

In Shizuoka, norovirus accounted for 55.6% of all 187 samples tested, showing a relatively high detection rate. Of note, norovirus GII.6 infection in the present study was identified with a high prevalence rate, accounting for 40.4%, being the second most common genotype following GII.4 infection (53.8%). The GII.6

infection rapidly emerged as a leading genotype during a short period of two months (November and December). At least three distinct GII.6 subclusters (a-c) appeared in different part of the world during the past 19 years. The GII.6 strains in this study clustered together in a single branch in subcluster c. Alignment of the full-length capsid genes revealed that GII.6 strains identified here had amino acid differences, particular in the P2 subdomain up to 10.9-17.5% compared with that of subcluster a and b viruses. These results indicate that the GII.6 that emerged in Shizuoka is a new variant. Homology modeling of the GII.6 capsid domain demonstrates that significant amino acid variations are positioned on loops of P2 and P1 subdomains. This also supports the idea that a new variant may possibly emerge from the accumulation of mutations in the P2 subdomain through immune pressure.

ABBRIVIATIONS

3D	: There dimensions
%	: Percent
^{o}C	: Degree Celsius
µl	: Microliter
µM	: Micromolar
BLAST	: Basic Local Alignment Search Tool
bp	: Base pair
cDNA	: Complementary deoxyribonucleic acid
CVR	: Central variable region
DTT	: Dithiothreitol
DNA	: Deoxyribonucleic acid
dNTP	: Deoxyrinonucleotide triphosphate
dsRNA	: Double stranded RNA
g	: Gravity
HBGA	: Histo-blood group antigen
i.e.	: id est
kb	: Kilobase
min	: Minute
mM	: Millimolar
$MgSO_4$	: Magnesium sulfate
nm	: Nanometer
nt	: Nucleotide
NSP	: Nonstructural protein
ORF	: Open reading frame

PCR	: Polymerase chain reaction
RdRp	: RNA-directed RNA polymerase
RT	: Reverse transcription
RNA	: Rinonucleic acid
RNase	: Ribonuclease
s	: Second
VLP	: Virus-like particles
VP	: Viral protein

LIST OF TABLES

Table 1. Specific primers used in multiplex PCR in order to detect diarrhea viruses ……..…27

Table 2. Specific primers used in PCR in order to characterize G-genotypes of group A rotaviruses……..29

Table 3. Specific primers used in PCR in order to characterize P-genotypes of group A rotaviruses……..……....30

Table 4. Specific primers used in PCR to amplify polymerase and capsid genes of noroviruses and sapoviruses …….....31

Table 5. Distribution of viral infection in children with acute gastroenteritis...36

Table 6. Descriptions of clinical signs of diarrhea diseases……...……37

Table 7. Distribution of GI and GII norovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009……..…39

Table 8. Distribution of G- and P-genotype combinations of rotavirus infection in children in five cities of Japan during 2007-2009……....53

Table 9. Distribution of G- and P-genotype combinations of rotavirus infection in children in five cities in Japan during 2003-2009……..……..61

Table 10. Distribution of sapovirus infection in children in five cities of Japan during 2007-2009…..……...……..68

Table 11. Results of structural validation of GII.14 and GII.6 models obtained from EasyModeller……...…..…105

LIST OF FIGURES

Figure 1. Structure and genome organization of rotavirus……..……..……14

Figure 2. Genomic organization and capsid structure of norovirus…….18

Figure 3. Two mechanisms of persistence of GII.4 noroviruses…….20

Figure 4. Structure and genome organization of sapovirus…….22

Figure 5. Age-related and seasonal distributions of norovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009…….……39

Figure 6. Phylogenetic analysis of the partial capsid sequences of norovirus GI detected in 2007-2008…….40

Figure 7. Phylogenetic analysis of the partial capsid sequences of norovirus GII detected in 2007-2009……...41

Figure 8. Changing genotypes of norovirus on the basis of phylogenetic trees of nucleotide sequences……...43

Figure 9. Identification of 8610/Saga/08 recombinant strain……..43

Figure 10. Amino acid alignment of the P domain of GII.14 strains showing mutations in P2 and P1-2 domains……...45

Figure 11. The 3D structure of a dimer of the GII.14 P domain…….....46

Figure 12. Age-related and seasonal distributions of group A rotavirus infections in children with acute gastroenteritis in five cities of during 2007-2009……....53

Figure 13. Phylogenetic analyses of VP7 nucleotide sequence genes of G1, G3, and G9 rotaviruses……..…55

Figure 14. Sequence alignments of the P[8] specific primer binding region showing five nucleotide differences……...…..58

Figure 15. Phylogenetic analyses of VP4 nucleotide sequence genes P[8]-nontypeable rotaviruses…….……...……59

Figure 16. Age-related and seasonal distributions of sapovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009…….....68

Figure 17. Phylogenetic tree of nucleotide sequences of sapovirus strains detected during 2007-2008…….……..69

Figure 18. Phylogenetic tree of nucleotide sequences of sapovirus strains detected during 2008-2009……..…….70
Figure 19. Changing genotypes of sapoviruses on the basis of phylogenetic trees of nucleotide sequences……..…….72
Figure 20. Identification of 8206/Maizuru/08 recombinant strain…..……..72
Figure 21. Amino acid alignment of the full-length capsid gene of 15 GI/1 sapovirus strains showing mutations along the aligned sequences..…74
Figure 22. Distribution of sapovirus genotypes detected in five locations of Japan during 2003 to 2009…….……...…76
Figure 23. Phylogenetic analysis of the partial capsid sequence of norovirus detected in Shizuoka city during 2008-2009…….……..80
Figure 24. Monthly distribution of norovirus infections in Shizuoka during 2008-2009…….……...81
Figure 25. Phylogenetic analysis of the full-length capsid sequences of GII.6 noroviruses…….…….83
Figure 26. Amino acid alignment of the P domain of GII.6 strains showing mutations across the region…….…....84
Figure 27. The 3D structure of a dimer of the GII.6 P domain…….…...86
Figure 28. Sequence alignment between GII.14 norovirus P domain (strain 8610) and template VA387 GII.4…….…..106
Figure 29. Superimposition of GII.14 and template VA387 GII.4 structure and Ramachandran plot analysis of the modeled structure……..…..106
Figure 33. Sequence alignment between GII.6 norovirus P domain (strain 8915) and template VA387 GII.4……...107
Figure 31. Superimposition of GII.6 and template VA387 GII.4 structure and Ramachandran plot analysis of the modeled structure……....107

LIST OF PUBLICATIONS

This thesis is based on the following papers:

1. Chan-it W, Thongprachum A, Khamrin P, Okitsu S, Mizuguchi M, Ushijima H. **Intergenogroup recombinant sapovirus in Japan, 2007-2008**. Emerging Infectious Diseases. 2009;15:1084-7.

2. Chan-it W, Thongprachum A, Okitsu S, Mizuguchi M, Ushijima H. **Epidemiology and molecular characterization of sapovirus and astrovirus in Japan, 2008-2009**. Japanese Journal of Infectious Diseases. 2010;63:302-3.

3. Chan-it W, Thongprachum A, Dey SK, Phan TG, Okitsu S, Nishimura S, Kikuta H, Baba T, Yamamoto A, Sugita K, Hashira S, Tajima T, Mizuguchi M, Ushijima H. **Detection and genetic characterization of rotavirus infections in non-hospitalized children with acute gastroenteritis in Japan, 2007-2009.** Infection Genetics Evolution. 2011. In press.

4. Chan-it W, Thongprachum A, Okitsu S, Nishimura S, Kikuta H, Baba T, Yamamoto A, Sugita K, Hashira S, Tajima T, Mizuguchi M, Ushijima H. **Detection and genetic characterization of norovirus infections in children with acute gastroenteritis in Japan, 2007-2009**. Clinical Laboratory. 2011. In press.

INTRODUCTION

Diarrhea remains one of the most common illnesses of children worldwide. In developing countries, it is the third most common cause of death, with an estimated 2 million deaths per year or 17% of all deaths in children younger than 5 years. In developed countries, diarrhea is the second most common cause of hospital admission and doctor visits. Although many different microorganisms (bacteria, parasites, and viruses) cause diarrhea, viral pathogens: rotavirus, norovirus and sapovirus are the most frequent causative agent for the most severe disease in children younger than 5 years worldwide [1].

Rotavirus

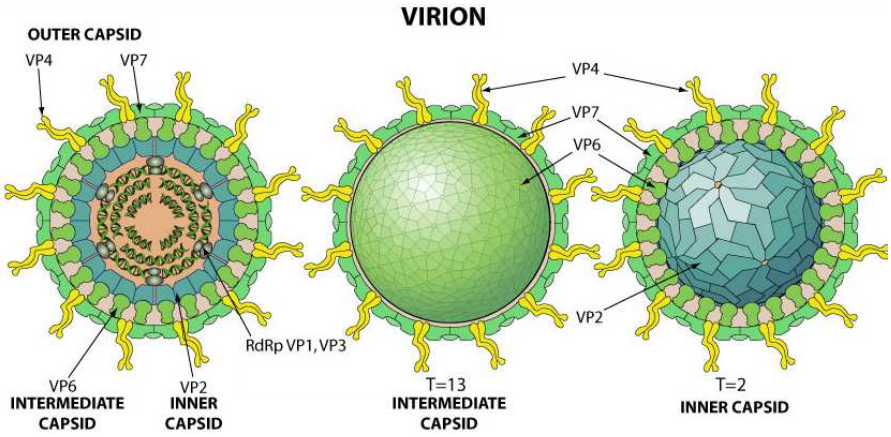
Rotavirus is the most common cause of severe gastroenteritis in infants and young children. By the age of 5 years, 95% of children will have experienced at least one rotavirus infection, with or without evidence of gastroenteritis symptoms. It is estimated that, globally, 1 in 5 cases will be presented to a doctor, 1 in 65 will require hospitalization, and 1 in 293 will die [1-3]. More than 90% of the rotavirus-associated deaths occur in developing countries [4], whereas morbidity due to rotavirus imposes considerable economic burden in the developed world. A typical symptom varies from mild illness with self-limiting (within 48 hours) watery diarrhea to severe diarrhea accompanied by up to 24 hours of vomiting (80%-90% of children) and fever, with risks for dehydration with electrolyte imbalance, shock, and death. Symptoms generally resolve after 3 to 7 days.

Rotaviruses are classified as belonging to the genus *Rotavirus*, family *Reoviridae*. They are non-enveloped, icosahedral viruses displaying a wheel-shaped appearance. Rotavirus is comprised of three concentric protein layers: the inner layer

consisting of VP2 encloses the genome and two minor structural proteins, VP1 and VP3, thus forming the 'core'; the intermediate layer consisting of VP6 surrounds the core, forming a double layered particle; the outer layer, consisting of VP7 and spike-like projections of VP4, enwraps the double layered particle to form the triple-layered particle or infectious virion (Fig. 1A) [5]. The genome of rotavirus consists of 11 segments of double-stranded RNA encoding six structural viral proteins (VP1-VP4, VP6-VP7) and six nonstructural proteins (NSP1-NSP6). With the exception of genome segment 11 which encodes 2 proteins (NSP5 and NSP6), the genome segments are monocistronic (Fig. 1B).

Seven groups (A–G) of rotavirus have been detected to date by antigenicity of the intermediate layer VP6. Group A-C mostly identified in humans and animals. Of these, group A rotavirus is a major cause of human rotavirus-associated gastroenteritis. Two outer capsid proteins, VP7 and VP4, independently elicit a neutralizing response and define different serotype specificities, G *sero*type (G for glycoprotein) and P *sero*type (P for protease-sensitive), respectively [5]. Classification of rotaviruses into P or G serotypes is performed by cross-neutralization assays using hyperimmune sera raised to prototype viruses and/or to laboratory-engineered mono-reassortants. Since antigenic characterization is time-consuming and requires virus collections and proper immunological reagents that are not available in all the laboratories, and due to the increasing ease of sequencing, the antigenic classification has gradually been replaced by a classification system of rotaviruses into P or G *geno*types, performed by sequence analyses and based on identities between sequences of cognate rotavirus gene segments. So far, 24 G and 33 P genotypes have been established for existing human and animal rotaviruses [6-11]. In human rotaviruses, the major genotypes are G1, G2, G3, G4, and G9, which are combined with P[4], P[6], and P[8] [12].

A



B

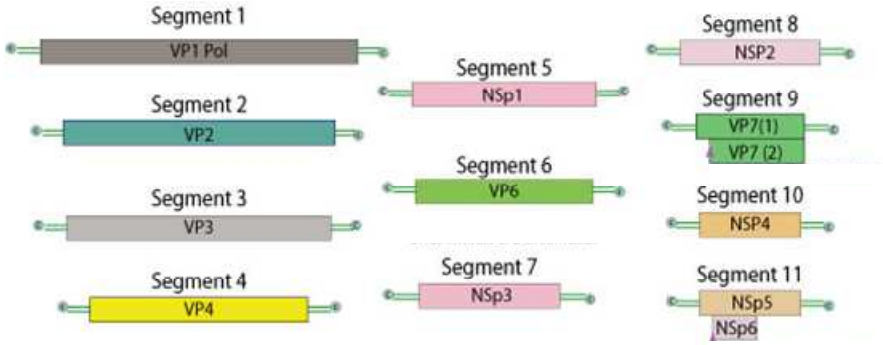


Figure 1. Structure and genome organization of rotavirus. (A) Rotavirus is a non-enveloped virus with an icosahedral capsid (about 70 nm in diameter). (B) Genome is made up of 11 segments of dsRNA held in the inner core of the three-layered virus (http://expasy.org/ viralzone/all_by_species/107.html).

The five G-P combinations, G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8], account for nearly 90% of strains worldwide, four of which share the same VP4 genotype; G1P[8] accounts for 60-80% of strains most years [12]. Other G genotypes such as G5, G6, G8, G10, and G12 in various combinations with P[4], P[6], P[8], P[9], and P[14] types have been detected in some regions of the world [5]. For example, G5P[8], G8P[6], and G9P[6] were detected as the major strains in Brazil, Malawi and India, respectively, but they do not appear to be of global impact [12, 13]. In addition, the emergence of G12 was also reported in several countries [14].

Despite improvement in hygiene and sanitation in recent years, the incidence of rotavirus infection remains unchanged [15]. Thus, developing a safe and effective vaccine to control rotavirus diarrhea in both developed and developing countries has been a high priority. Two live attenuated oral rotavirus vaccines, a pentavalent bovine derived vaccine (RotaTeq; Merck Vaccines, Whitehouse Station, NJ) and a monorovirusalent human rotavirus vaccine (Rotarix; GlaxoSmithKline Biologicals, Rixensart, Belgium) have recently licensed in many countries. In 2006 the World Health Organization (WHO) strongly recommended the inclusion of rotavirus vaccines into national immunisation programmes in American and European countries, and in April 2009 WHO extended this recommendation to all regions of the world after review of clinical trial data from Africa and Asia, and postlicensure data from the Americas [16, 17].

In making a global recommendation for rotavirus vaccines and noting the variable performance of the vaccines in different populations, WHO emphasized the need for further efficacy and effectiveness data, particularly in developing countries [17]. For example, the efficacy of monorovirusalent vaccine against severe rotavirus diarrhea was 96% (95% confidence interval 83% to 99%) in Europe, 83% (73% to 90%) in Latin America, 72% (40% to 88%) in South Africa, and 49% (11% to 72%) in Malawi [18-21]. Similarly, the pentavalent vaccine prevented 98% (88% to 100%) of severe rotavirus infections in the United States and Finland, but effectiveness was 46% (18% to 64%) in Nicaragua, the only developing country setting for which data are available on the performance of the pentavalent vaccine [22, 23].

Although both vaccines could have a substantial impact on the global burden of severe diarrhea in childhood in reducing the hospitalization of the most common rotavirus infections, different vaccines may vary in efficacy against various strains.

Therefore, vaccine efficiency against strains that do not share any G-P combination with the vaccine strains has been questioned. Thus, surveillance of rotavirus G and P genotypes in different areas of the world is important in order to monitor new G-P combinations and arrive at a better understanding of the effect of strain variation on vaccine efficacy.

Norovirus

Globally, noroviruses are the second most important cause of severe viral gastroenteritis in young children after rotavirus, cause traveler's diarrhea in all ages, and are especially virulent in the elderly. The Centers for Disease Control and Prevention report that noroviruses account for over 96% of all viral gastroenteritis cases, with at least 23 million infections occurring annually in the United States alone [24]. In addition, noroviruses cause up to half of all outbreaks of gastroenteritis [25], making this virus the most common cause of sporadic diarrhea in community settings [26]. There are at least two factors that contribute to the virus' ability to cause outbreaks. First, norovirus infections require a very low infectious dose of <10 virions per individual to infect 50% of those individual (ID50) [27]. Second, the virus is extremely stable in the environment, showing resistance to freezing, heating to 60ºC, disinfection with chlorine, acidic conditions, vinegar, alcohol, aseptic hand solutions, and high sugar concentration.

Noroviruses are transmitted mainly via the fecal-oral route through person-to-person contact or consumption of contaminated food. Clinical norovirus infection generally has an incubation period of 27-48 hours and is characterized by acute onset of nausea, vomiting, abdominal cramps, myalgias, and non-bloody diarrhea. Norovirus illness can present with relatively severe symptoms of vomiting and non-

bloody diarrhea, with symptoms usually resolving in 2-3 days, but can last longer (i.e., 4-6 days) in nosocomial outbreaks and among children younger than 11 years of ages [28].

Noroviruses belong to the family *Caliciviridae*, genus *Norovirus*, and are small non-enveloped, icosahedral viruses with a diameter of 38 nm. The virus was first described in 1968 during an outbreak in an elementary school in Norwalk, Ohio [29]. Norovirus encodes a 7.5-7.7 kb positive-sense, single-stranded RNA genome. The RNA genome is organized into three open reading frames (ORF) 1, 2, and 3. ORF1 encodes a large polyprotein that is cleaved by the viral proteinase into six non-structural proteins; ORF2 encodes major capsid protein (VP1), whereas ORF3 encodes minor capsid protein (VP2) (Fig. 2). The capsid protein is further divided into two domains known as the shell domain (S), which forms the inner core and the protruding domain (P), linked by a flexible hinge. The P domain forms prominent protrusions that extend away from the structure and is subdivided into two subdomains (Fig. 2). These include P1, which acts as a stem region between S and the second P subdomain, P2, which is the most surface exposed region of the capsid protein. The sequence of the P2 subdomain is the most variable region in the norovirus genome, and it contains the carbohydrate ligand-binding site(s) and a motif resembling an RNA-binding domain. Immune pressure and variant histo-blood group antigen (HBGA) carbohydrate–receptor interactions likely drives the evolution of the distal regions of P2 as well as some surface exposed P1 residues, which can result in norovirus patterns of evolutionary and structural changes in the protein [30-34].

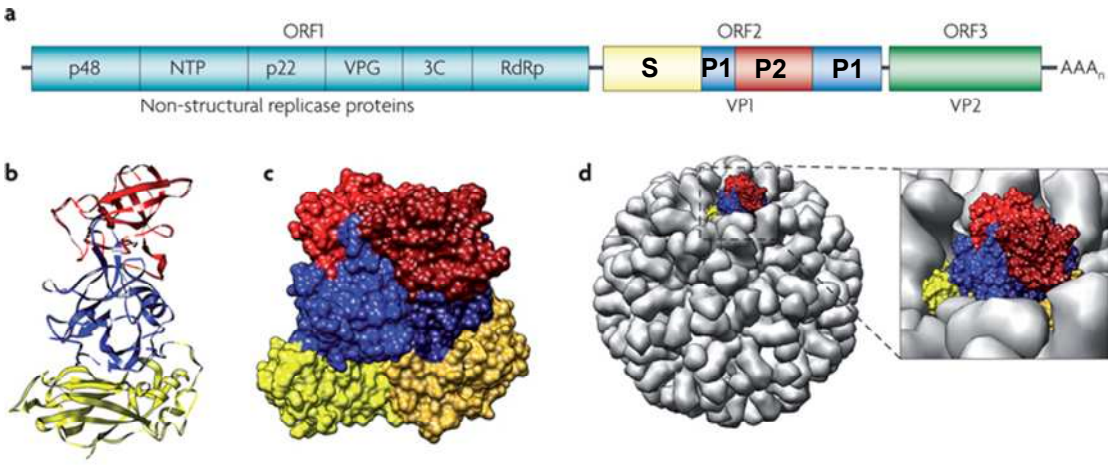


Figure 2. Genome organization and capsid structure. a | The norovirus genome is composed of three open reading frames (ORFs). ORF1 (~5 kb) is located in the first two-thirds of the genome and encodes a ~200 kDa polyprotein that is auto-processed by a virally encoded 3C-like protease (3C) to yield the non-structural replicase proteins that are essential for viral replication89. The resultant proteins are: p48, an amino-terminal protein of unknown function (~48 kDa); nucleoside triphosphatase (NTP), a 2C-like protein; p22, a 22 kDa 3A-like protein; viral genome-linked protein (VPG), a protein that is covalently linked to the 5′ end of the genome; and RNA-directed RNA polymerase (RdRp), a 3D-like protein. ORF2 is 1.8kb in length and encodes the 57 kDa major structural capsid protein, viral protein 1 (VP1). VP1 is divided into two domains, the shell domain (yellow) and the protruding domain, which is further divided into two subdomains known as P1 (blue) and P2 (red). ORF3 is ~0.6 kb in length and encodes a 22 kDa minor basic structural protein, VP2 (REF. 89). **b** | The structure of the VP1 monomer is shown, with protein domains coloured as for part **a**. **c** | Two capsid protein monomers form the A–B dimmer (indicated with the A monomer in lighter shades and the B monomer in darker shades), which allows the P2 domain to protrude from the viral particle. **d** | The virus-like particle is formed of 180 monomers of the capsid protein that assemble through different dimers. The A–B dimer, shown in colour, extends away from the capsid and provides the receptor-binding region and the sites of antigenic variation. In the virus particle, VP2 is incorporated in low copy number. Structural models were generated and pictures were rendered using MacPyMOL (Delano Scientific LLC, Palo Alto, California, USA) [35].

Noroviruses are genetically diverse and presently classified, based on the VP1 capsid sequence, into five different genogroups (GI-GV). GI and GII genogroups are further subdivided into 8 and 17 genotypes, respectively [36]. The most clinically significant of the five genogroups is GII, as it is the most prevalent human norovirus genogroup detected and more frequently associated with epidemics compared with other genogroups. Of particular interest is GII genotype 4, (GII.4), because this lineage accounts for 62% of all norovirus outbreaks globally [37, 38] and has also caused all five major norovirus pandemics in the last decade (1995/1996, U5-95_US

strain; 2002, Farmington Hills; 2004, Hunter; 2006, 2006a virus; and 2007, 2006b virus) [39-41]. The basis for the increased epidemiological fitness of the GII.4 strains, as determined by its high incidence and ability to cause pandemics, is currently unknown.

Donaldson et al.'s observations [42] suggest that the GII.4 viruses persist in the human population via two different mechanisms. First, changes in and around the receptor-binding domain allow for subtle remodeling of the structure, which can lead to altered HBGA binding, suggesting that the virus utilizes a receptor switching mechanism to persist by targeting unique populations for infection. Second, additional changes on the P2 surface appear to correlate with altered antigenicity, suggesting that the second mechanism of persistence is by escape from pre-existing herd immunity (Fig. 3). These two mechanisms likely occur in an overlapping fashion, such that changes may alter both receptor usage and antigenicity simultaneously. Obviously, detailed analyses of human HBGA phenotypes from outbreak settings in different years and locations would provide important support for this hypothesis, while also recognizing that even minor differences in capsid sequence within the time-ordered subclusters might impact susceptibility phenotypes.

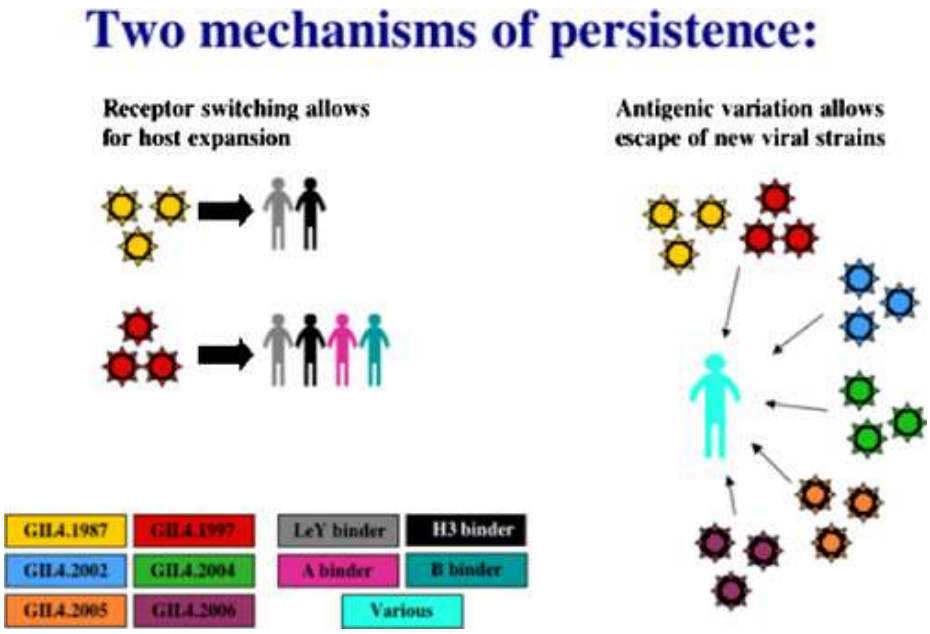


Figure 3. Two mechanisms of persistence of GII.4 noroviruses. GII.4 viruses persist in the human population by two mechanisms. First, adaptation, driven by herd immunity, results in viruses capable of binding different and sometimes novel HBGA receptors, allowing an expansion of host range or penetrance into a previously naïve population. Second, the variation allows escape from the predominant herd immunity, resulting in a virus competent to infect the same population [42].

Recombination is a common event in RNA viruses [43]. It has been reported in retroviruses [44], flaviviruses [45] and, to a much lesser extent, in negative-sense RNA viruses [46]. In human caliciviruses, the first evidence of homologous recombination was shown in 2000, as Jiang et al. characterized a norovirus recombinant bearing ORFs belonging to norovirus genetic clusters GII.3 and GII.4 [47]. This strain, Arg320/1995/AR, has since been reported to circulate in the community [48]. Recombination is also a major determinant of viral virulence, being implicated in the emergence of new viral strains [43]. Its implications for the development of vaccines and virus control programs are essential.

Preclinical studies have shown that virus-like particles (VLPs) administered to mice as parenteral, oral, or intranasal vaccines are highly immunogenic [49, 50]. Among volunteers, recombinant VLPs expressed in transgenic plants and given orally

and VLPs expressed in baculovirus and administered orally have both been safe and immunogenic [51]. However, many challenges to developing norovirus vaccines remain, including an incomplete understanding of the immune correlates of protection, the lack of good long-term immunity and heterotypic protection against antigenically distinct strains, and the existence of multiple genetic and antigenic types of virus.

Sapovirus

Sapoviruses belong to the family *Caliciviridae*, and are a causative agent of acute gastroenteritis. Sapoviruses were previously detected by electron microscopy (EM). The prototype strain of human sapovirus, the Sapporo virus, was originally discovered from an outbreak in an orphanage in Sapporo, Japan, in 1977 [52]. In that study, Chiba et al. identified viruses with the typical animal calicivirus morphology, the Star-of-David structure, by EM. Besides having this classical structure, sapovirus particles are typically 41-46 nm in diameter and have a cup-shaped depression and/or 10 spikes on the outline (Fig. 4A). The sapovirus genomes are predicted to contain either 2 or 3 main open reading frames (ORF1-3) (Fig. 4B). Sapovirus ORF1 encodes for nonstructural proteins and the major capsid protein (VP1), and ORF2 (VP2) and ORF3 encode proteins of yet unknown functions. On the basis of VP1 nucleotide sequences, sapoviruses have been divided into five genogroups (GI-GV), of which GI, GII, GIV, and GV strains infect humans and GIII strains infect porcine species [53]. Sapovirus genogroups can be further subdivided into genotypes: 8 GI and 5 GII [54]. Genotype GI/1 has played a predominant role in sporadic gastroenteritis occurring in both developed and developing countries [55, 56]. Recently, several recombinant sapovirus strains have been identified [57, 58].

Sapovirus has been implicated in acute sporadic gastroenteritis and in gastroenteritis outbreaks primarily affecting children [59], with a high rate of disease (17.6%) reported in Osaka in 2004–2005 [60]. Sapovirus gastroenteritis appears to produce symptoms milder than those by rotavirus and norovirus, so hospitalization is often unnecessary [59]. Immunological and seroepidemiological studies have found a worldwide distribution of sapovirus as an agent of acute sporadic gastroenteritis [61].

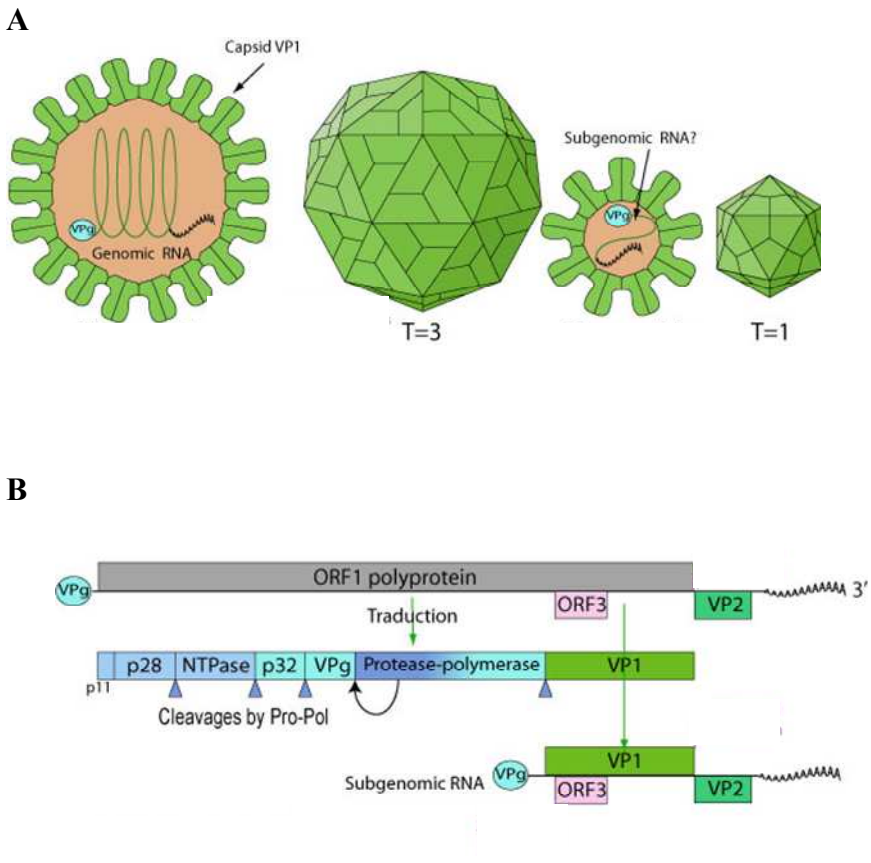


Figure 4. Structure and genome organization of sapovirus. (A) Sapovirus is non-enveloped, icosahedral with T=3 symmetry. The capsid is composed of 180 capsid proteins. (B) Genome is monopartite, linear, single-stranded, positive-sense RNA genome of 7.3 to 8.3 kb. At 5'-terminus a virus protein (VPg) is covalently linked to genome, whereas 3'-terminus is polyadenylated (http://expasy.org/ viralzone/all_by_species/196.html).

In 2003, multiplex RT-PCR was developed and introduced in diagnosis of seven diarrhea viruses, including group A, B, and C rotaviruses, adenovirus, norovirus, sapovirus, and human astrovirus, among non-hospitalized infants and children in the same locations (Sapporo, Saga, Tokyo, Osaka and Maizuru) in Japan [62]. Since its introduction, the identification of viral infections in all five cities in 2003-2007 revealed that group A rotavirus was the most prevalent pathogen (16.7%) [63-66]. Norovirus as the second prevailing virus was detected in 12.5% [67-70], whereas sapovirus was much less common and represented in 3.7% [55, 57, 58, 60, 71, 72]. Regarding the age and seasonal patterns of rotavirus, norovirus, and sapovirus in these five areas, a 1-year-old group was commonly affected by all viruses. While norovirus and sapovirus showed the main peak in winter, rotavirus peaked in early spring.

To follow up the molecular epidemiology of diarrhea viruses from previous studies, I genetically characterized the diarrhea viruses in the earlier samples from Japanese children in five locations (Sapporo, Saga, Tokyo, Osaka and Maizuru) from 2007-2009. In addition, I also investigated the viruses in new stool samples collected from patients in Shizuoka from 2008-2009.

AIMS OF THE STUDY

The general aim of this study was to investigate the molecular epidemiology of diarrhea viruses in Japanese children with acute gastroenteritis.

The specific aims were the followings:

- To identify the presence of diarrhea viruses in fecal specimens collected from children with acute gastroenteritis
- To determine the age-related distribution and seasonal patterns of children infected with norovirus, rotavirus, and sapovirus
- To characterize genotypes of norovirus, rotavirus, and sapovirus and to describe the distribution of these genotypes
- To investigate genetic diversity and recombination in norovirus and sapovirus
- To analyze genetic evolution and to predict the 3D structures of P domain of noroviruses by homology modeling

MATERIALS AND METHODS

Fecal specimens

A total of 1,195 fecal specimens were collected from non-hospitalized children with acute gastroenteritis in pediatric clinics in six different cities (Sapporo, Saga, Tokyo, Osaka, Maizuru, and Shizuoka) of Japan. Of these 1,195 specimens, 1,008 were obtained from five locations (Sapporo, Saga, Tokyo, Osaka, Maizuru) in 2007-2009 and additional 187 from Shizuoka (Fujieda city) in 2008-2010. The case definition of diarrhea used in this study was less than or at least three times of unformed (loose and watery) stool a day. Acute gastroenteritis was defined as the occurrence of diarrhea and other symptoms such as nausea, vomiting, fever, and abdominal cramp. The ages of the subjects ranged from 1 month to 15 years, with a median of 25 months. The Ethical Committees of the University of Tokyo and Aino University approved the study. A parent or guardian of each child provided informed consent. Specimens were stored as undiluted stools at 4°C.

Extraction of viral genome

The fecal specimens were prepared as 10% suspensions in distilled water. Viral genome was extracted from the 10% suspension using the QIAamp Viral RNA Mini Kit, according to the manufacturer's instructions (Qiagen, Hilden, Germany). This QIAamp Viral RNA Mini Kit was also used to extract the DNA genome of adenorovirus. One hundred and forty µl of the fecal supernatant were added into a 1.5 ml microcentrifuge tube containing 560 µl of AVL buffer containing carrier RNA. After incubation at room temperature for 10 min, 560 µl of 100% ethanol were added to the sample and mixed by pulse-vortexing for 15 s. Next, the tube was briefly centrifuged to remove drops from inside the lid. Six hundred and thirty µl of the

solution were carefully transferred to the QIAamp spin column. After centrifuged at 6,000 x *g* for 1 min, 500 µl of buffer AW1 were added to the spin column. The spin column was further centrifuged at 6,000 x *g* for 1 min and 500 µl of buffer AW2 were added. After centrifuged at 20,000 x *g* for 3 min, the spin column was transferred into a new 1.5 ml microcentrifuge tube. The viral genome was dissolved with 50 µl of AVE buffer and kept at –30°C until use.

Reverse transcription

Reverse transcription (RT) was performed with Superscript reverse transcript III (Invitrogen, Carlsbad, CA). Briefly, 5 µl of viral genome extract were first denatured by heating at 95ºC for 5 min in the presence of 1 µl of 50% dimethyl sulfoxide (DMSO) and then immediately chilled on ice. Five µl of the heated genome were added to a reaction mixture (final, 15 µl) containing 3 µl of 5x first strand buffer (Invitrogen, Carlsbad, CA), 0.8 µl of 10 mM dNTPs (Roche, Mannheim, Germany), 0.8 µl of 10 mM DTT (Invitrogen, Carlsbad, CA), 0.8 µl of random primer (hexa-deoxyribonucleotide mixture) (Takara, Shiga, Japan), 0.8 µl of superscript reverse transcriptase III (Invitrogen, Carlsbad, CA), 0.5 µl of RNase inhibitor (Toyobo, Osaka, Japan) and 3.3 µl MilliQ water. The RT step was carried out at 50^{o}C for 1 h, followed by 95°C for 5 min and then chilled on ice for 5 min. The cDNA was kept at –30°C.

Virus detection by multiplex PCR

The presence of diarrhea viruses was detected by multiplex PCR with two sets of specific primers as summarized in Table 1. The first primer set included equimolar mixes of Beg9 [73] and VP7-1' [74], ADG9-1F and ADG9-1R [75], NG8S1 and NG8S2 [76] for group A, B, and C rotaviruses, respectively; Ad1 and Ad2 [77] for

Table 1. Specific primers used in multiplex PCR in order to detect diarrhea viruses in children with acute gastroenteritis in Japan

Group	Target virus	Primer	Polarity[a]	Sequence (5' to 3')[c]	Position[b]
I	Group A rotavirus	Beg9	+	GGCTTTAAAAGAGAGAATTTCCGTCTGG	1-29
		VP7-1'	-	ACTGATCCTGTTGGCCATCCTTT	395-373
	Group B rotavirus	ADG9-1F	+	GGCAATAAAATGGCTTCATTGC	1-22
		ADG9-1R	-	GGGTTTTTACAGCTTCGGCT	814-795
	Group C rotavirus	NG8S1	+	ATTATGCTCAGACTATCGCCAC	353-374
		NG8S2	-	GTTTCTGTACTAGCTGGTGAAC	704-683
	Adenovirus	Ad1	+	TTCCCCATGGCICAYAACAC	19374-19393
		Ad2	-	CCCTGGTAKCCRATRTTGTA	19855-19836
II	Astrovirus	PreCAP1	+	GGACTGCAAAGCAGCTTCGTG	4235-4255
		82b	-	GTGAGCCACCAGCCATCCCT	4953-4934
	Norovirus GI	G1SKF	+	CTGCCCGAATTYGTAAATGA	5342-5361
		G1SKR	-	CCAACCCARCCATTRTACA	5671-5653
	Norovirus GII	COG2F	+	CARGARBCNATGTTYAGRTGGATGAG	5002-5028
		G2SKR	-	CCRCCNGCATRHCCRTTRTACAT	5389-5367
	Sapovirus	SLV5317	+	CTCGCCACCTACRAWGCBTGGTT	5083-5105
		SLV5749	-	CGGRCYTCAAAVSTACCBCCCCA	5516-5494

a: +, Forward primer; -, Reverse primer.
b: The position numbers of primer pairs in the first group correspond to representative strains Wa (K02033), CAL (AF184083), KW408 (AB086968) and Tak (DQ315364) for group A, B and C rotaviruses and adenorovirusirus, respectively. The position numbers of primer pairs in the second group correspond to representative strains Oxford (L23513), Norwalk/68 (M87661), Lordsdale (X86557) and Manchester (X86560) for astrovirus, norovirus (GI, GII) and sapovirus, respectively.
c: Within nucleotide sequence of primers, B=C, G or T; H=A, C or T; N: Any base; R= A or G; S= G or C; V=A, C or G; W= A or T; Y= C or T.

adenovirus. These primer pairs amplified VP7 regions of rotavirus and hexon region of adenovirus. The respective products of these reactions consisted of amplicons of 395 bp, 814 bp, 352 bp and 482 bp for group A, B, and C rotaviruses and adenovirus, respectively. The second primer set included equimolar mixes of PreCAP1 and 82b [78] for astrovirus, G1SKF and G1SKR [40] for norovirus GI, COG2F and G2SKR [79] for norovirus GII, and SLV5317 and SLV5749 [80] for sapovirus. The primer pairs for astrovirus and sapovirus amplified their capsid regions, while the primer pairs for norovirus amplified its polymerase and capsid regions. The respective products of these reactions consisted of amplicons of 719 bp, 330 bp, 387 bp, and 434

bp for astrovirus, norovirus (GI, GII) and sapovirus, respectively. Exactly, 1 µl of cDNA was added to a reagent mixture containing 2.5 µl of 10x Taq DNA polymerase buffer (Promega, Madison, WI), 1 µl of 2.5 mM dNTPs (Roche, Mannheim, Germany), 0.25 µl of each specific primer (25 µM) and 0.05 µl of Taq DNA polymerase (Promega, Madison, WI). MilliQ water was added to make up a total volume of 12.5 µl. The PCR was performed at 94^{o}C for 3 min, followed by 35 cycles of 94°C 30 s, 55°C 30 s, 72°C 60 s, and a final extension at 72°C for 7 min, and then held at 4°C [62].

Semi-nested PCR was performed using G2SKF (5'-CNTGGGAGGGCGATC GCAA-3') [79] and G2SKR primers to detect norovirus GII in samples that were found negative by the second primer set. In this step, to avoid a high risk of cross-contamination, a norovirus positive control was not included in the semi-nested PCR, whereas a negative control was included in the assays throughout the experiment.

Norovirus and sapovirus genotyping

Genotypes of norovirus and sapovirus were directly determined by DNA sequencing of their PCR products obtained from multiplex PCR and semi-nested PCR. Primers G2SKR and SLV5749 were used for the DNA sequencing of the positive norovirus and sapovirus samples, respectively.

Group A rotavirus G-genotyping

The full-length VP7 gene (1,062 bp) was amplified using the consensus primers Beg9 and End9 [73] in the first RT-PCR. The second amplification was performed using the first PCR product as the template with End9 and G-genotype specific mixed primers (BT1, CT2, ET3, DT4, and FT9) [73] (Table 2). These primers

Table 2. Specific primers used in PCR in order to characterize G-genotypes of group A rotaviruses detected in children with acute gastroenteritis in Japan

Target G-genotype	Primer	Polarity[a]	Sequence (5' to 3')	Position[b]
Complete VP7 gene				
Consensus	Beg9	+	GGCTTTAAAAGAGAGAATTTCCGTCTGG	1-28
Consensus	End9	-	GGTCACATCATACAATTCTAATCTAAG	1062-1036
Specific G-genotype				
G1	BT1	+	CAAGTACTCAAATCAATGATGG	314-335
G2	CT2	+	CAATGATATTAACACATTTTCTGTG	411-435
G3	ET3	+	CGTTTGAAGAAGTTGCAACAG	689-709
G4	DT4	+	CGTTTCTGGTGAGGAGTTG	480-498
G9	FT9	+	CTAGATGTAACTACAACTAC	757-776

a: +, Forward primer; -, Reverse primer.
b: The position numbers of primer pairs correspond to representative strains Wa (K02033), L-2 (DQ478581), L243 (DQ873672), R479 (DQ873680) and KUMS04-96 (DQ056298) for G1-G4 and G9, respectively.

specifically generated five different sizes of amplicons of 749 bp, 652 bp, 374 bp, 583 bp, and 306 bp for G1, G2, G3, G4, and G9, respectively. Exactly, 1 μl of the first PCR product was mixed in a reagent mixture containing 2.5 μl of 10x Taq DNA polymerase buffer (Promega, Madison, WI), 1 μl of 2.5 mM dNTPs (Roche, Mannheim, Germany), 0.25 μl of each specific primer (20 μM) and 0.05 μl of Taq DNA polymerase (Promega, Madison, WI). MilliQ water was added to make up a total volume of 12.5 μl. The second PCR was performed at 94°C for 3 min, followed by 35 cycles of 94°C 30 s, 55°C 30 s, 72°C 60 s, and a final extension at 72°C for 7 min, and then held at 4°C.

Group A rotavirus P-genotyping

The partial VP4 gene (879 bp) was amplified using the consensus primers Con2 and Con3 [81] in the first RT-PCR. The second amplification was performed using the first PCR product as the template with Con3 and P-genotype specific mixed primers (1T-1, 2T-1, 3T-1, 4T-1, and 5T-1) [81] (Table 3). These primers specifically generated five different sizes of amplicons of 346 bp, 484 bp, 268 bp, 392 bp, and 584

Table 3. Specific primers used in PCR in order to characterize P-genotypes of group A rotaviruses detected in children with acute gastroenteritis in Japan

Target P-genotype	Primer	Polarity[a]	Sequence (5' to 3')	Position[b]
Partial VP4 gene				
Consensus	Con3	+	TGGCTTCGCTCATTTATAGACA	11-32
Consensus	Con2	-	ATTTCGGACCATTTATAACC	887-868
Specific P-genotype				
P[8]	1T-1	-	TCTACTTGGATAACGTGC	356-339
P[4]	2T-1	-	CTATTGTTAGAGGTTAGAGTC	494-474
P[6]	3T-1	-	TGTTGATTAGTTGGATTCAA	278-259
P[9]	4T-1	-	TGAGACATGCAATTGGAC	402-385
P[10]	5T-1	-	ATCATAGTTAGTAGTCGG	594-575

a: +, Forward primer; -, Reverse primer.
b: The position numbers of primer pairs correspond to representative strains KU (M21014), L26 (M58292), Gottfried (M33516), K8 (D90260) and 69M (M60600) for P[8], P[4], P[6], P[9] and P[10], respectively.

bp for P[8], P[4], P[6], P[9], and P[10], respectively. Exactly, 1 µl of the first PCR product was mixed in a reagent mixture containing 2.5 µl of 10x Taq DNA polymerase buffer (Promega, Madison, WI), 1 µl of 2.5 mM dNTPs (Roche, Mannheim, Germany), 0.25 µl of each specific primer (20 µM) and 0.05 µl of Taq DNA polymerase (Promega, Madison, WI). MilliQ water was added to make up a total volume of 12.5 µl. The second PCR was performed at 94°C for 3 min, followed by 35 cycles of 94°C 30 s, 45°C 30 s, 72°C 60 s, and a final extension at 72°C for 7 min, and then held at 4°C.

Amplification of polymerase and capid genes of noroviruses

To investigate the possible recombination in norovirus, a partial RNA polymerase region was first amplified using three pairs of specific primers: GV21/Hep171 [39], P290/P289 [82], or Yuri22F/CB2 [83, 84] (Table 4). To confirm the recombinant strain, a long genomic fragment that included the junction between polymerase and capsid genes was generated using forward primers GV21, P290, or Yuri22F and reverse primer G2SKR. The PCR was carried out at 94°C for 3 min,

Table 4. Specific primers used in PCR to amplify polymerase and capsid genes of noroviruses and sapoviruses detected in children with acute gastroenteritis in Japan

Target gene	Virus	Primer	Polarity[a]	Sequence (5' to 3')[b]	Position
Partial polymerase	**Norovirus**				
		GV21	+	GTBGGNGGYCARATGGGNATG	3368-3388[d]
		Hep171	-	CGACGCCATCTTCATTCAC	5099-5081[d]
		P290	+	GATTACTCCAAGTGGGACTCCAC	4295-4317[d]
		P289	-	TGACAATGTAATCATCACCATA	4613-4592[d]
		Yuri22F	+	ATGAATGAGGATGGACCCAT	4232-4251[d]
		CB2	-	GTTTYARCCCGTATTCCTTG	4676-4657[d]
Partial polymerase	**Sapovirus**				
		SR80/2[c]	+	TGGGATTCTACACAAACCC	4366-4384[e]
		SLV5794	-	CGGRCYTCAAAVSTACCBCCCCA	5516-5494[e]
Full-length genome	**Norovirus**				
		5EndGII.6[c]	+	GTGAATGAAGATGGCGTCTAA	1-21[d]
		Pol3464[c]	-	GCTGAGTTTAGGTGCCTTACC	3640-3620[d]
		GV21	+	GTBGGNGGYCARATGGGNATG	3368-3388[d]
		G2SKR	-	CCRCCNGCATRHCCRTTRTACAT	5389-5367[d]
		COG2F	+	CARGARBCNATGTTYAGRTGGATGAG	5002-5028[d]
		TX30SXN	-	GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)30	3'end
Full-length genome	**Sapovirus**	SV1-21[c]	+	GTGATTGGTTAGTATGGCTTC	1-21[e]
		SV2098[c]	-	GGCTGCCATGTAGGATTTCAGAC	2098-2076[e]
		SV1973[c]	+	GGGTTGTCACGCTTAATGGAGACC	1973-1996[e]
		SV3444[c]	-	CCTTCCGCTATTGTCGTCACTGGG	3444-3421[e]
		SV3329[c]	+	GGGGAAACTACCTGGGTTAAC	3329-3349[e]
		SV4525[c]	-	CCCGGACTTGGTGACAAAAA	4525-4506[e]

a: +, Forward primer; -, Reverse primer.
b: Within nucleotide sequence of primers, B=C, G or T; H=A, C or T; N: Any base; R= A or G; K=T or G; W= A or T; Y= C or T
c: New primers were designed in this study.
d, e: The position numbers of primer pairs correspond to representative strains Lordsdale (X86557), and Manchester (X86560), respectively

followed by 35 cycles of 94°C 30 s, 55°C 30 s, 72°C 1-2 min, and a final extension at 72°C for 7 min, and then held at 4°C.

The full-length genome of two representatives of GII.6-polymerase/GII.14-capsid recombinant strains was characterized using three pairs of primers: 5EndGII.6/Pol3464, GV21/G2SKF, and COG2F/TX30SXN [79]. The partial polymerase and full-length capsid region of a new GII.6 variant was amplified using two pairs of primers, GV21/G2SKF and COG2F/TX30SXN. Long-range PCR

(>1,200 bp) was amplified with High Fidelity DNA polymerase KOD-Plus (Toyobo, Osaka, Japan). The PCR was carried out with 2 µl of cDNA, which has been reverse transcribed using TX30SXN primer, in the reagent mixture containing 2.5 µl of 10x buffer for KOD-Plus, 1 µl of 2 mM dNTPs, 1 µl of 25 mM $MgSO_4$, 0.75 µl of each specific primer (20 µM) and 0.5 µl of KOD Plus enzyme. MilliQ water was added to make up a total volume of 25 µl. The PCR was performed at 94°C for 2 min, followed by 35 cycles of 98°C 20 s, 50°C or 55°C 30 s, 68°C 60 s/2kb, and a final extension at 68°C for 10 min, and then held at 4°C.

Amplification of polymerase and capid genes of sapoviruses

To investigate the possible recombination in sapovirus, a long genomic fragment that included the junction between polymerase and capsid genes was generated using SR80/2 and SLV5794, which generated a 1,151-bp product (Table 4). The PCR was performed at 94^{o}C for 3 min, followed by 35 cycles of 94°C 30 s, 55°C 30 s, 72°C 1.30 min, and a final extension at 72°C for 7 min, and then held at 4°C.

The full-length genome of a GII-polymerase/GIV-capsid recombinant strain, 8206/Maizuru/08, was characterized using four pairs of primers: SV1-21/SV2098, SV1973/SV3444, SV3329/SV4525, and SV5317/TX30SXN (Table 4). The full-length genome of GI/1 sapovirus strains was amplified using SLV5371 and TX30SXN primers. Long-range PCR (>1,200 bp) was amplified with High Fidelity DNA polymerase KOD-Plus (Toyobo, Osaka, Japan). The PCR was carried out with 2 µl of cDNA, which has been reverse transcribed using TX30SXN primer, in the reagent mixture containing 2.5 µl of 10x buffer for KOD-Plus, 1 µl of 2 mM dNTPs, 1 µl of 25 mM $MgSO_4$, 0.75 µl of each specific primer (20 µM) and 0.5 µl of KOD Plus enzyme. MilliQ water was added to make up a total volume of 25 µl. The PCR

was performed at 94°C for 2 min, followed by 35 cycles of 98°C 20 s, 50°C or 55°C 30 s, 68°C 60 s/2kb, and a final extension at 68°C for 10 min, and then held at 4°C.

Electrophoresis

The PCR products were analyzed by electrophoresis on a 1.5% agarose gel, followed by staining with SYBR Safe (Invirtogen) for 20 min and then visualized under ultraviolet light. The result was recorded by photography.

Nucleotide sequencing and phylogenetic analyses

The nucleotide sequences of PCR products were determined with the Big-Dye terminator cycle sequencing kit and an ABI Prism 310 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA). Nucleotide or amino acid sequence alignments were performed with MEGA 4 [85]. Phylogenetic trees with 1,000 bootstrap replicates were constructed by the neighbor-joining method, and the genetic distances were calculated according to the Kimura two-parameter method [86]. SimPlot software (Version 1.3) [44] was used to compare the recombinant norovirus and sapovirus sequences.

Protein structure modeling

Homology models of GII.14 and GII.6 capsid P domains were constructed using the X-ray crystal structure of the P domain of GII.4 virus VA387 in complex with A and B trisaccharides, respectively [31] (Protein Data Bank accession numbers 2OBS and 2OBT [http://www.rcsb.org/pdb/home/home.do]) as templates. Briefly, alignments of the template and targets (GII.14 and GII.6) P-domain regions were generated and demonstrated sequencing identity greater than 55%, indicating that the

GII.4 was an appropriate template for generating GII.14 and GII.6 homology models. The GII.14 and GII.6 homology models were then predicted using the program EasyModeller, version 2.0 [87], which is a graphical interface to MODELLER [88] using the automodel class [89, 90]. The 3D structures were thermodynamically optimized by energy minimization using steepest descent (200 cycles) with the GROMOS 43B1 force field of the Swiss-Pdb Viewer program (http://www.expasy.org/spdbv/) [91]. A physically unacceptable local structure of the optimized 3D models was further refined on the basis of Ramachandran plot evaluation by using EasyModeller. The Protein Data Bank files generated for the GII.14 and GII.6 P domains were visualized on the YASARA view program (http://www.yasara.org/) [92]. Dimers of obtained models were generated by superimposing two predicted monomers onto a dimeric template of the VA387 P domain [31] by MUSTANG [93] in YASARA view program. The final homology models were further evaluated using PROCHECK [94], WHATCHECK [95], VERIFY 3D [96], ERRAT [97], and PROVE [98] programs (http://nihserver.mbi.ucla.edu/SAVES/).

CHAPTER I

DIVERSITY OF DIARRHEA VIRUS INFECTIONS

In the study, the existence of six diarrhea viruses co-circulating in children with acute gastroenteritis was shown in Table 5. None of group B rotavirus was identified. Overall, diarrhea viruses were detected in 54.5% (652 of 1,195) fecal specimens tested. Of these, monoinfection and mixed infection accounted for 52.5% (579 of 1,195) and 2% (22 of 1,195), respectively. Mixed infection between group A rotavirus and the other viral agents was dominant (16 of 22). The findings suggested that from acute gastroenteritis cases in children from six localities in Japan, diarrhea viruses were responsible for 54.5% and 45.5% caused by other etiologic agents. Of note, norovirus was the most predominant with the detection rate of 30.1% (360 of 1,195), followed by group A rotavirus with 15.1% (181 of 1,195), sapovirus with 2.9% (33 of of 1,195), and other viruses with lower frequencies.

The clinical features of diarrhea diseases associated with common viruses were summarized in Table 6. However, the clinical manifestation of viral infections in this study could not be demonstrated fully because only medical records were obtained from patients visiting a clinic in Maizuru. Clinical signs included diarrhea, vomiting, and fever. Diarrhea was the dominating sign seen in 90% for group A rotavirus infection, 88% for norovirus infection, and 65% for sapovirus infection. Vomiting and fever were more common in group A rotavirus and norovirus infections than sapovirus infections.

Table 5. Distribution of viral infection in children with acute gastroenteritis visiting pediatric clinics encompassing six localities (Sapporo, Saga, Tokyo, Osaka, Maizuru and Shizuoka) in Japan from July 2007 to June 2009

Year	No. of specimen tested	No. of viral positive (%)	No. of monoinfection (%)							No. of mixed infection (%)							
			RV			NoV	SaV	Ade	AstV	Ade SaV	SaV RA	Ade NoV	RA NoV	Ade RA	SaV NoV	SaV AstV	AstV RA
			A	B	C												
2007-2008[a]	477	253 (53)	86 (18)		4 (0.8)	120 (25.2)	16 (3.4)	17 (3.6)	1 (0.2)	1 (0.2)	2 (0.4)	2 (0.4)	3 (0.6)	1 (0.2)			
2008-2009[a]	531	241 (45.3)	55 (10.4)		4 (0.8)	136 (25.6)	15 (2.8)	12 (2.3)	7 (1.3)				8 (1.5)		2 (0.4)	1 (0.2)	1 (0.2)
2008-2009[b]	187	158 (84.5)	40 (21.4)		1 (0.5)	104 (55.6)	2 (1.1)	2 (1.1)	8 (4.3)								1 (0.5)
Total	1,195	652 (54.5)	181 (15.1)		9 (0.4)	360 (30.1)	33 (2.9)	31 (2.7)	16 (1.3)	1 (0.1)	2 (0.2)	2 (0.2)	11 (0.9)	1 (0.1)	2 (0.2)	1 (0.1)	2 (0.2)

No., Number; RV, Rotavirus; NoV, Norovirus; SaV, Sapovirus; Ade, Adenovirus; AstV, Astrovirus; RA, Group A rotavirus; %, Refer to total number of specimens tested

[a] Results based on specimens obtained from five localities (Sapporo, Saga, Tokyo, Osaka, and Maizuru).
[b] Results based on specimens obtained from Shizuoka.

Table 6. Descriptions of clinical signs of diarrhea diseases associated with different types of virus infections in the surveillance in Maizuru during 2007-2009.

Signs and symptoms	RA (n=96)	NoV (n=87)	SaV (n=19)
Diarrhea	90%	88%	65%
Vomiting	48.8%	35.6%	5.3%
Fever (≥38°C)	46.8%	31%	26.3%

RA, Group A rotavirus; NoV, Norovirus; SaV, Sapovirus

DETECTION AND GENETIC CHARACTERIZATION OF NOROVIRUS IN FIVE LOCATIONS OF JAPAN DURING 2007 to 2009

Molecular epidemiology of norovirus infection

A total of 1,008 fecal specimens collected from infants and children with acute gastroenteritis from five different regions (Sapporo, Saga, Tokyo, Osaka, and Maizuru) of Japan during 2007-2009 were examined for norovirus by multiple RT-PCR and semi-nested PCR. Of the 1,008 specimens tested, norovirus was detected in 271 specimens (26.9%), in which 100 (9.9%) were positive by semi-nested PCR. Of the 271-positive specimens, 7 (2.6%) were positive for norovirus GI and 264 (97.4%) for norovirus GII. Noroviruses were detected for 125 of 477 (26.2%) specimens in 2007-2009 and for 146 of 531 (27.5%) in 2008-2009. The highest prevalence of norovirus was found in children aged 12-23 months (117, 43.2%) and lowest in infants aged <6 months (7, 2.6%) (Fig. 5A). It was also found that children under 3 years had a high rate of norovirus infection (232 cases, 85.6%). Infections increased slightly from October to February in both 2007-2008 (94, 75.2%) and 2008-2009 (91, 62.3%) (Fig. 5B).

Phylogenetic analysis of norovirus sequences

For the 271-positive specimens, noroviruses were sequenced and analyzed by phylogenetic analysis based on the partial capsid region. Table 7 shows the distribution of norovirus GI and GII genotypes. All 7 GI sequences (100%), which were detected only in 2007-2008 clustered with Chiba407 known as GI.4 genotype (Fig. 6). Most of the norovirus GII sequences in 2007-2008 belonged to GII.4 (77.9%),

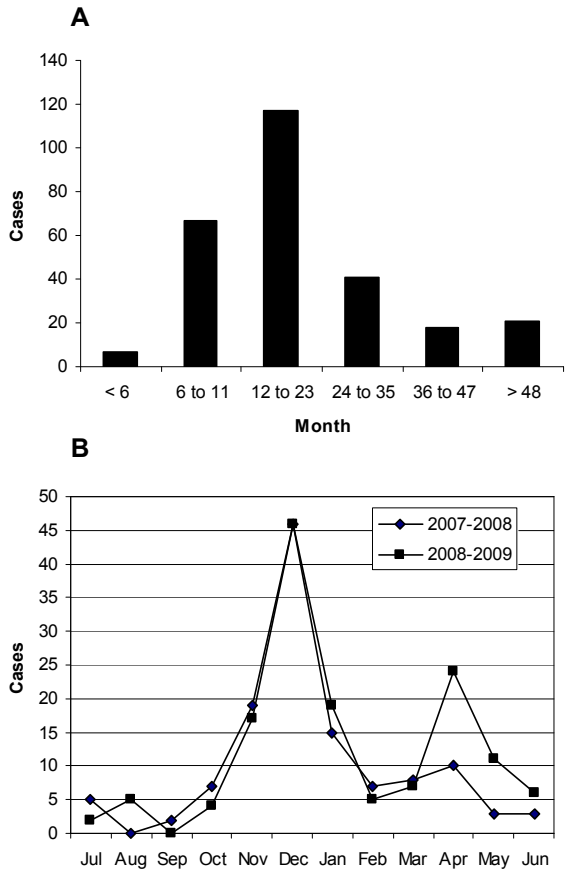


Figure 5. Age-related (A) and seasonal distributions (B) of norovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009.

Table 7. Distribution of GI and GII norovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009

Year	No. of specimens tested	No. (%) positive norovirus	GI (%[a])	GII (%[b])							
			GI.4	GII.1	GII.2	GII.3	GII.4	GII.6	GII.7	GII.12	GII.14
2007-2008	477	125 (26.2)	7 (100)	0	0	6 (5.1)	92 (77.9)	6 (5.1)	0	0	14 (11.9)
2008-2009	531	146 (27.5)	0	1 (0.7)	6 (4.1)	9 (6.1)	97 (66.4)	26 (17.8)	1 (0.7)	3 (2.1)	3 (2.1)

[a] Refers to number of GI-positive specimens.
[b] Refers to number of GII-positive specimens.

followed by GII.14 (11.9%), and GII.3 and GII.6 (5.1% each). In 2008-2009, norovirus sequences were classified into 8 distinct genotypes (GII.1, GII.2, GII.3, GII.4, GII.6, GII.7, GII12, and GII.14) with GII.4 (66.4%) being the most prevalent. GII.14 suddenly decreased from 11.9% in 2007-2008 to 2.1% in 2008-2009, while GII.6 increased sharply from 5.1% to 17.8%. A phylogenetic tree revealed that all of the GII.4 sequences identified in this study clustered with the recent GII.4/2006b variant in both seasons (Fig. 7).

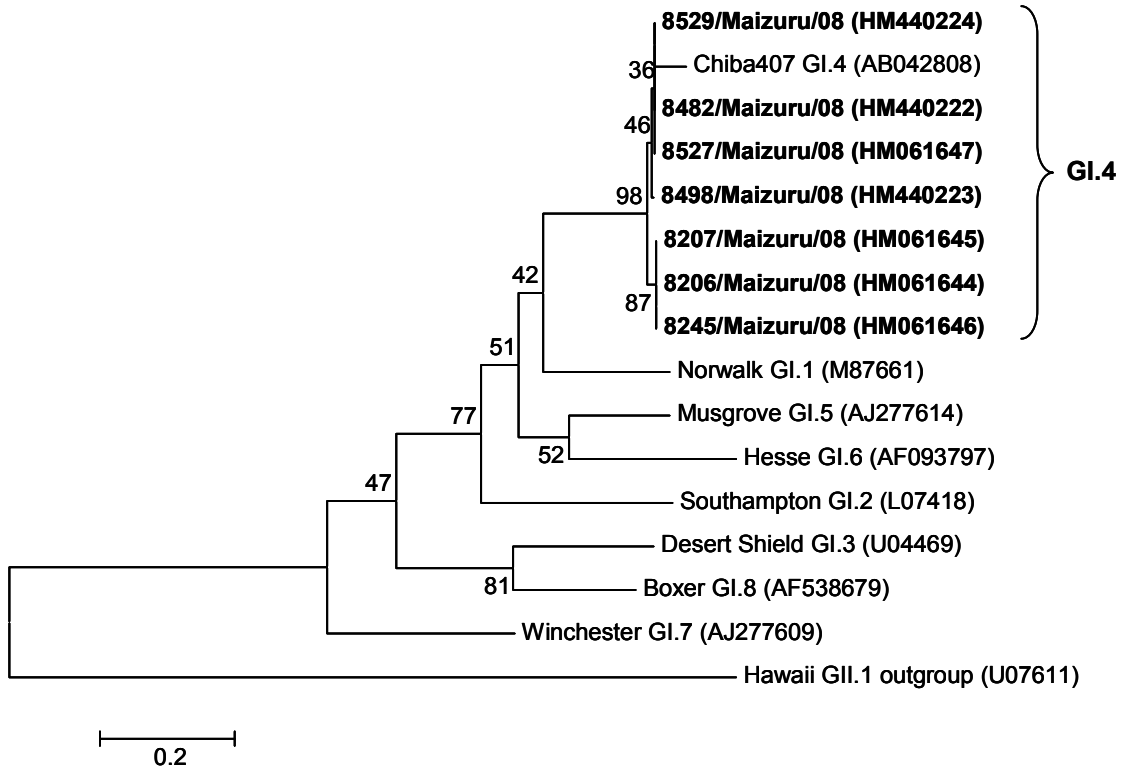


Figure 6. Phylogenetic analysis of the partial capsid sequences of norovirus GI detected in 2007-2008. All of the detected GI sequences are represented in boldface. Genbank accession numbers are given in parentheses.

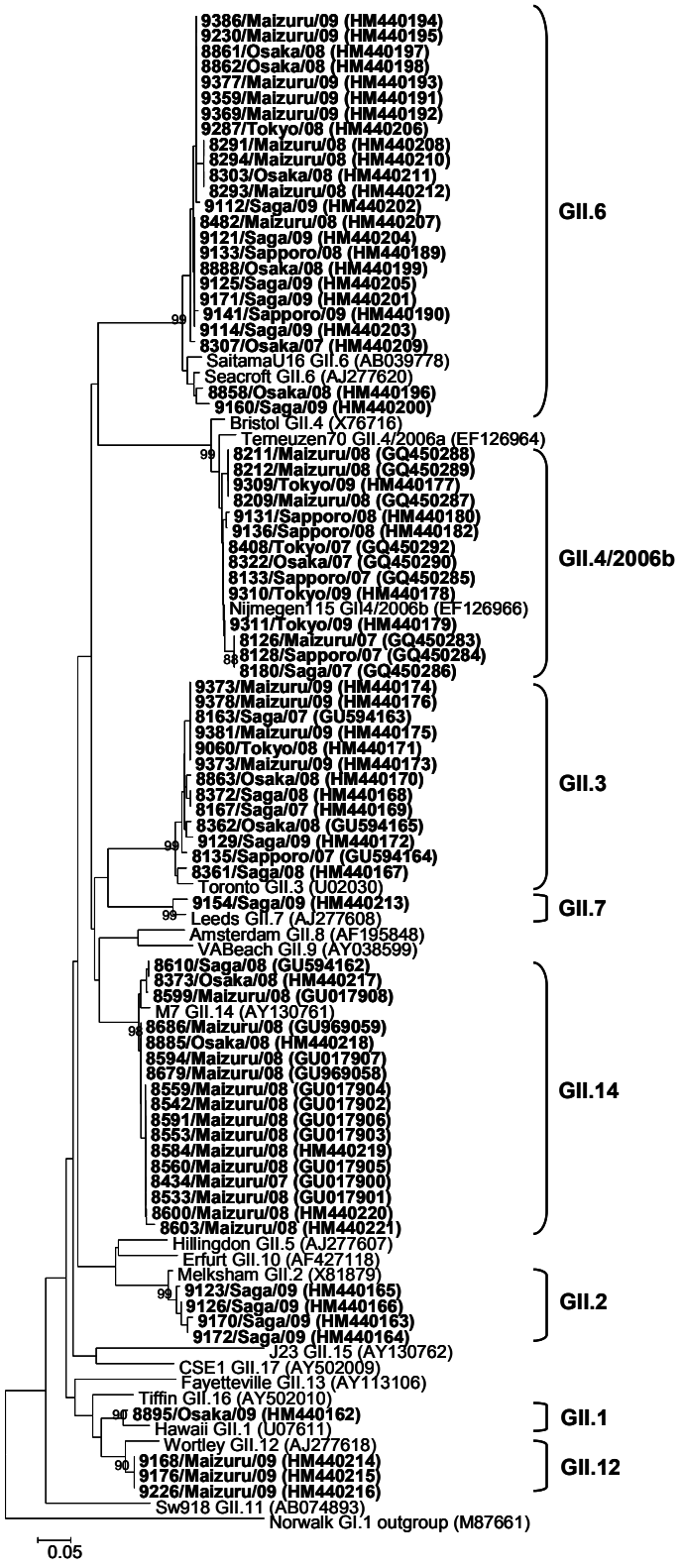


Figure 7. Phylogenetic analysis of the partial capsid sequences of norovirus GII detected in 2007-2009. Representative GII sequences are represented in boldface. Genbank accession numbers are given in parentheses. Only bootstrap values >70 are shown.

Occurrence of recombinant noroviruses

To investigate possible recombination noroviruses, partial polymerase regions of 3 GII.3 and 12 GII.14 norovirus strains were sequenced and analyzed. Of the 3 GII.3-capsid strains, one strain (8163/Saga/07) was classified into Lordsdale cluster (known as GII.4) and the other two strains (8135/Sapporo/07 and 8362/Osaka/08) were grouped with SydneyC14, Pont de Roide673, and VannesL169 cluster (known as GII.b) when polymerase-based grouping was performed. Interestingly, the polymerase sequences of the 12 GII.14 strains including 9 and 3 strains detected in 2007-2008 and 2008-2009, respectively, were classified into Saitama U3 and GIFU99 cluster (known as GII.6), indicating that all GII.14 strains isolated in both 2007-2008 and 2008-2009 were GII.6/GII.14 recombinant strains (Fig. 8). 8610/Saga/08 strain was selected as representative for analyzing the full-length genome sequence (6.7 kb) in order to localize the potential recombination site and to understand a possible recombination mechanism of this strain. SimPlot analysis using Saitama U3 GII.6 as a parental strain showed a sudden drop in nucleotide similarity after ORF1, suggesting the ORF1/ORF2 overlap as the potential cross-over site (Fig. 9). This analysis was performed based on a single parental strain (Saitama U3, GII.6) because the polymerase region of the M7 GII.14 reference strain (GenBank accession no. AY130761) was not available in GenBank database.

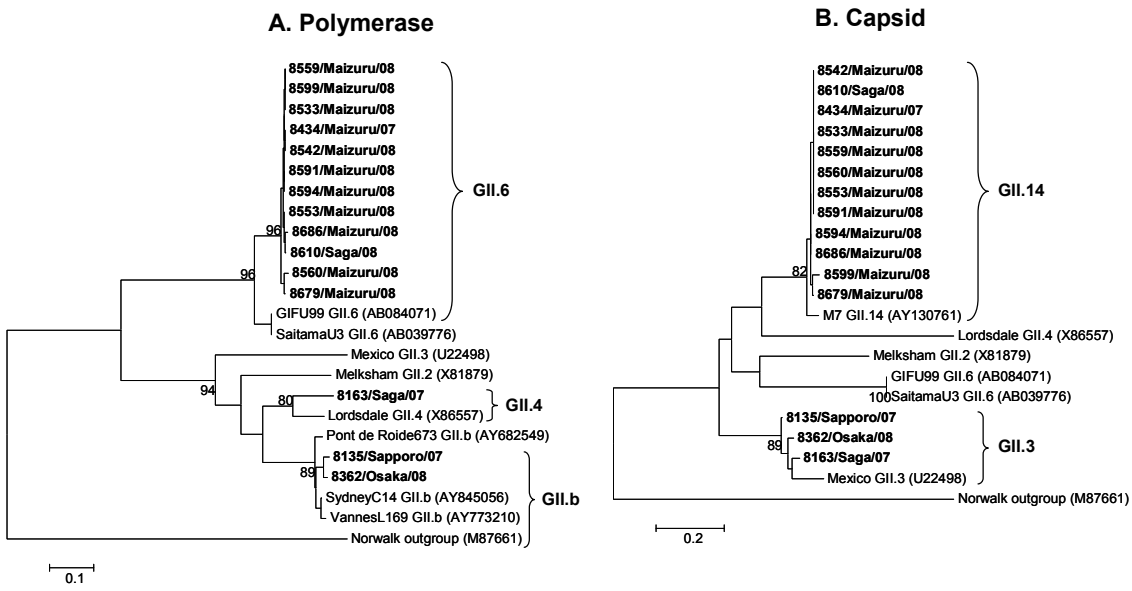


Figure 8. Changing genotypes of norovirus on the basis of phylogenetic trees of nucleotide sequences. The trees were constructed from partial nucleotide sequences of polymerase (A) and capsid (B) regions.

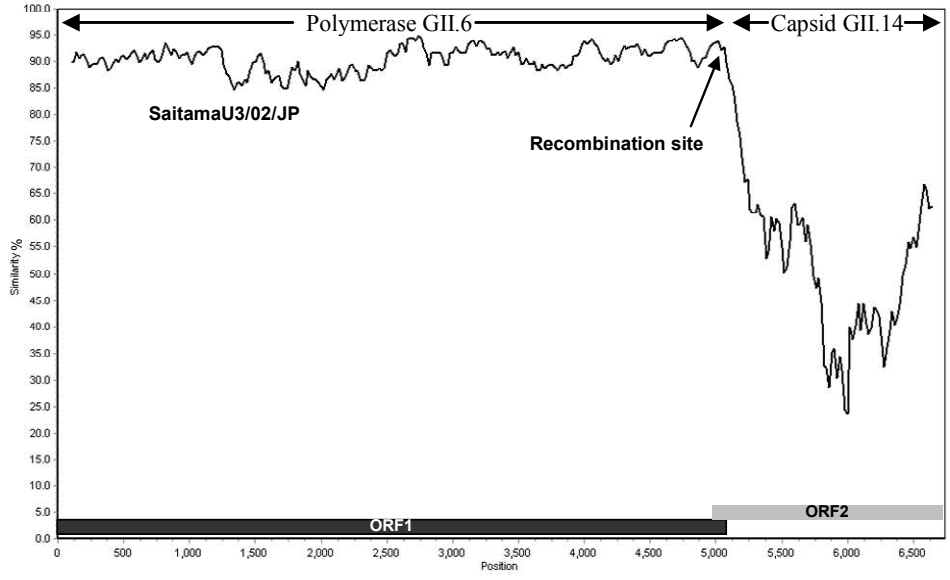


Figure 9. Identification of 8610/Saga/08 recombinant strain. SimPlot analysis was performed using the 6.7-kb fragment covering the ORF1 and ORF2 with the putative parental strain, SaitamaU3 (GII.6). A window size of 100 bp with an increment of 20 bp was shown. The recombinant site is suspected to be located within the ORF1/2 overlap, as shown by the arrow.

Genetic analysis of capsid protein of GII.14 noroviruses

Little information is known regarding the genetic evolution in capsid protein of GII.14 genotype. Therefore, the complete capsid genes of the GII.14 were sequenced and analyzed. The amino acid sequences of the capsid genes were compared for 11 GII.14 strains, comprising 9 GII.14 strains detected in this study, the M7 prototype strain, which was detected in 1999, and the Shanxi/50106 strain, which was reported recently in China in 2006-2007 [99]. Sequence comparison showed ≥94% nucleotide and ≥98% amino acid identities among these M7-like strains. In addition, ten amino acid changes were observed in the P domain of the capsid protein as shown in Fig. 10. These amino acids included 277 (T→I), 287 (R→K), 297 (Q→E), 341 (R→I), and 373 (D→N) in the P2 subdomain and 433 (H→Y), 445 (I→M), 454 (I→V), 483 (L→F), and 495 (V→I) in the P1-2 subdomain.

To obtain structural insights into the roles of the amino acid substitutions specific to GII.14 capsid protein, a 3D structure of the most variable P domain of a dimer of the GII.14 strain, 8610/Saga/08, was constructed by homology modeling. The 8610/Saga/08 had 55% amino acid sequence identity in the P domain with the template VA387/98/US GII.4 (2OBS). The homology model was generated by EasyModeller 2.0 program and verified with structural validation programs. Results of the model validation were shown in APPENDIX (Table 11, Fig. 28-29). A comparison of the positions of the 10 mutations to the predicted 3D structure showed that 5 residues (277, 287, 297, 341, and 373) and 2 residues (433 and 445) were predicted to be located at the surface-exposed P2 and P1-2 subdomains, respectively (Fig. 11). All GII.14 strains detected in this study contained an altered RGD-like motif (RGT→KGT) in position 287-289. This motif has been known to be recognized as a receptor binding site of GII noroviruses [30].

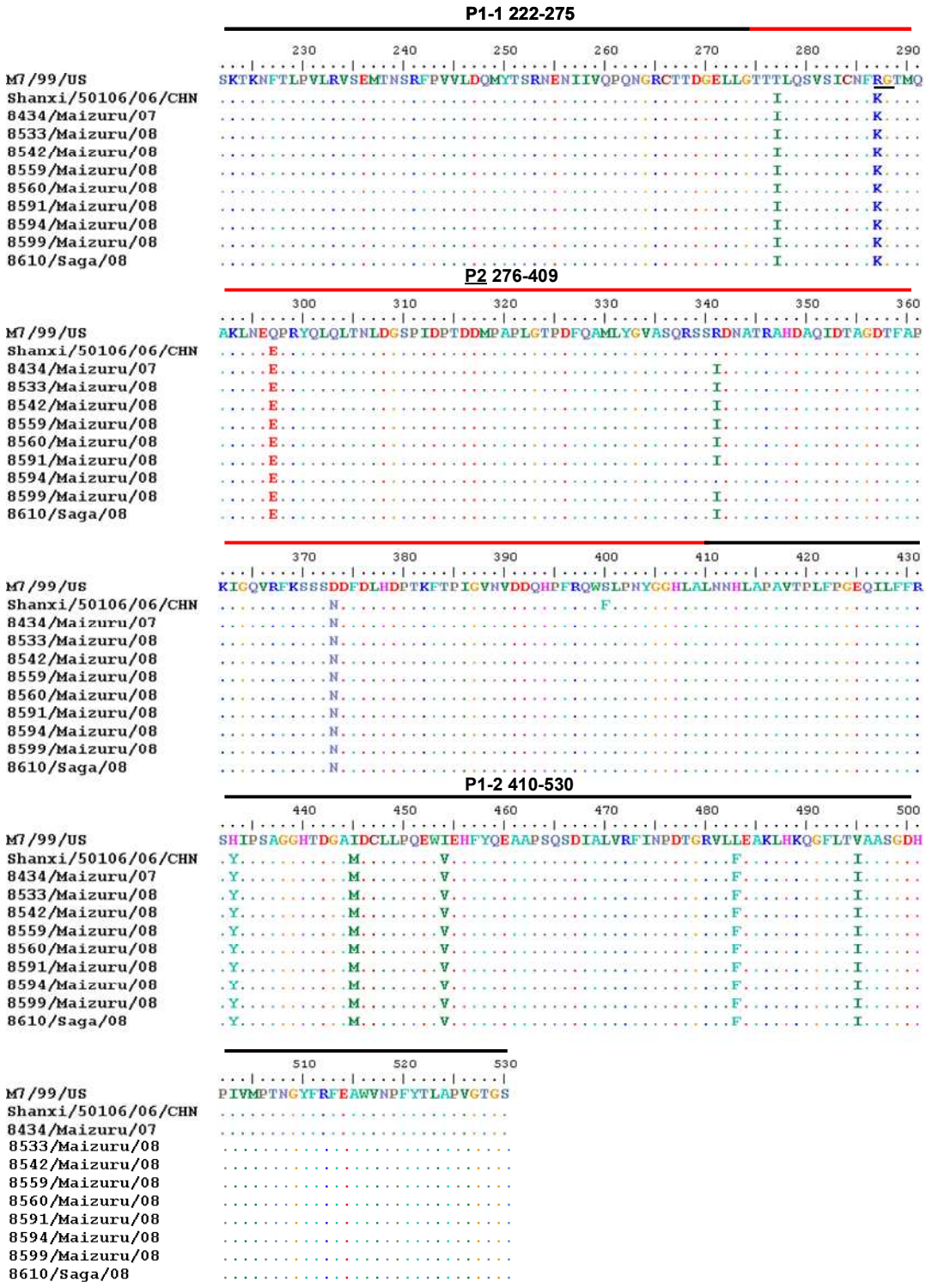


Figure 10. Amino acid alignment of the P domain of GII.14 strains showing mutations in P2 and P1-2 subdomains. Underlined residues indicate the RGT motif that mutated to KGT motif in the GII.14 detected in this study.

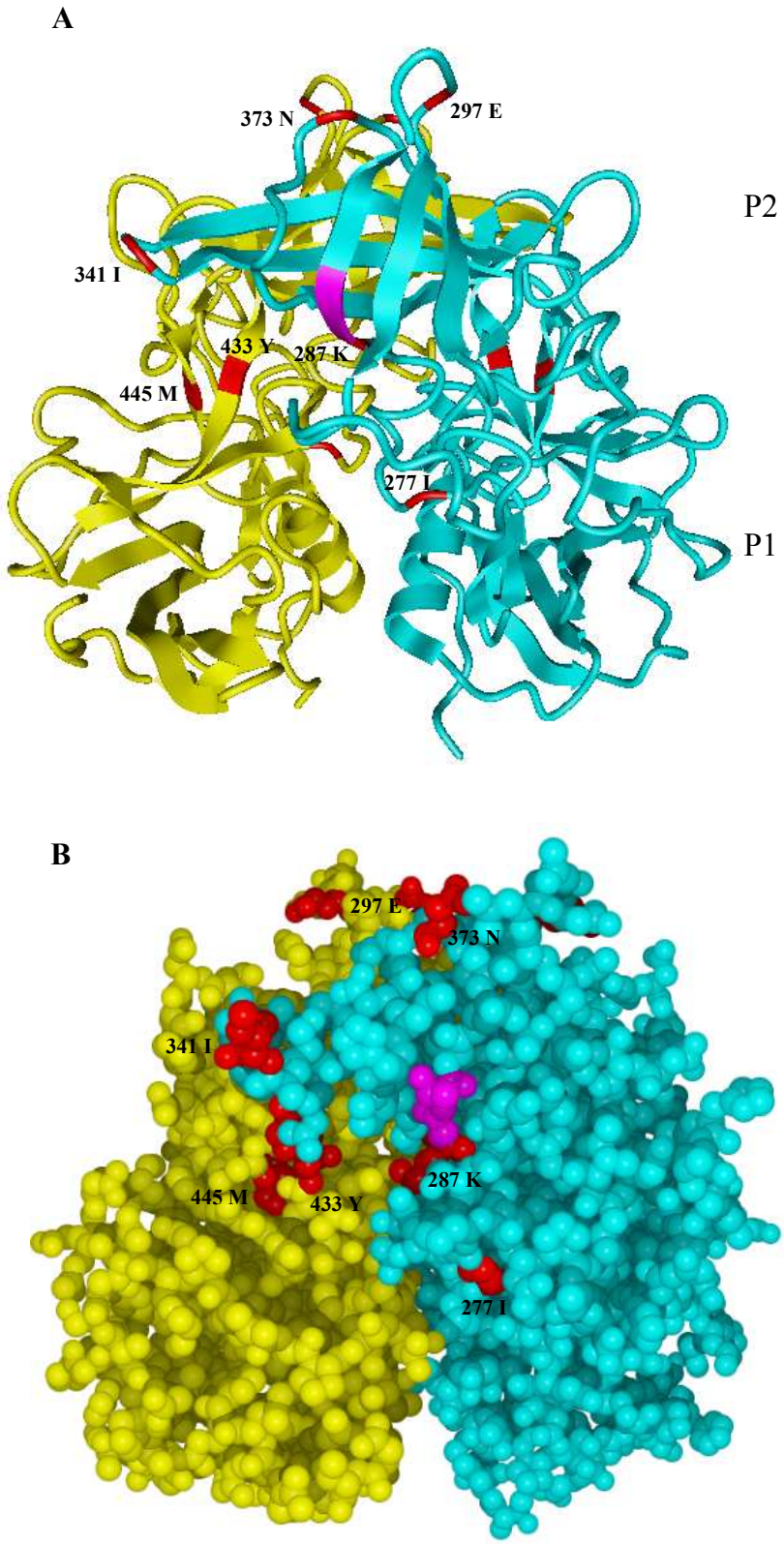


Figure 11. The 3D structure of a dimer of the GII.14 P domain, as predicted by EasyModeller, showing the location of the amino acid changes. The predicted structures are presented in ribbon (A) and ball (B). Monomer A is shown in yellow, while the monomer B is shown in cyan. Locations of the changes that occurred in the surface-exposed P2 (277, 287, 297, 341, and 373) and P1-2 (433 and 445) subdomains are shown in red. Additional KGT motif in monomer B (amino acid 287-289, magenta) is also shown.

DISCUSSION

In the present study, norovirus was recognized as the leading pathogen of illness in 26.6% of the children with acute gastroenteritis. Of these, 16.7% were positive by multiplex RT-PCR. This detection rate was in line with our previous epidemiological studies, which demonstrated that norovirus was detected in 8.5-16.5% by multiplex RT-PCR [67-70]. However, this study showed that semi-nested PCR assay enhanced sensitivity up to 9.9% for the detection of norovirus GII in negative samples that failed to be detected by multiplex RT-PCR. Recently, Kittigul and co-workers reported that semi-nested PCR (44.7%) provides seven times higher sensitivity than multiplex RT-PCR (6.5%) [100]. Generally, multiplex RT-PCR assay has limitation due to the fact that the presence of several pairs of primers in a PCR increases the probabilities of mispairing and nonspecific amplification, particularly the formation of primer-dimers. On this basis, therefore, it is suggested that a sensitive semi-nested PCR should be performed to avoid false-negative results that may occur during the multiplex RT-PCR step.

The higher prevalence of GII (97.4%) over GI (2.6%) is in accordance with previously published data on norovirus from Japan and elsewhere [34, 67-69, 101], confirming that norovirus GII is the dominant group associated with acute gastroenteritis in Japan. Several studies have found that norovirus infection occurs frequently in young children under 3 years of age. In agreement with this, the results showed a marked trend toward higher rates of infection in children under 3 years of age (85.6%) versus children over 3 years of age (14.4%), with children between 12 to 23 months of age (43.2%) being the most frequently affected. Quite possible, 12-23-month-old children might lack antibody protection to norovirus, whereas by the time they have reached the age of 24 months they have begun to acquire viral immunity.

Similar to many reports, these data showed that norovirus infection is prevalent in winter months, peaking in December (Fig. 6B) [67, 68, 102, 103], although norovirus infections with spring or summer months have been observed [83, 101, 104, 105].

According to other reports, norovirus belonging to GII/4 represented the highest detection in acute gastroenteritis among children not only in Japan but also in many other countries who run norovirus surveillance [69, 100, 102, 106]. These findings were similar with my study in which GII.4 dominated over any other genotypes, accounting for 66.4-77.9% (Table 7). During the past 15 years, five pandemics have been occurred by new GII.4 variants: 1995/1996, U5-95_US strain; 2002, Farmington Hills; 2004, Hunter; 2006, 2006a virus; and 2007, 2006b virus) [39-41]. The GII.4/2006b variant was first described in Europe and thereafter has been the most commonly identified variant worldwide [34, 38, 102, 106-108]. In the study, GII.4/2006b variant was responsible for 100% among the detected GII.4 strains, which supports that this variant is still prevalent in Japan. A recent study showed that the predominant GII.4 strains had a higher mutation rate and rate of evolution compared to the less frequently detected GII.b, GII.3 and GII.7 strains. And, this study suggests mutation rate in combination with a high replication rate are key determinates in epidimiological fitness. Therefore, it would seem that GII.4 viruses, as opposed to GII.b/GII.3 and GII.7 viruses, have reached a balance in their replication rate and mutaion rate that is better suited to viral adaptaion [109]. However, the presence of other genotypes such as GII.1, GII.2, GII.3, GII.6, GII.7, GII.12, and GII.14 was identified in this study, indicating the continued worldwide emergence of these unusual norovirus strains.

Recombination is also a major determinant of viral virulence, being implicated in the emergence of new viral strains [43]. This study found that two GII.b/GII.3

recombinant strains were similar to those previously reported in Japan and India [68, 110]. One GII.4/GII.3 recombinant strain detected in this study showed more than 97% nucleotide identity with those detected in China, South Korea, and Japan [67, 99, 110]. GII.6/GII.14 recombinant strains were first identified from stool sample of a Chinese patient [99] and then from sewage water in Japan [111]. However, this kind of recombinant virus seems to be very rare as shown by a relatively low detection rate in those two studies. Remarkably, in the present study, GII.6/GII.14 recombinant strains emerged, for the first time in Japan, and became the second leading strain (11.9%) after the GII.4/2006b variant in 2007-2008. Moreover, these recombinant strains were mainly detected in Maizuru (13 of 14 cases). However, this virus has decreased rapidly to 2.1% a year later (3 cases: one in Osaka and two in Maizuru). This sudden disappearance of strain might indicate that the virus appeared at the time that pediatric population lacked antibodies against this strain, and the virus disappeared as the population began to acquire viral immunity. It is also possible that such recombinant virus could not sustain the environmental conditions that contribute to its ability to cause illness.

All of the norovirus recombinants from GI, GII and GIII, with the exception of one norovirus GII recombinant, had a crossover point either within or close to the ORF1/2 overlap [112]. In agreement with other studies, the recombinant strains identified in my study had a recombination point at ORF1/2 junction. It has been reported that the ORF1/2 overlap includes the subgenomic promoter [113] and consequently has a stem–loop structure that is 100 % conserved within each genogroup [114]. This is consistent with the proposed model [114], which suggests that recombination occurs when the polymerase switches templates mid-transcription due to complex secondary structure at the start of ORF2. Consequently, polymerases

with poor processivity would switch templates at a higher frequency than other RdRps. The ability of polymerases to switch templates at the start of ORF2 is advantageous because it can help viruses escape evolutionary bottlenecks [115, 116]. This is because ORF2 encodes the capsid protein, VP1, which contains the antigenic regions. Therefore, viruses that are able to swap their capsid coat are able to escape immune responses and possible viral extinction.

A recent bioinformatics study of GII.2 strains suggests that this genotype may be evolving new strains in a time-ordered manner over a 12-year period. It appears that the novel strains are defined by targeted variation primarily occurred in the surface-exposed P2 subdomain, suggesting that heterogeneity in the GII.2 strains is mediated by an immune response-driven evolution [33]. In addition, similar studies with the GII.3 and GII.4 genotypes show evidence of immune response-driven evolution [34, 117]. Supporting this conclusion, the comparison of the capsid sequences between the present GII.14 strains and the prototype M7 strain revealed 10 amino acid substitutions in the P2 and P1 subdomains during the past 10 years. Furthermore, the 3D structure of the GII.14 strains shows that most of the changes were located on the surface-exposed P2 subdomain. This can be suggested that the amino acid changes in the P2 subdomain among these GII.14 strains might be suggestive of immune response.

Noroviruses have been found to recognize the histo-blood group antigens (HBAGs) such as A, B, and H blood group antigens as receptors [34, 35, 42]. Earlier studies have suggested that RGD-like motif (located at the beginning of the P2 subdomain) might be of structural importance for norovirus binding to HBGAs [30]. In the present study, arginine residue of RGT motif (amino acid 287-289, magenta in Fig. 10) that corresponds to the RGD-like motif of other noroviruses was substituted

to a lysine residue in all GII.14 strains. To determine if this mutation leads to changes in host binding specificity that allows an expansion of host range or penetrance into a previously naïve population, further experiments such as VLP-saliva binding assay is needed.

CONCLUSION

The data showed that GII.4/2006b norovirus emerged as the major pathogen causing acute gastroenteritis among infants and children less than 3 years of age during 2007-2009. The seasonality for norovirus infections peaks in cold months (October to February). This is also the first report of an emergence of GII.6/GII.14 recombinant strain in Maizuru in 2007-2008. Furthermore, the homology model of GII.14 P domain showed several mutations in the surface exposed P2 subdomain. This indicates evidence for immune response-driven evolution that may lead the virus to escape from pre-exiting immunity. Understanding how noroviruses evolve is necessary for finding more effective ways to control this disease and particularly its impact in healthcare settings.

CHAPTER II.

DETECTION AND GENETIC CHARACTERIZATION OF ROTAVIRUS IN FIVE LOCATIONS OF JAPAN DURING 2007 to 2009

Molecular epidemiology of group A rotavirus infection

Rotavirus was detected in 156 out of 1,008 (15.5%) specimens. Rotavirus infection in 2007-2008 (19.3%) was more prevalent than those in 2008-2009 (12.1%). Detection rates were different between the five localities: 8.5% in Tokyo, 10.5% in Saga, 11% in Osaka, 18.7% in Maizuru, and 28.8% in Sapporo. Most of the rotavirus infections were in a 12-23-month-old group (81 of 156, 51.9%), followed by a 6-11-month-old group (27 of 156, 17.3%) (Fig. 12A). Rotavirus was detected continuously in a 8-month period lasting from December through July. However, infections increased most rapidly from March to May with the highest peak in April (77 of 156, 49.4%) (Fig. 12B).

Distribution of G- and P-genotypes of rotaviruses

Four different rotavirus G-genotypes, G1, G2, G3, and G9, were detected during the study period. The majority of G genotypes were G1 (62.8%), followed by G3 (21.8%), G9 (14.7%), and G2 (0.7%) as shown in Table 8. Two common P genotypes, P[4] and P[8], were identified. Rotavirus P[8] strain (99.3%) was the most prevalent. The most frequent G-P combination was G1P[8] and represented 62.8%, followed by G3P[8] (21.8%), G9P[8] (14.7%), and G2P[4] (0.7%). The data showed that the trend of changes in the distribution of G- and P-genotypes between 2007-2008 and 2008-2009 was almost the same.

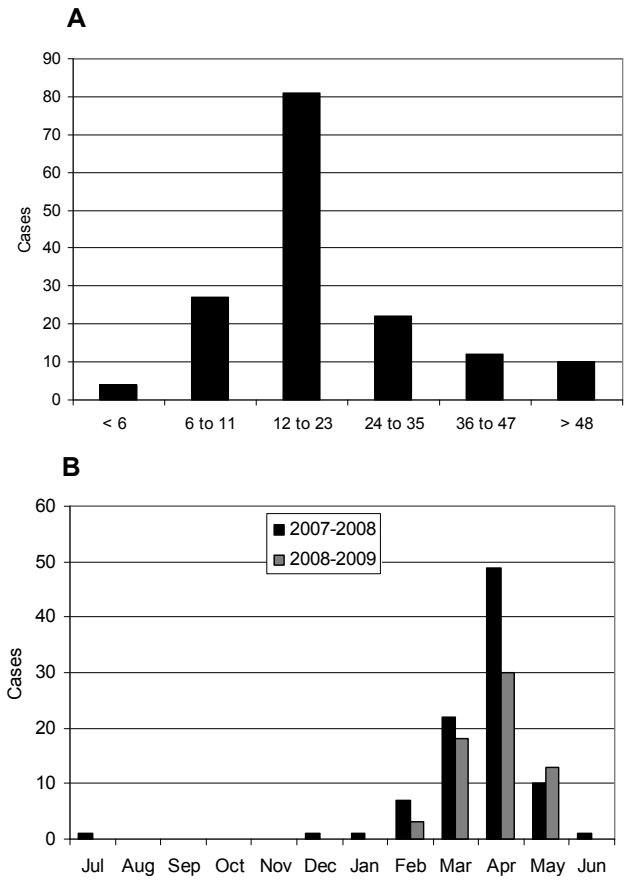


Figure 12. Age-related (A) and seasonal distributions (B) of group A rotavirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009.

Table 8. Distribution of G- and P-genotype combinations of rotavirus infection in children in five cities of Japan during 2007-2009

Year	No. of specimens tested	No. (%) positive rotavirus	Genotype (%[a])			
			G1P[8]	G2P[4]	G3P[8]	G9P[8]
2007-2008	477	92 (19.3)	56 (60.9)	0	20 (21.7)	16 (17.4)
2008-2009	531	64 (12.1)	42 (65.6)	1 (1.6)	14 (21.9)	7 (10.9)
Total	1,008	156 (15.5)	98 (62.8)	1 (0.7)	34 (21.8)	23 (14.7)

[a] Refers to number of rotavirus-positive specimens in each year.

Phylogenetic analyses of G1, G3, and G9 rotavirus sequences

In order to characterize rotaviruses that circulated during 2007-2009, the full-length VP7 genes of 15 representative G1 strains were determined. The identity was very high (99-100% at nucleotide level) among the 15 Japanese G1 strains, represented by JP-9453/08-09. The phylogenetic analysis clearly showed that there were five lineages (1-5) of G1 strains detected worldwide (Fig. 13A). The Japanese G1 strains analyzed in the present study were located in lineage 5e of lineage 5 together with some other strains reported in China, Japan, USA, and Bangladesh. JP-9453/08-09 exhibited more than 99% VP7 gene sequence identities to SY5, AS4, and E050 strains circulating during 2006-2008 in China [78]. In contrast, JP-9453/08-09 showed lower identities (96.7-98.3%) to strains in lineage 5a through 5d and 92% identities to Wa USA in lineage 1.

Twelve G3 rotaviruses, represented by JP-9313/08-09, exhibited high sequence identities in the VP7 gene (99-100%) and were found in lineage 3d of lineage 3, which includes many strains from Asian countries including Japan, China, Russia, and Viet Nam (Fig. 13B). The VP7 genes of the 12 Japanese G3 strains were identical (100% at amino acid level) to the "new variant G3" strain 5019, which previously emerged in Japan during 2003-2004 [79]. Between the JP-9313/08-09 and prototype G3 strain YO in lineage 3a, there were 95% identities at nucleotide level.

Six G9 rotaviruses were almost identical (>98% at nucleotide and amino acid levels) among each other and were clustered into two separate branches in lineage 3, which comprises globally spreading G9 strains (Fig. 13C). Comparison of the VP7 amino acid sequences between the 2007-2009 strains and 2008-2009 strains revealed the amino acid differences in four positions (position 8, 46, 65, and 268) of the non-antigenic region (data not shown).

A

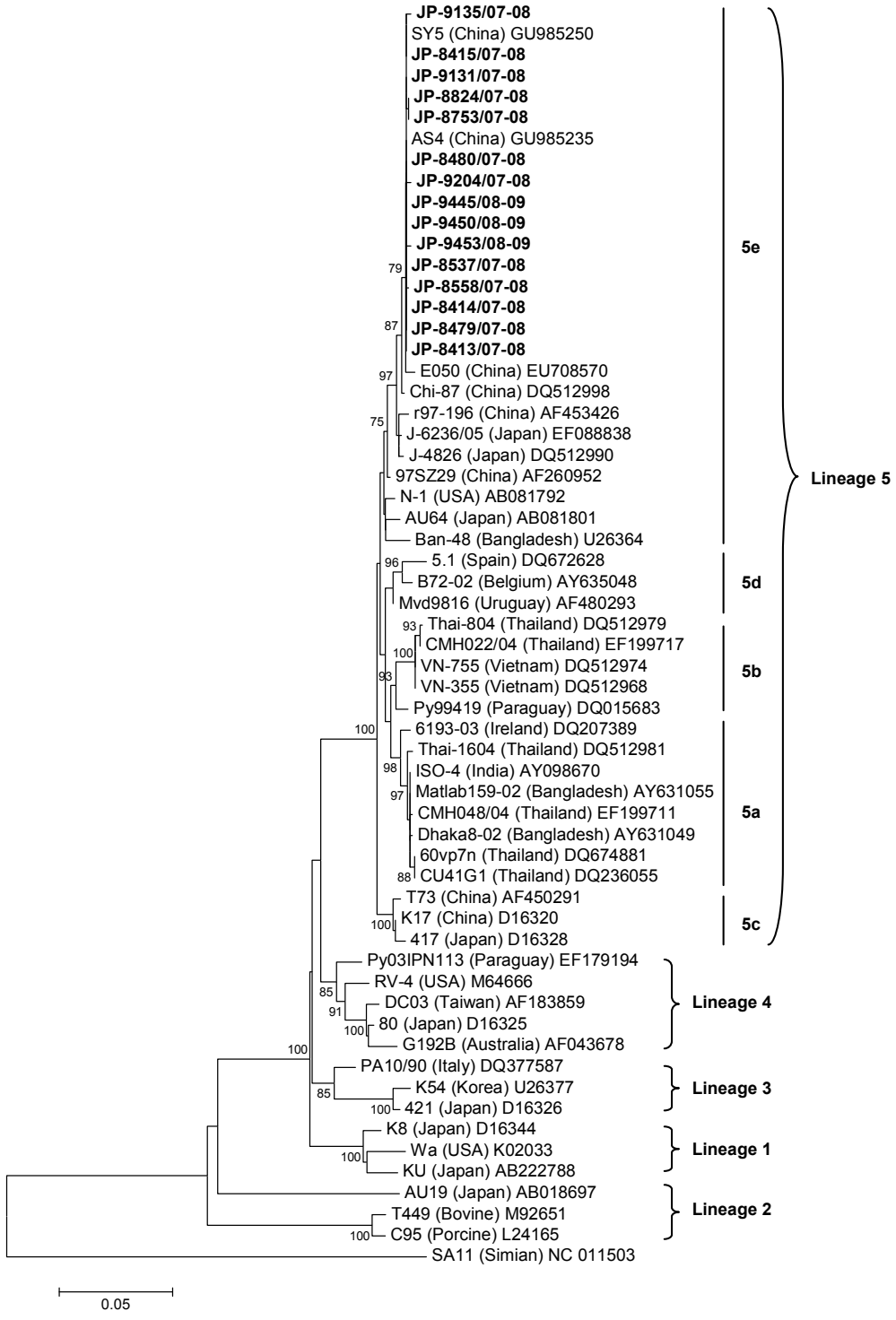


Figure 13. Phylogenetic analyses of VP7 nucleotide sequence genes of (A) G1, (B) G3, and (C) G9 rotaviruses from five locations of Japan. Representative strains detected in this study are represented in boldface. Only bootstrap values >70 are shown. The nucleotide sequences of rotavirus strains described in this study had been submitted to the GenBank database under accession numbers HQ230024-HQ230028 (G1), HQ230029-HQ230033 (G3), and HQ230034-HQ230038 (G9).

B

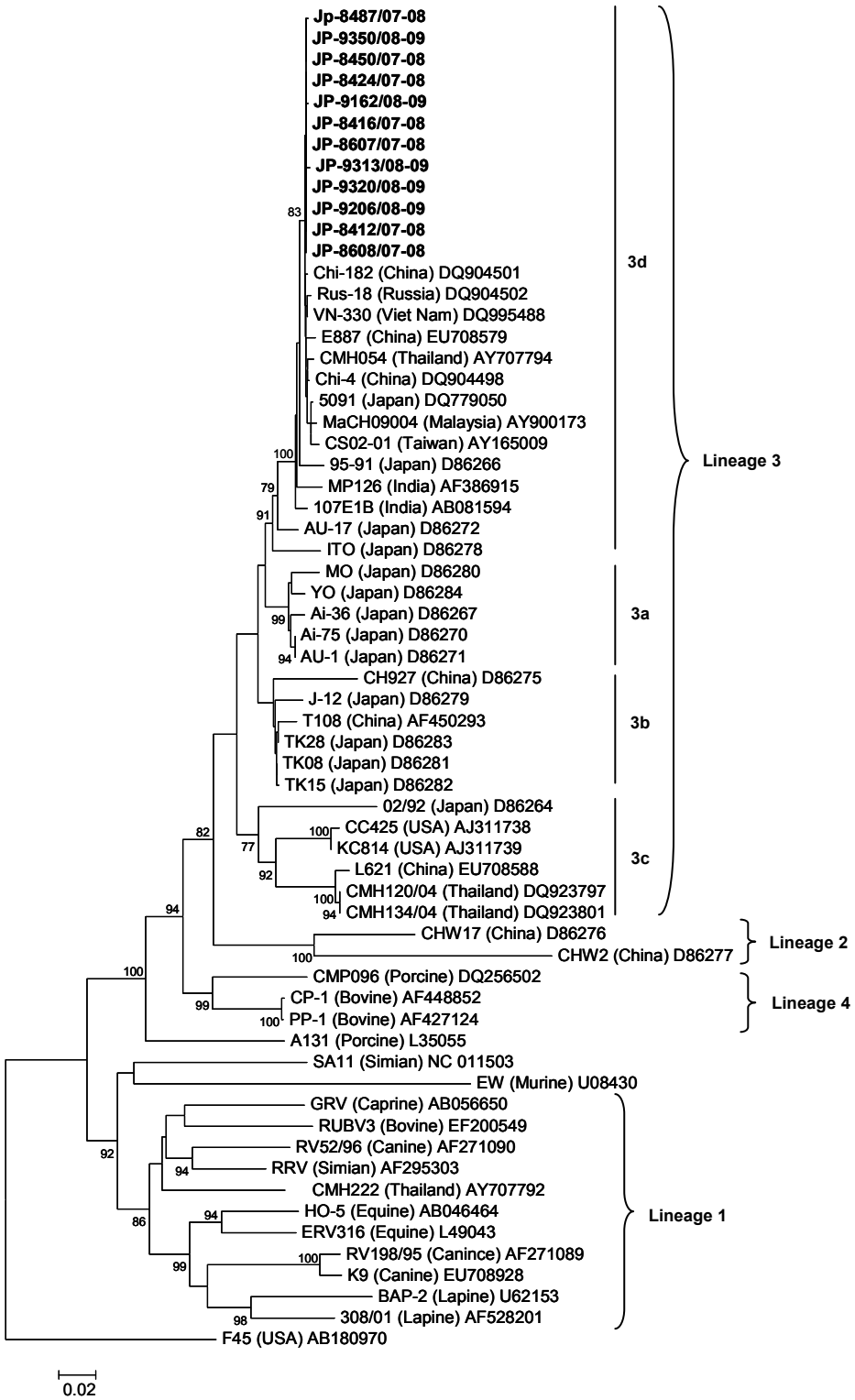


Figure 13-Continued.

C

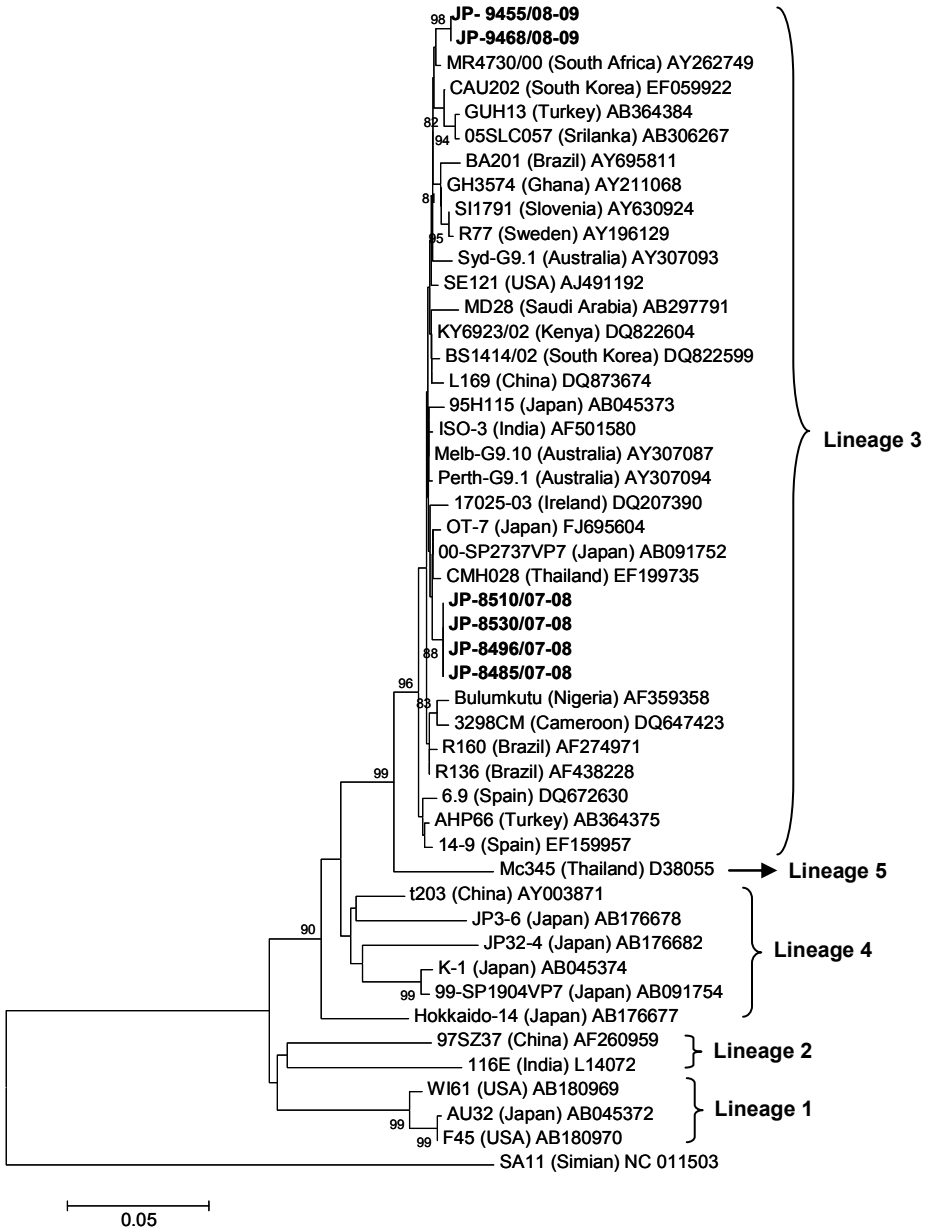


Figure 13-Continued.

Sequence analysis of P-nontypeable rotaviruses

Among the identified 156 rotaviruses, there were 33 strains (21.2%) whose P-genotypes could not be determined by multiplex semi-nested PCR. Nucleotide sequence analysis based on the partial VP4 gene (876 bp) of these 33 P-nontypeable strains turned out to be P[8]. All 33 P[8] strains exhibited 96.6-100% nucleotide sequence identities to each other. The nucleotide sequence alignment between the VP4 sequences of these nontypeable strains and the 1T-1 primer specific for P[8] revealed 5 nucleotide mismatches in all P-nontypeable strains (Fig. 14). The finding indicates, that due to these mismatches, the 1T-1 P[8] primer failed to detect up to 21.2% of the P strains. The phylogenetic tree clearly showed that P[8] could be classified into 3 distinct lineages, in which all Japanese P[8] strains were grouped into lineage 3 (Fig. 15).

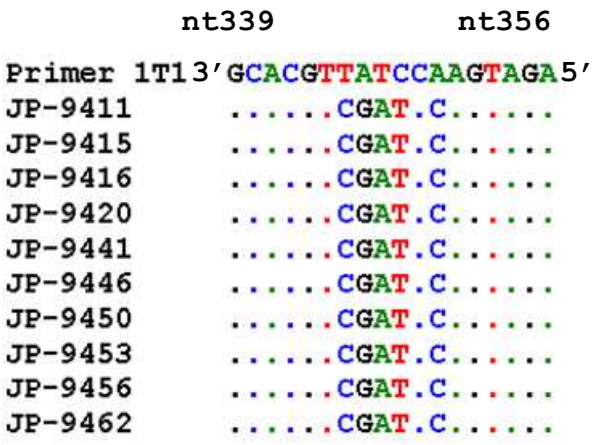


Figure 14. Sequence alignments of the P[8] specific primer binding region showing five nucleotide differences between the reverse complementary sequence of the conventional primer 1T-1 and P[8] strains. Dots represent identical nucleotides.

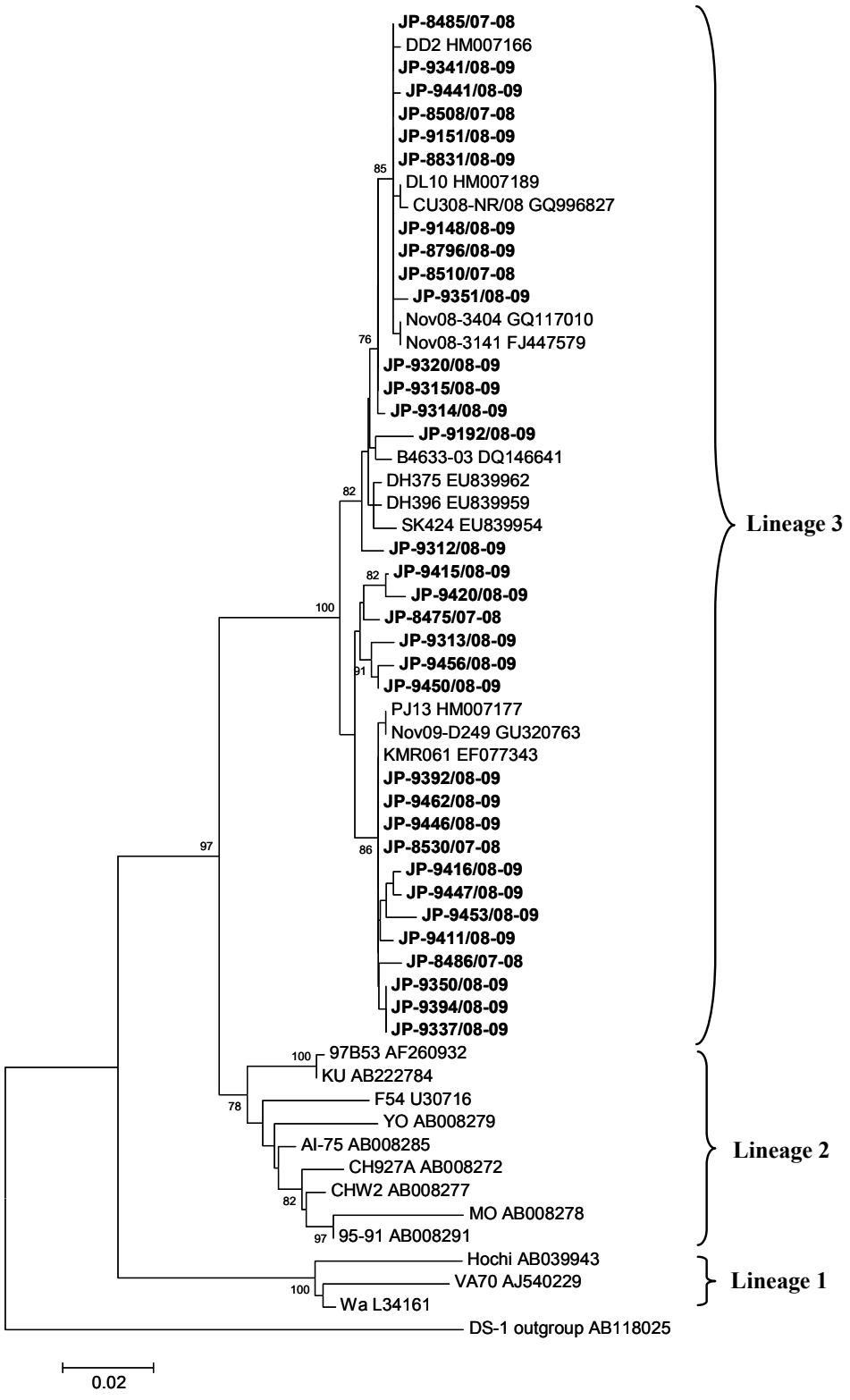


Figure 15. Phylogenetic analysis of VP4 nucleotide sequence genes of P[8]-nontypeable rotaviruses from five locations of Japan. Strains detected in this study are represented in boldface. The strain identities precede the accession numbers. Only bootstrap values >70 are shown. The nucleotide sequences of rotavirus strains described in this study had been submitted to the GenBank database under accession numbers HQ230039-HQ230043.

Distribution of G-and P-genotypes in individual regions (Sapporo, Saga, Tokyo, Osaka, and Maizuru) of Japan during 2003-2009

In 2003, multiplex RT-PCR was introduced in diagnosis of rotavirus and other diarrhea viruses among non-hospitalized children in the same locations (Sapporo, Saga, Tokyo, Osaka, and Maizuru) in Japan. Since its introduction, the overall prevalence of rotavirus infection in all five cities have been published, but they have not been addressed how prevalent it was in individual areas [63-66]. Since it is important to describe the trend of prevalent rotaviruses in each area, the previously published (2003-2007) [63-66] and present (2007-2009) data are summarized and shown in Table 9. The highest detection rate of rotavirus infection was found in Tokyo (20.6%), followed by Osaka (18.7%), Maizuru (16.9%), Sapporo (11.4%), and Saga (9.1%). In Sapporo, Osaka, and Maizuru, G1P[8], as the most common strain, and G3P[8], as the second predominant strain, were detected at frequencies of 52.6%, 47.2%, 57.7% and 47.4%, 37.1%, 31.7%, respectively. In Tokyo, G1P[8] (47.4%) was the predominant strain, followed by G9P[8] (20.6%), whereas in Saga, G3P[8] (38.9%) and G9P[8] (36.1%) were identified as the most dominant types. None of G9P[8] were detected in Sapporo. Moreover, many other uncommon genotypes including G1P[4], G2P[8], G3P[4], G3P[9], and G9P[6] were also reported during 2003-2009.

Table 9. Distribution of G- and P-genotype combinations of rotavirus in children in five cities in Japan during 2003-2009[a]

Area[c]	No. of samples	No. (%) of rotavirus	No. of G- and P-genotypes (%[b])					
			G1P[8]	G2P[8]	G3P[8]	G4P[8]	G9P[8]	Uncommon
Sapporo	166	19 (11.4)	**10 (52.6)**	0	**9 (47.4)**	0	0	0
Saga	394	36 (9.1)	7 (19.4)	0	**14 (38.9)**	0	**13 (36.1)**	2 (6.5)
Tokyo	472	97 (20.6)	**46 (47.4)**	11 (11.3)	13 (13.4)	3 (3.1)	**20 (20.6)**	4 (4.2)
Osaka	518	97 (18.7)	**46 (47.2)**	8 (8.2)	**36 (37.1)**	0	5 (5.1)	2 (2.4)
Maizuru	1843	312 (16.9)	**180 (57.7)**	6 (2)	**99 (31.7)**	0	16 (5.1)	11 (3.5)

[a] The 2003-2007 data were published previously [63-66].
[b] Refers to number of rotavirus-positive specimens in each area.
[c] Areas were ordered according to number of samples.
The first and second most detection rates were highlighted in boldface.

DISCUSSION

In this study, the follow-up surveillance of rotavirus infection was conducted in non-hospitalized infants and children with acute gastroenteritis in five different places in Japan from 2007 to 2009. Rotavirus was recognized as the causative agent of disease in 19.3% cases in 2007-2008 and 12.1% cases in 2008-2009, which are consistent with the results of previous reports in the same population and the observations in Japan with rotavirus infection rates ranging from 10.9% to 28.4% [63-66]. Several studies have found that rotavirus infection occurred frequently in young children under 3 years of age [63, 65]. In agreement with this, the results showed a marked trend toward higher rates of infection in children under 3 years of age (85.8%) versus children over 3 years of age (14.2%), with children of the 1-year-old group (51.9%) being the most frequently affected. The finding was in line with the surveillance on pediatric cases of rotavirus gastroenteritis conducted in Japan, which demonstrated that the main peak from March to April [63, 65, 118].

The results in the present and preceding studies (2004-2007) [63-66] indicated that the predominant G1P[8] strain has been detected at a relatively constant frequency, ranging from 46.7% to 70.1%. In contrast, G3P[8] strains have gradually decreased in number since 2004, causing 32.9% cases in 2004-2005, 17.9% cases in 2005-2006 and 7.3% cases in 2006-2007. In this study, interestingly, the number of G3P[8] strains increased threefold from the last season, 2006-2007 [63], from 7.3% to 21.8%, whereas G2P[4] decreased from 11.4% to 0.7%. A recent study, on the other hand, reported that G2P[4] (20.2%) was more common than G3P[8] (0.6%) in children with acute gastroenteritis admitted in four hospitals of Thailand during the same period of time (2007-2009) [119]. This suggests that the change of the distribution of rotavirus genotypes in each geographic area may be explained by the

co-existence of multiple factors such as anti-rotavirus immunity in children, climate, and water.

Most G1 rotaviruses, which have been reported recently from several Asian countries, belong to lineage 5 (Fig. 13A). It is noted that G1 strains isolated in this study were closely related to those isolated in China from 2006 to 2008 [120], clustering in the same branch within lineage 5e. Accordingly, it is suggested that G1 strains isolated in Japan and China seem to share the same evolutionary ancestor.

Phan et al. described the Japanese G3 rotaviruses detected in 2003-2004, represented by the strain 5091, as having significant amino acid sequence differences compared with G3 rotaviruses in Japan in 1990-1995, and called the latest G3 virus a "new variant" rotavirus [66]. They noted some amino acid substitutions in the antigenic regions in the VP7 of the new variant viruses. The present study indicated that the representative JP-9313/08-09 strain had identical VP7 sequence to the new variant strain 5091 [66] and some amino acid differences were detected in the antigenic regions compared with G3-lineage 3a-c. However, the strain JP-9313/08-09, together with the strain 5091, was found to have only one different amino acid at position 99 in the antigenic regions, compared with G3 prototype strain YO which was isolated in Japan in 1977 [121]. Amino acids at position 96, 213 in the antigenic regions were the same between the new variant G3 rotaviruses (JP-9313/08-09 and 5091) and old prototype strain YO. This means that amino acid divergence in the antigenic regions may not be specific traits of the recently prevailing new variant G3 viruses. In contrast, three amino acids located at the constant region (positions 108, 266, 278) were commonly found in the lineage 3d rotaviruses containing the new variant G3 viruses, but were distinct from other sublineages. Therefore, amino acids at these positions are considered to be a feature of the new variant G3 viruses,

although the influence of these amino acids to the infectivity or virulence of rotavirus has yet to be elucidated. Recently, the 5091-like virus has emerged as the most frequent genotype in China [120]. It is suggested that the 5091-like strain may become a reemerging strain in the near future.

Rotavirus G9 has been recognized as the most widespread of the emerging genotypes since 1996 and to be the frequent cause of severe acute gastroenteritis in many countries, covering all continents of the world. In Japan, only one rotavirus G9 was first detected in Tokyo in 1996-1997 [122]. And then, rotavirus G9 was determined to be the prevailing genotype in 1998-2003 with the detection rate of 15.3% [122, 123]. However, rotavirus G9 appeared to vanish in 2003-2005 [65, 66]. These strains were detected in 6.8% in 2005-2006 and increased sharply to 20.3% in 2006-2007 [63, 64]. In this study, G9 viruses were the third most common genotype in 14.7% rotavirus infection. The phylogenetic tree clearly showed that G9 rotaviruses isolated globally could be classified into three lineages in which Japanese G9 strains were clustered into discrete branches in lineage 3 of the globally spreading G9 viruses (Fig. 13C). By BLAST analysis, it was shown that the 2007-2008 G9 strains had a close genetic relationship with the Thai CBH058/20 strain isolated recently in Thailand [119], whereas the 2008-2009 G9 strains appeared to be more similar to a recent Chinese L720 strain (data not shown) [120]. It is possible that the G9 strains circulating in Japan during 2007-2008 and 2008-2009 might have originated from two different ancestors. It is also suggested by these findings that the G9 rotavirus from the global lineage may become an important genotype in Japan in the near future, considering that G9 rotavirus scarcely affected Japanese people previously. In addition, it should be also noted that the G9 rotavirus was reported to be associated with more severe diseases in children in Latin American countries where G9 became

prevalent [124]. Thus, further surveillance of the rotavirus genotyping and cautions against an increase of G9 will be necessary in Japan.

Nucleotide sequence analysis of P-nontypeable strains revealed that all 33 strains turned out to be P[8]. Nucleotide sequence alignment of these P[8] strains demonstrated that 5 point mutations were observed at the primer binding region for P[8] (1T-1 primer). This may lead to the loss of complementarity with the typing primers and result in either a failure of genotyping or mistyping of rotavirus isolates. In fact, a failure of rotavirus genotyping by RT-PCR due to natural variation in the primer binding sites of VP7 and VP4 genes has been documented previously for G1, G3, G6, G8, G9, G10 and P[4], P[6], P[8], P[11] [65, 125-132]. For this reason, the PCR-based method used for rotavirus typing needs to be monitored and updated regularly to avoid the point mutations that may accumulate at the binding sites for type-specific primers.

There were some differences in the distribution of rotavirus genotypes depending on the regions where collections were undertaken (Table 9). For example, the most variety of rotavirus genotypes had been found in Tokyo where all G1-G4 rotaviruses were detected. G1P[8], G3P[8], and G9P[8] were highly prevalent in almost all regions, whereas G2P[4] and G4P[8] were less frequent than other genotypes and circulated in 3 (Tokyo, Osaka, and Maizuru) out of the 5 regions of Japan.

CONCLUSION

Rotavirus was the second most prevalent following norovirus infection. This report stresses the emergence of G1P[8] during 2007-2009 and re-emergence of G3 strains, which were identical to the “new variant G3” strain 5091 reported in Japan

during 2003-2004. The data on seasonality confirm the previously reported trend that rotavirus infections spread in Japan from March to April. Because of the natural variation in the rotaviral gene sequences, close monitoring of rotavirus genotyping methods is important. Moreover, this study provides knowledge on the diversity of rotavirus genotypes found in different areas of Japan during 2003-2009. Continuous surveillance for group A rotavirus strains and monitoring the diversity of G- and P-genotype nature of strains in an area over time are essential to formulate future vaccine strategies.

CHAPTER III.

DETECTION AND GENETIC CHARACTERIZATION OF SAPOVIRUS IN FIVE LOCATIONS OF JAPAN DURING 2007 to 2009

Molecular epidemiology of sapovirus infection

Sapovirus was detected in 37 of 1,008 (3.7%) specimens tested: 19 of 477 (4%) in 2007-2008 and 18 of 531 (3.4%) in 2008-2009. The highest incidence of infection was in the 12-23-month-old group (20, 54%). Most (29, 78.4%) of these infections occurred in infants and children <3 years of age (Fig. 16A). None of sapovirus infection was found in the under 6-month old group. Infections increased slightly from November to February (27 cases, 73%) (Fig. 16B).

Phylogenetic analysis of sapovirus sequences

In order to further characterize sapoviruses into genogroups and genotypes, the PCR products of sapoviruses in the study were sequenced. Nineteen sapovirus sequences detected in 2007-2008 were analyzed by phylogenetic analysis and grouped using the recent sapovirus capsid region classification scheme [54]. Surprisingly, sapovirus genogroup IV (79%) emerged as the most prevalent, followed by GI/4 (16%), and GI/1 (5%) (Table 10). A phylogenetic tree showed that three sequences of GI/4 genotype having 98%–100% nucleotide identity with each other were clustered with the Karachi/872/91/PK and Osaka/5836/JP strains known to belong to GI/4 genotype. One GI/1 sequence had 97% nucleotide identity and clustered with Manchester sequence (Fig. 17).

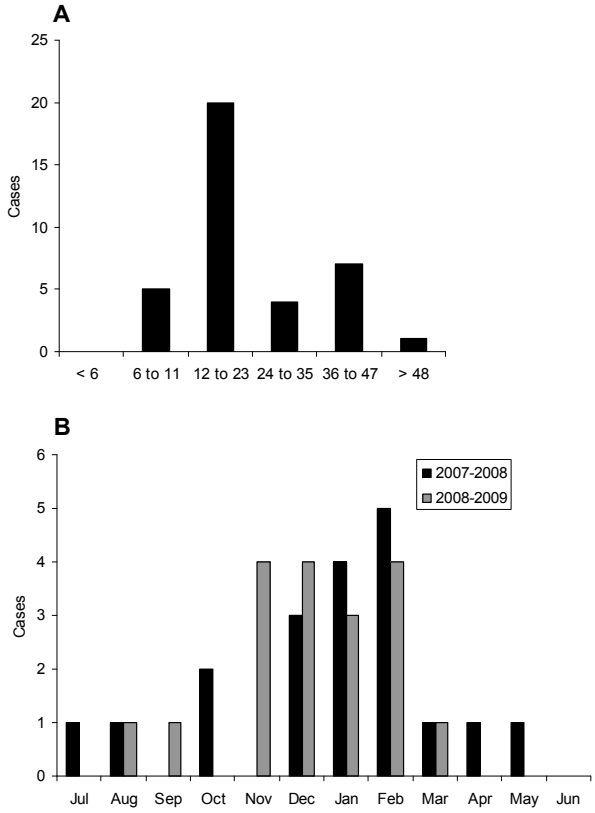


Figure 16. Age-related (A) and seasonal distributions (B) of sapovirus infections in children with acute gastroenteritis in five cities of Japan during 2007-2009.

Table 10. Distribution of sapovirus infection in children in five cities of Japan during 2007-2009

Year	No. of specimens tested	No. (%) positive sapovirus	Genotype (%[a])			
			GI/1	GI/4	GII/1	GIV
2007-2008	477	19 (4)	1 (5)	3 (16)	0	15 (79)
2008-2009	531	18 (3.4)	14 (77.8)	0	4 (22.2)	0

[a] Refers to number of SaV-positive specimens in each year.

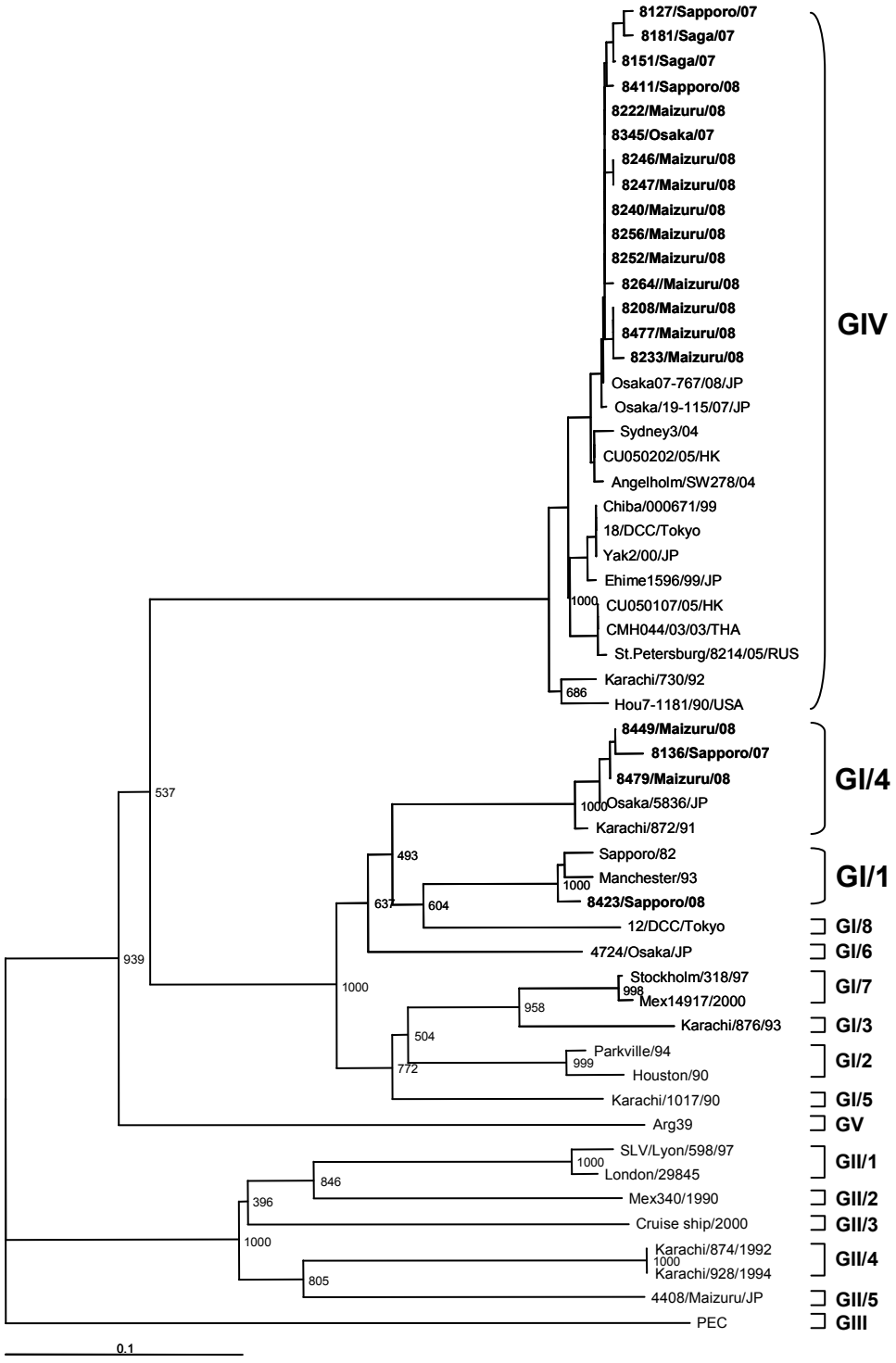


Figure 17. Phylogenetic tree of nucleotide sequences of sapovirus strains detected during 2007-2008 (represented in boldface). The tree was constructed from partial nucleotide sequences of the capsid region, using PEC strain (a porcine sapovirus) as an outgroup. The numbers on each branch indicate the bootstrap values. The distance scale in nucleotide substitutions per position is shown. GenBank accession numbers of reference strains are as follows: Osaka07-767/08/JP (AB433785), Osaka/19-115/07/JP (AB327280), Sydney3/04/AU (DQ104357), CU050202/05/HK (DQ155647), Angelholm/SW278/04/SE (DQ125333), Chiba/000671/ 99/JP (AJ786349), 18/DCC/Tokyo/43/JP (AB236378), Yak2/00/JP (AB046353), Ehime1596/ 99/JP (AM049952), CU050107/05/HK (DQ155646), CMH044/ 03/03/THA (EF600796), St.Pertersburg/8214/05/RUS (FJ214057), Karachi/730/92/PK

(AB126249), Hou7-1181/90/ USA (AF435814), Osaka/5836/04/JP(AB242324), Karachi/ 872/91/PK (AB181231), Sapporo/ 82/JP (U65427), Manchester/93/UK (X86560), 12/DCC/ Tokyo/44/JP (AB235380), 4724/ Osaka/02/JP (AB180212), Stockholm/318/97/SE (AF194182), Mex14917/00/USA (AF435813), Karachi/867/93/PK (AB181132), Parkville/ 94/UK (U73124), Houston/90/USA (U95644), Karachi/1017/90/PK (AB181227), Arg39/ 95/ARG (AY289803), Lyon/598/97/F (AJ271056), London/29845/92/UK (U95645), Mex340/90/USA (AF435812), Cruise ship/ 00/USA (AY289804), Karachi/874/92/PK (AB181129), Karachi/928/94/PK (AB181128), 4408/Maizuru/03/JP, and PEC (AF182760). The sequences of sapovirus strains detected in 2007-2008 have been submitted to GenBank under accession numbers FJ445092-FJ445110.

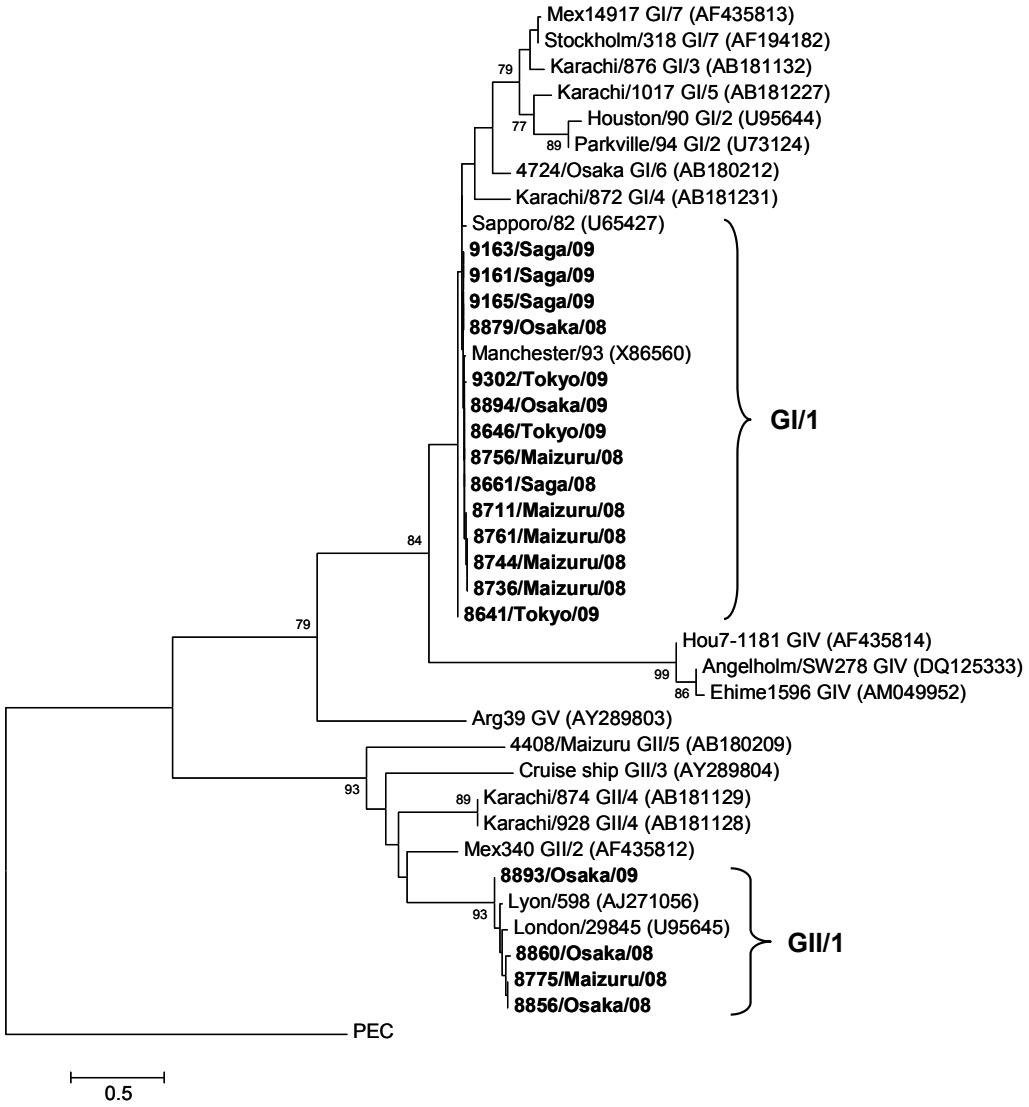


Figure 18. Phylogenetic tree of nucleotide sequences of sapovirus strains detected during 2008-2009 (represented in boldface). The tree was constructed from partial nucleotide sequences of the capsid region, using PEC strain (a porcine sapovirus) as an outgroup. The numbers on each branch indicate the bootstrap values. The distance scale in nucleotide substitutions per position is shown. The strain identities precede the accession numbers. Only bootstrap values >70 are shown. The sequences of sapovirus strains detected in 2008-2009 have been submitted to GenBank under accession numbers HM030920-HM030923 and HM106431-HM106446.

In 2008-2009, interestingly, GI/1 (77.8%) re-emerged to be the predominant genotype, followed by GII/1 (22.2%). GIV and GI/4, which were detected in 2007-2008 disappeared in 2008-2009. All sequences of GI/1 genotype exhibited more than 98%-100% nucleotide identity to each other and grouped with Manchester/93 and Sapporo/82 strains known to belong to GI/1 genotype, as shown in Fig. 18. Four GII/1 sequences had 98% nucleotide identify and clustered with Lyon/598/97 and London/29845 strains.

Intergenogroup recombinant sapovirus during 2007-2008

Nucleotide sequence comparison of the 15 GIV strains detected in 2007-2008 shared little divergence among themselves (98%–100% identity), even when they were detected in regions of Japan distant from each other. They are likely to represent a single strain, 8208/Maizuru/08/JP. The 8208/Maizuru/08/JP sequence closely matched Ehime1107, SW278 [53], and Yak2 [133], which were previously established as intergenogroup recombinant sapovirus strains with the GII polymerase region and GIV capsid region, with 97% and 96% nucleotide identities, respectively. To determine whether these GIV strains were the recombinant sapovirus, 5 of the 15 GIV strains were randomly selected as representative, and long genomic fragments that included part of the RNA polymerase and part of the capsid genes were amplified by using primers SR80/2 and SLV5749, which generated a 1,151-bp product. The products were directly sequenced, and capsid- and polymerase-based phylogenetic trees confirmed these strains as the recombinant sapoviruses (Fig. 19). It is suggested that the GIV strains isolated in 2007-2008 were intergenogroup recombinants. In addition, strain 8208/Maizuru/08 was selected as representative for analyzing the full-length genome sequence (7.4 kb) in order to localize the potential recombination

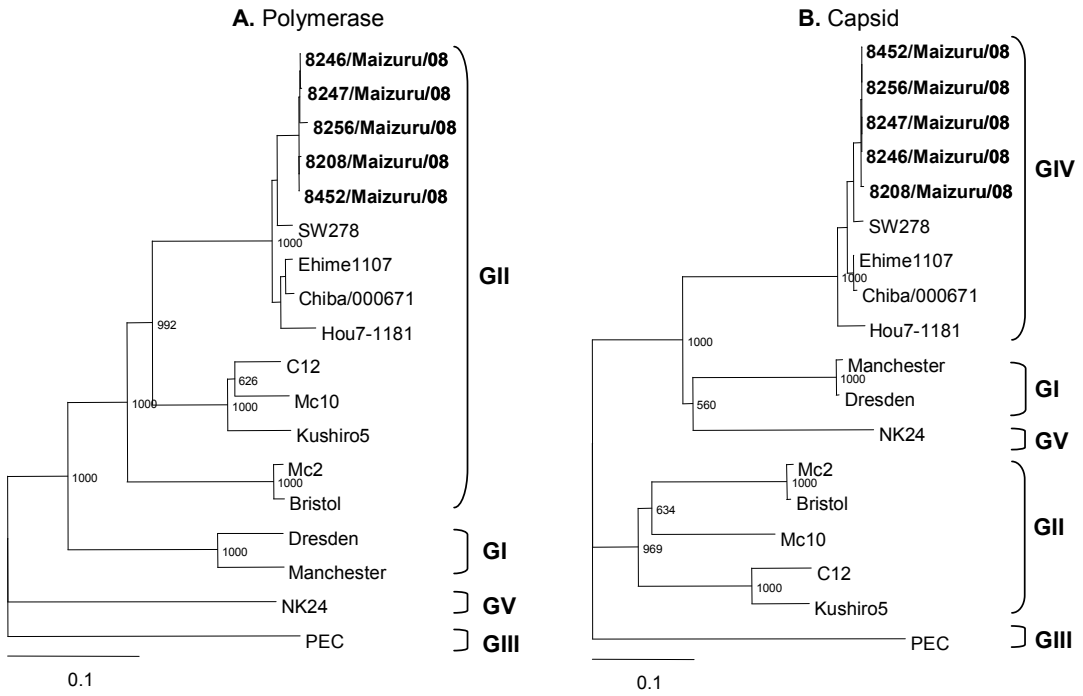


Figure 19. Changing genotypes of sapoviruses on the basis of phylogenetic trees of nucleotide sequences. The trees were constructed from partial nucleotide sequences of polymerase (A) and capsid (B) regions.

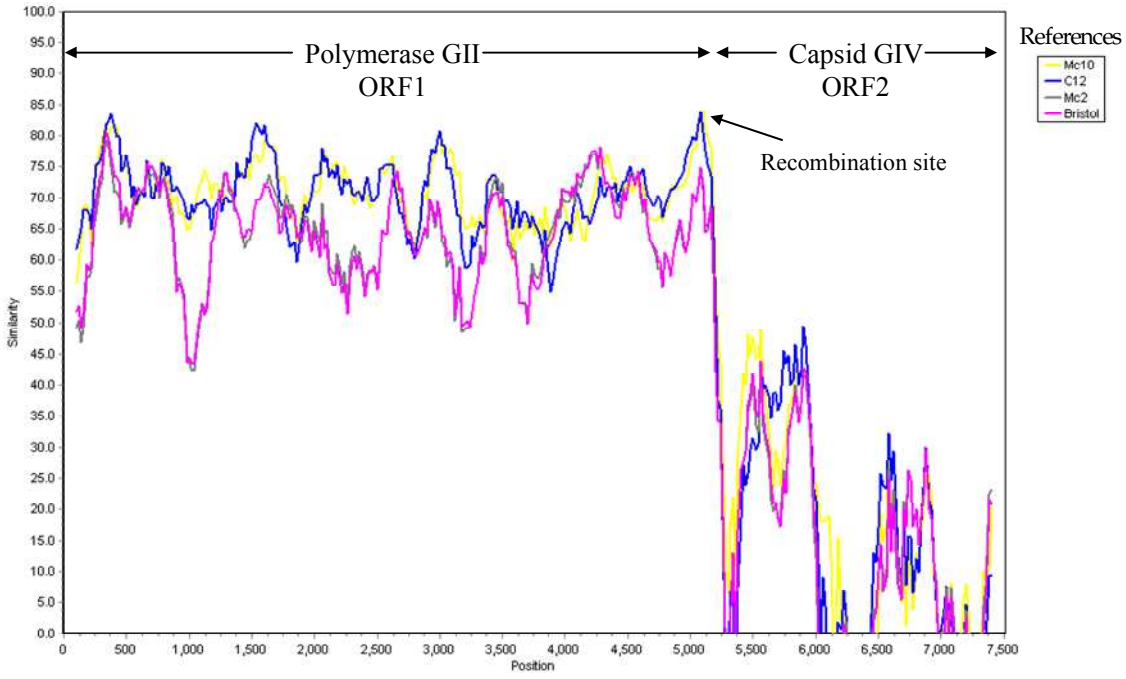


Figure 20. Intergenogroup recombination observed in 8206/Maizuru/08 strain. SimPlot analysis was performed using the 7.4-kb fragment covering the ORF1 and ORF2 with the putative parental strains, Mc10, C12, Mc2, and Bristol as references. A window size of 100 bp with an increment of 20 bp was shown. The recombinant site is suspected to be located within the ORF1/2 overlap, as shown by the arrow.

site and to understand a possible recombination mechanism of this strain. SimPlot analysis showed a sudden drop in nucleotide similarity after ORF1, suggesting the ORF1/ORF2 overlap as the potential cross-over site (Fig. 20).

Genetic analysis of capsid protein of GI/1 sapovirus in 2008-2009

The complete capsid genes were compared for 15 GI/1 strains collected between 1982 and 2009 (26 years), including 4 representative strains detected in this study and 11 from GenBank database. Sequence comparison showed ≥92% nucleotide and ≥99% amino acid identities among these GI/1 strains. Several nucleotide mutations were found and scattered throughout the capsid gene (Fig. 21). These mutations resulted in sporadic or permanent amino acid changes in 25 of the affected codons. Except for sporadic amino acid changes, 6 amino acids were accumulative: amino acid positions 8 in N-terminal variable region, 100 in N-terminal region, 333 in a central variable region (CVR), and 505, 524, and 557 in C-terminal region. Of the 6 accumulative sites, positions 100 (S→A), 333 (V→I), 505 (A→S), and 557 (K→R) appeared to be fixed in the genome overtime.

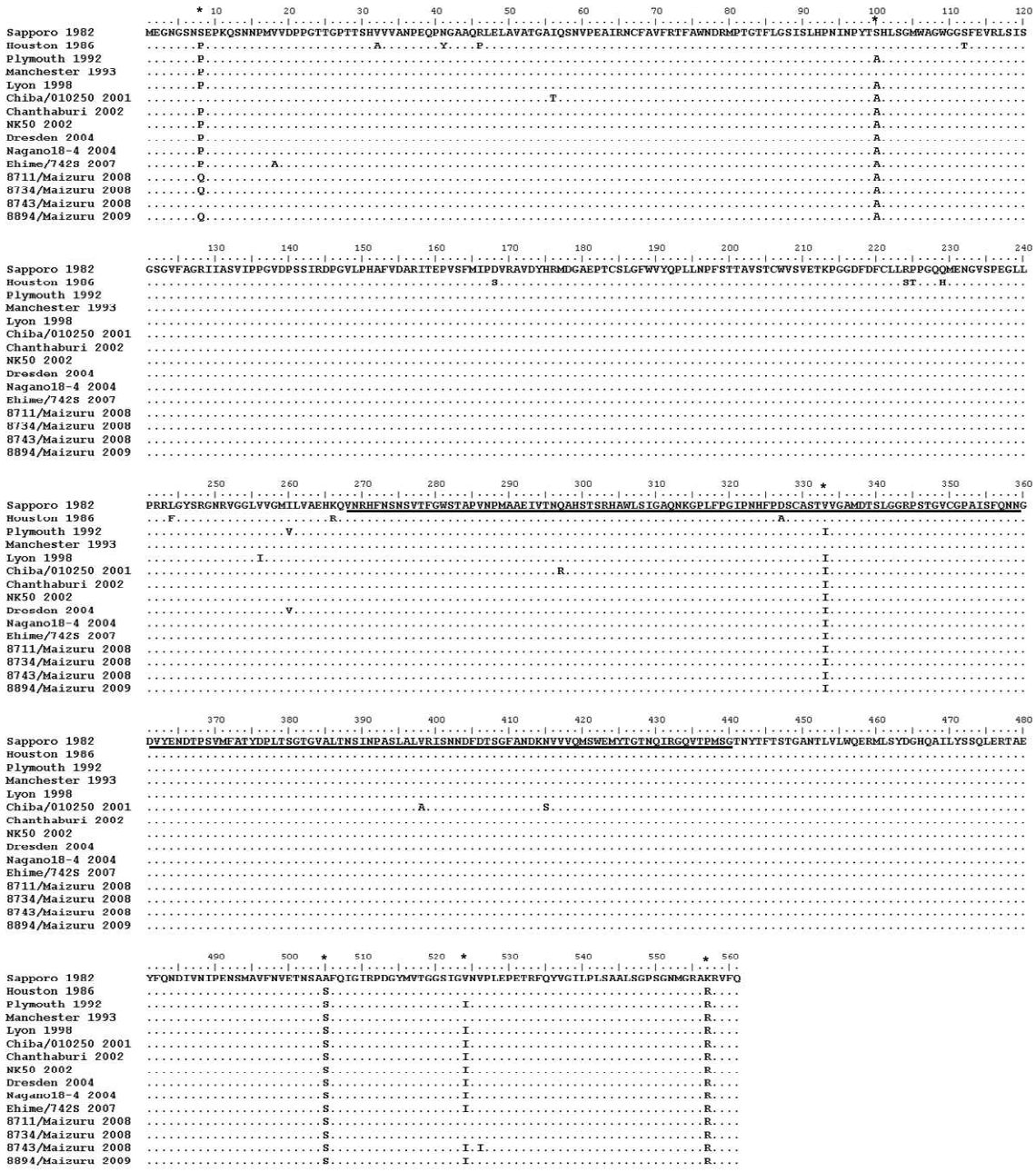


Figure 21. Amino acid alignment of the full-length capsid gene of 15 GI/1 sapovirus strains showing mutations along the aligned sequences. Underlined sequence refers to the central variable region (CVR) corresponding to the P2 domain of norovirus. Asterisk indicates accumulative sites.

DISCUSSION

Sapovirus infection causes acute gastroenteritis in all age groups, through it occurs predominantly in infants and young children [34, 55, 57, 58]. Overall, 37 of 1,008 (3.7%) fecal specimens tested were positive for sapovirus, and positive specimens were found in all age groups of the subjects included in the study. However, most (78.4%) of the sapovirus infections occurred in infants and children <3 years of age. These results were in line with previous reports on sapovirus epidemiology worldwide, in which sapovirus prevalence was shown to range from 0.3 to 9.3%, far below the prevalence of either rotavirus or norovirus [80, 134, 135]. However, sapoviruses have been reported as causes of occasional outbreaks, but their prevalence seems to be increasing. Similar findings were reported recently from Canada (17.6%) [136], Japan (17.6%) [60], and Hungary (12.6%) [137]. The difference of the detection rate may also be partly explained by differences in the target country, patients (outpatients vs. hospitalized), settings (sporadic vs. outbreak), and detection methods (ELISA, and multiplex or uniplex PCR with different primer sets).

For the past 4 seasons (2003-2007) of sapovirus surveillance conducted in the same setting and population in Japan, sapovirus GI/1 was the most common genotype during 2003–2004, and thereafter genotype GI/6 dominated over the GI/1 in 2004–2005 (Fig. 22) [55]. Then, the GI/6 genotype was replaced by the predominant sapovirus GI/1 since 2005 until 2007 [57, 71]. GIV genogroup emerged and became the predominant strain in the first season (2007-2008) of this study. Interestingly, the GIV sapovirus was found to rapidly become the predominant strain recovered from infected children or adults in different locations of Japan [138, 139] and other countries (Canada, the Netherlands) [136, 140] in 2007-2008. These findings indicate

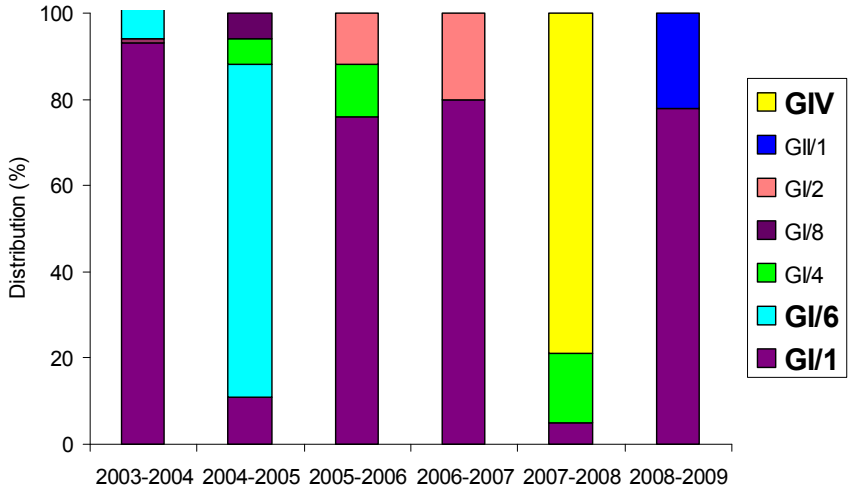


Figure 22. Distribution of sapovirus genotypes detected in five locations of Japan during 2003 to 2009. The 2003-2007 data were published previously [55, 57, 71].

that GIV strains are stable in environment and may be globally distributed. My study also showed that the changing distribution of sapovirus continued in the next season (2008-2009) with the re-emergence of GI/1 and the disappearance the GIV strain. Taken together, these data indicate that the distribution of sapovirus genotype in Japan has changed with time.

Recently, several types of recombinant sapoviruses, including intergenogroup and intragenogroup, have occurred frequently in Japan [53, 57, 58, 72, 133, 141], which indicates that recombination between sapovirus genomes is another important feature of the evolution of sapovirus. This study demonstrated for the first time of the emergence of intergenogroup (GII/GIV) sapovirus in Japan during 2007-2008. Also, these recombinant strains were detected in 4 locations distant from each other: Maizuru city (10 cases), Sapporo and Saga (2 cases each), and Osaka (1 case), which suggests that the recombinant strains were widely spread through the country. Although GIV sapovirus strains had been isolated in some countries, including Japan, Thailand, Pakistan, and Hong Kong [54, 134, 142, 143], they were detected less often than strains from the other genogroups, and whether they are the recombinant strains

has not been confirmed. Genotyping based solely on the capsid sequence may not be sufficient with the growing identification of naturally occurring recombinant sapoviruses, GIV in particular, as the recombination breakpoint is close to or within the ORF1/2 overlap. Thus, recombinants would be missed by sequencing only the capsid region. Therefore, I recommend that sequencing in both the capsid gene and RNA polymerase sequences is needed to identify such recombinant viruses.

The re-emergence of GI/1 strains was clearly demonstrated in second season of this study. Sequence analysis of the capsid encoding gene and in particular the analysis of possible amino acid changes was perfumed. Chen et al. [144] and Okada et al. [142] previously reported that the central variable region (CVR) of sapovirus capsid gene was homologous to the P2 domain of norovirus virus, which is predicted to be located on a surface-exposed part of the capsid and is involved in the receptor binding function and antigenic properties. Bull et al. [39] also described that more than 5% amino acid diversity, as a consequence of an immune pressure, in the P2 domain in the capsid gene between past GII.4 noroviruses and variants. In contrast to what has been described for the GII.4 norovirus, there was only 1% amino acid diversity across the complete capsid gene of GI/1 sapoviruses over a 26-year period, showing a clearly low rate of change. This result confirmed that sapoviruses are less diverse than norovirus and are conserved within the same genotype.

CONCLUSION

In this study, the molecular epidemiology of sapovirus was described. During the 2-year period from 2007-2009, the detection rate of sapovirus was 3.9%. The highest presence of sapovirus infection was found in the 1-year-old children. The rapid changing distribution of sapovirus genotypes was shown for the first time with

the emergence of intergenogroup recombinant sapovirus (GII/GIV) in 2007-2008 and the re-emergence of GI/1 in 2008-2009. Further epidemiological studies should be conducted to determine the trend of sapovirus infection in Japan in the coming year.

CHAPTER IV.

EMERGENCE OF A NEW VARIANT OF GII. 6 NOROVIRUS IN SHIZUOKA

Molecular epidemiology of norovirus infection

A total of 187 fecal specimens collected from non-hospitalized children with acute gastroenteritis in a pediatric clinic in Shizuoka (Fujieda city) during the period from July 2008 to June 2009 were investigated for the presence of diarrhea viruses. Diarrhea viruses were overall detected in 158 of 187 (84.5%). Of the viruses detected, norovirus was the most prevalent (55.6%), followed by group A rotavirus (21.4%), adenovirus (4.3%), sapovirus (1.1%), astrovirus (1.1%), and group C rotavirus and mixed infection (0.5% each). The highest prevalence of diarrhea viruses was found in infants and children aged 12-23 months (50.6%) and lowest in infants aged <6 months (3.9%). It was also found that children under 3 years of age had a high rate of diarrhea infection (79.8%).

Phylogenetic analysis of norovirus sequences

Sequence analysis of norovirus-positive samples showed evidence for the presence of several genotypes, with the distribution as follows: GII.4 (53.8%), GII.6 (40.4%), GII.3 (4.8%), and GII.7 (1%). Interestingly, GII.6, which has been known as a rare genotype among GII noroviruses, emerged for the first time as the second most common strain (40.4%) in this study. The phylogenetic tree revealed that all GII.4 sequences (100%) clustered with the recent GII.4/2006b variant. In addition, GII.6 genotypes were divided into 3 subclusters (a, b, and c) and all GII.6 strains detected in this study clustered together in a single branch within subcluster c (Fig. 23).

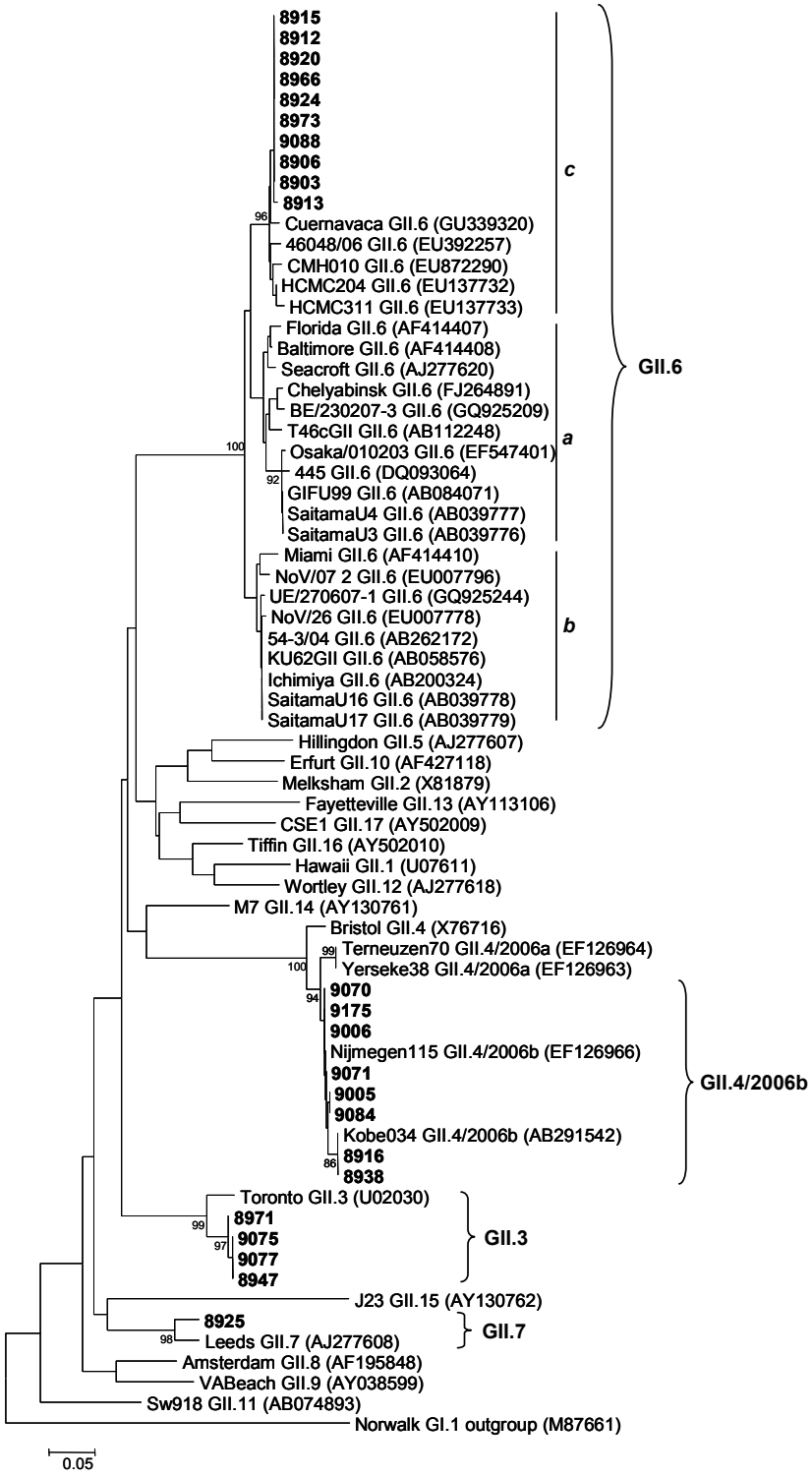


Figure 23. Phylogenetic analysis of the partial capsid sequence of norovirus detected in Shizuoka during 2008-2009. Representative strains are represented in boldface. Reference strain identities precede the accession numbers. Only bootstrap values >70 are shown.

Outbreak of GII.6 norovirus in Shizuoka

Norovirus infection was detected continuously in a 8-month period lasting from October to May and reached its peak between November and January. However, the picture was different with regard to individual norovirus genotypes during these three months. Norovirus GII.6 emerged, dominated over other genotypes during the first two months (November and December) and then disappeared in January, when GII.4 appeared as the most prevalent (Fig. 24). This pattern of infection indicated an outbreak of norovirus GII.6 in these subjects, and this would be the first outbreak of acute gastroenteritis attributed to norovirus in Shizuoka. The results also showed that norovirus infection (89%) caused by GII.4 occurred mainly in infants and young children under 3 years of age. In contrast, the detection rates of norovirus GII.6 in children under 3 years (52.4%) were similar to children over 3 years of age (47.6%).

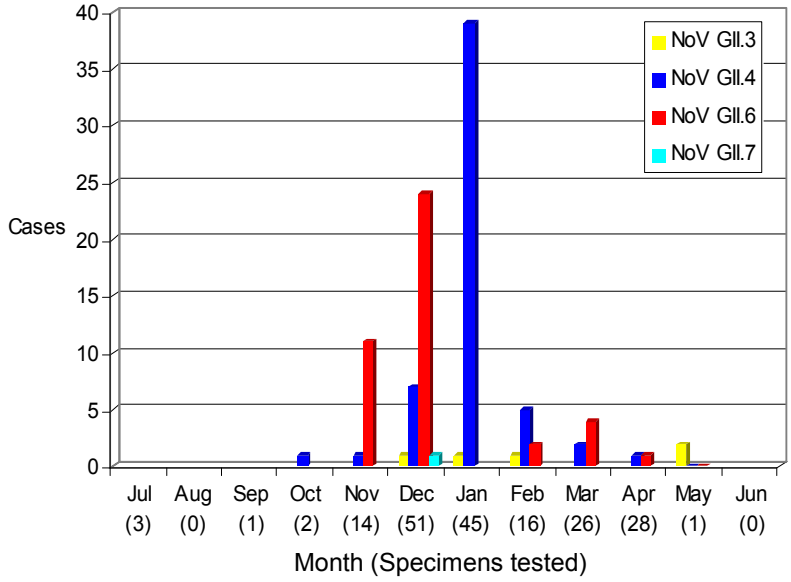


Figure 24. Monthly distribution of norovirus infections in Shizuoka during 2008-2009.

Genetic analysis of capsid protein of GII.6 noroviruses

In an attempt to understand the molecular basis of norovirus genetic diversity within the GII.6 genotype, the full-length capsid genes of five representative strains were analyzed and compared with previously published GII.6 sequences in GenBank database. The 5 GII.6 strains were 99% identical to each other at the amino acid level. The phylogenetic tree based on the complete amino acid sequence confirmed 3 subclusters (Fig. 25). By amino acid sequence comparison, strains in subcluster c shared 91.2-92.5%% and 94.3-94.8% similarities with those in subcluster a and b, respectively, indicating 7.5-9.8% and 5.2-5.7% divergence between these subclusters. Amino acid differences mostly occurred in the P2 subdomain up to 16.1-17.5% (between subcluster c and a) and 10.9-12.9% (between subcluster c and b). Alignment of amino acid sequences of the P domain revealed 52 hotspots of variation across the P domain (amino acid positions where ≥2 amino acid changes between strains were observed), as well as random point mutations (Fig. 26). Strains in subcluster b and c had 20 substitutions in the same positions and also contained 3 amino acid deletions in position 358-360. In addition to the 20 substitutions identified in both subcluster b and c, 8 substitutions were unique for subcluster c: 291 (S→G), 297 (S→A), 306 (A→R), 308 (N→D), 370 (D→N), 390 (N→D), (393 (N→V), and 398 (I→V). Taken together, the sequence analysis suggested that the GII.6 that emerged and caused outbreak in Shizuoka was a new GII.6 variant.

A 3D structure of the most variable P domain of a dimer of the new GII.6 variant strain 8915 was constructed by homology modeling, based on the known 3D structure of the VA387/98/US GII.4 genotype (2OBT), which has 56% amino acid sequence identity in the P domain to the 8915 strain. The homology model was generated by EasyModeller and verified with structural validation programs. Results

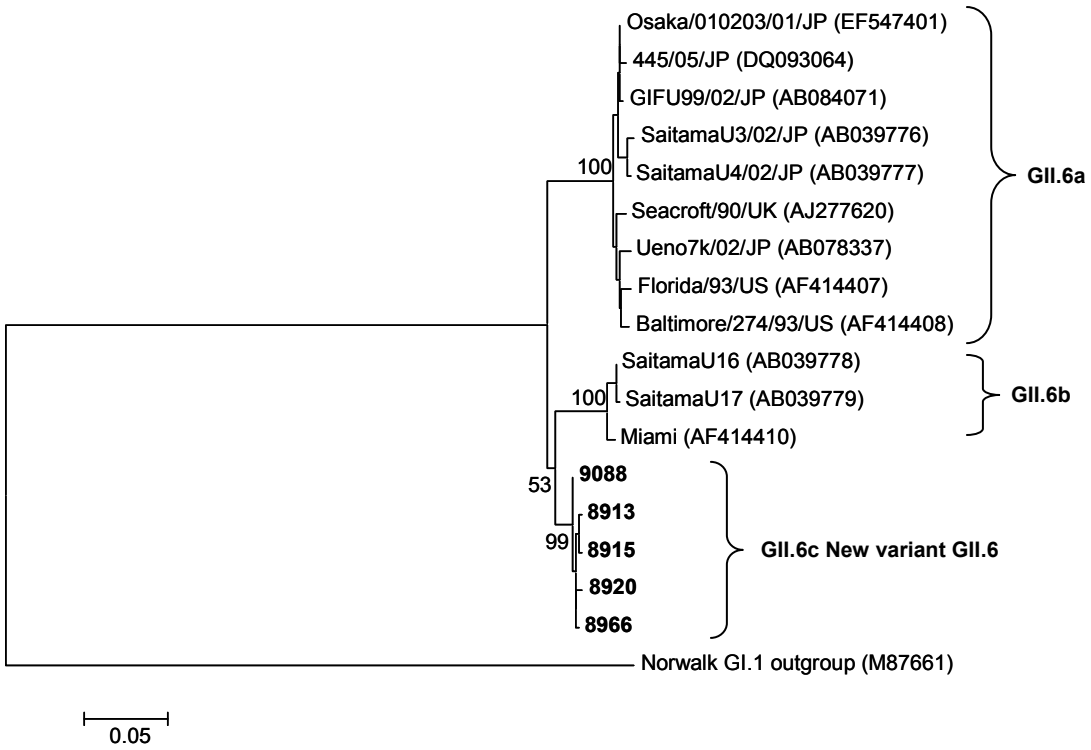


Figure 25. Phylogenetic analysis based on the full-length capsid sequences of GII.6 noroviruses. Representative strains detected in this study are represented in boldface.

of the model validation were shown in APPENDIX (Table 11, Fig. 30-31). The structural model of the dimer of the GII.6 strain as shown in Fig. 27 revealed a large loop that probably extends away from the P2 surface, which is similar with the structural models of GII.3 [42, 145, 146]. The model also showed that the 8 unique substitutions of subcluster c viruses were positions on loops and β-sheet of the P2 subdomain (Fig. 27, red). Several changes found in GII.6 strains of both subcluster b and c were predicted to be located at the surface-exposed P1 and P2 subdomains (Fig. 27, blue).

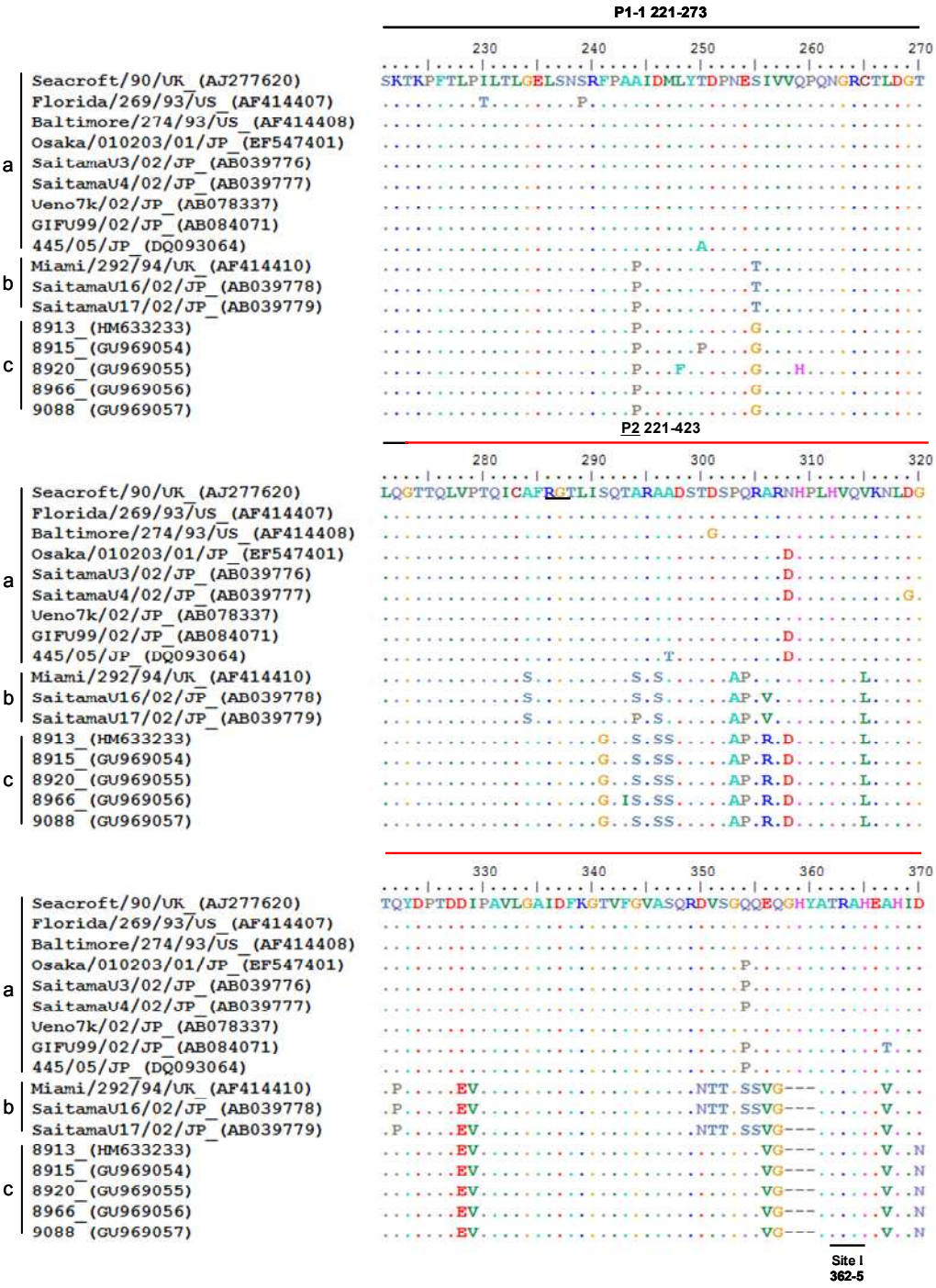


Figure 26. Amino acid alignment of the P domain of GII.6 strains showing mutations across the region. Underlined residues indicate the RGT motif that corresponds to the RGD-like motif of other noroviruses. Sites I, II, and III indicate putative receptor-binding sequences of GII noroviruses.

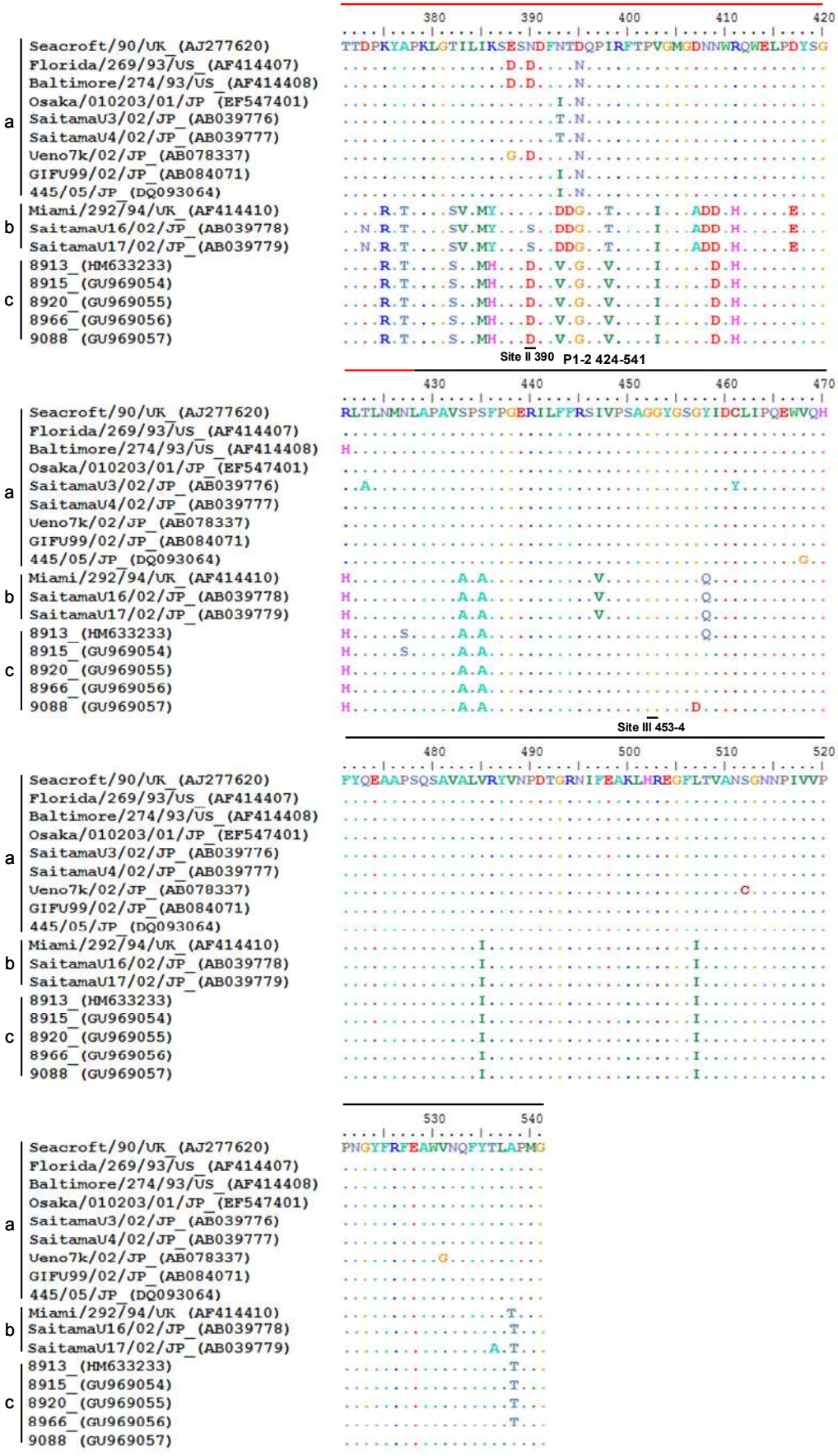
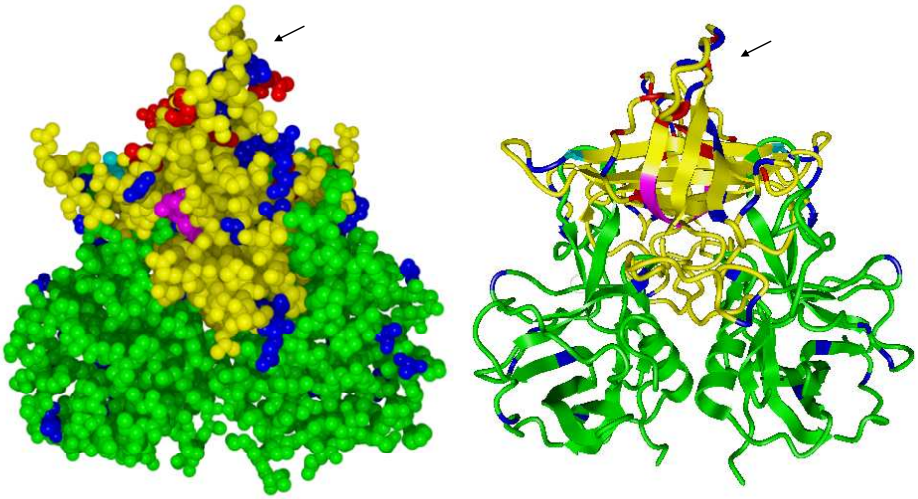
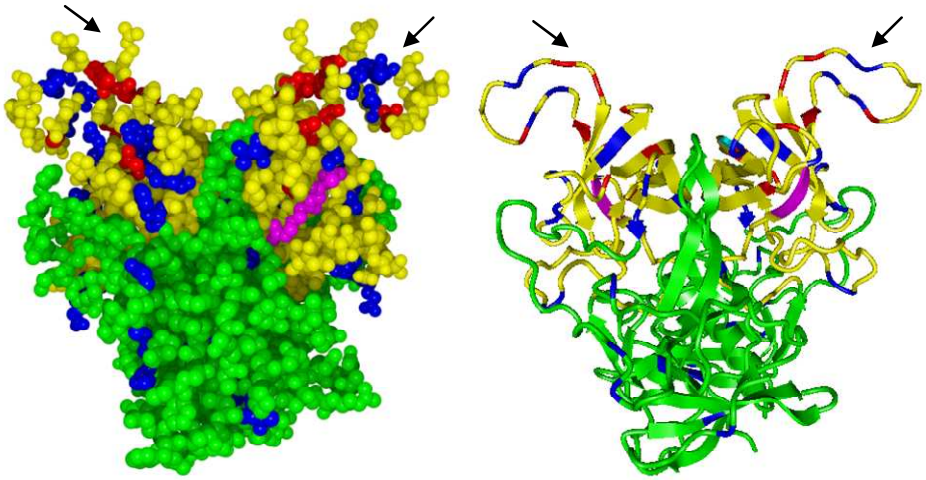


Figure 26-continued.

A



B



C

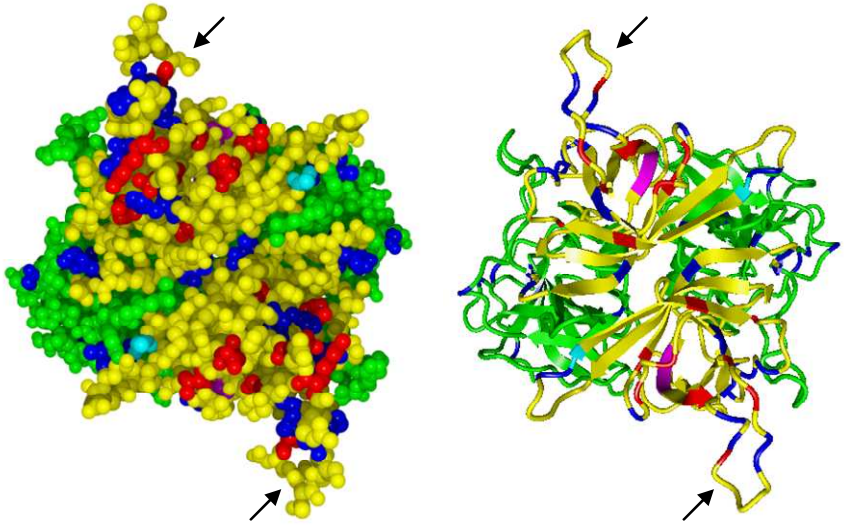


Figure 27. The 3D structure of a dimer of the GII.6 P domain, as predicted by EasyModeller, showing the location of the amino acid changes. Ball and ribbon formations were shown in (A) front view, (B) side view, and (C) top view. P1 subdomain is shown in green, while P2 subdomain is shown in yellow. Location of amino acid changes in GII.6 strains in both subcluster b and c is shown in blue, while the unique amino acid changes of the new variant GII.6 strain are shown in red. Additional RGT motif (magenta) is also shown. A long loop inserted in the P2 region is shown by arrow. Position next to the three amino acid deletions in GII.6 of sucluster b and c is shown in cyan.

DISCUSSION

Although other viral agents were detected in the samples studied, norovirus was by far the most important viral agent associated with acute gastroenteritis reported here. Norovirus corresponded to 55.6% of all the samples tested, showing a relatively high detection rate in Shizuoka. Although norovirus GII.6 strains had been isolated in some countries, including Japan, they were less common than strains from the other genotypes [54, 68, 69, 99, 101]. Of note, norovirus GII.6 infection in the present study was identified with a high prevalence, accounting for 40.4% compared with 53.8% for the GII.4 infection, indicating the emergence of GII.6 strain. The GII.6 infection rapidly emerged as a leading genotype during a short period of two months (November and December), although thereafter some cases of GII.6 infection were identified during February to April. Taken together, this indicates an outbreak attributed to the GII.6 among infants and children, and this would be the first outbreak of acute gastroenteritis caused by GII.6 in Shizuoka. Furthermore, the result also showed that children over 3 years of age were commonly affected by GII.6 in this outbreak, indicating the immune invasion of this GII.6 strain.

At least 3 distinct GII.6 subclusters (a-c) appeared in different part of the world during the past 19 years. The GII.6 strains in this study clustered together in subcluster c. As the full-length sequences from strains in subcluster c were not available in GenBank, 5 GII.6 strains detected in this study were selected as representative for further analysis in order to gain further insights into the genetic variability within the GII.6 genotype. Alignment of the full-length capsid genes revealed that Shizuoka GII.6 strains had a great variability, particular in the P2 subdomain up to 10.9-17.5% compared with viruses in subcluster a and b, showing a clearly high rate of change. Zheng et al. [36] reported that there were 0-14.1%

differences in complete capsid sequences for norovirus strains within a cluster (genotype). The Shizuoka GII.6 had 5.2-9.8% amino acid differences in the capsid protein compared with that of subcluster a and b viruses, indicating that GII.6 emerged in Shizuoka is a new genetic GII.6 variant. Comparison of the structural models of the dimer of GII.4 and GII.6 P domains demonstrated the P2 subdomain of the GII.6 strain probably contains a large loop that extends away from the surface and away from the region that contains the reported receptor binding domain in the GII.4 P domain. The model also shows that 8 specific amino acid positions occur in exposed loops and surface of the P domain. This supports the idea that a new variant may possibly emerge from the accumulation of mutations in the P2 domain through immune response-driven evolution [32, 35, 42, 147, 148]. Additionally, a comparison of the structural models between Ueno 7k (subcluster a) and 8915 Shizuoka strain demonstrates three deletions found in Shizuoka strains caused a loop to become smaller (data not shown). Taken together, these structural changes may be sufficiently prominent to obscure the site from recognition by existing antibodies, thus permitting the GII.6 virus to escape the host pre-existing immunity. This analysis presented in this study has predicted amino acid changes within the P2 domain that may be immunologically significant but it remains to be proven that mutations result in antibody-escape mutants.

CONCLUSION

This study shows for the first time that norovirus infections in Shizuoka were mainly caused by GII.4/2006b variant and a new GII.6 variant. Simultaneous analysis of the full-length capsid sequences of a new GII.6 variant resulted in the discovery of amino acid variations located in the exposed-surface on the P2 subdomain. The

homology model of the GII.6 demonstrates that several amino acid changes are located in the exposed-surface on the P2 domain. This indicates the possible immune pressure, driving evolution.

SUMMARY

The present study has shown the difference of diarrhea virus distribution among non-hospitalized children with diarrhea in six areas (Sapporo, Saga, Tokyo, Osaka, Maizuru, and Shizuoka) of Japan between 2007 and 2009. Norovirus was the major pathogen in surveillance, dominating over group A rotavirus. The increased detection rate of norovirus resulted from using a sensitive semi-nested PCR. The prevalence of rotavirus and sapovirus were consistent to previous studies conducted in the same setting and population. The season distribution of norovirus and sapovirus gastroenteritis showed a winter peak, while rotavirus infection was common in early spring. Even noroviruses, rotaviruses and sapoviruses were detected in six different localities in Japan; however, they shared the great sequence identities each other in each year when they belonged to the same cluster, suggesting they came from the same origin of infection. A significant increase in the prevalence of the G3 rotavirus in Japan and other counties indicate that this strain may be globally distributed. The epidemiological data on rotavirus infection described in this study would be useful in comparison of the genotype distribution after the start of rotavirus vaccination in Japan in 2011. One highlight in this study was the first report of an emergence of intergenogroup GII/GIV recombinant sapoviruses, suggesting that this strain may have widely spread throughout Japan as a new dominant strain. Another highlight was the first report of an outbreak due to a new variant of GII.6 noroviruses in Shizuoka. In addition, GII.6/GII.14 recombinant noroviruses emerged as the second most prevalent in 2007-2008. The structural models of the new variant GII.6 and the GII.14 demonstrated for the first time that several variations were observed in exposed loops and surface of the P2 domain. These findings indicate that these viruses are likely to be mediated by herd immunity. Therefore, it will be essential to define the

evolutionary patterns of each genotype in order to identify appropriate contemporary isolates for vaccine design. Although vaccine studies should focus on the predominant GII.4 genotype, control of a single genotype through vaccination would probable leave an evolutionary niche that would be exploited by another GII genotype, perhaps a GII.6 or a GII.14 variant. Thus, the design of efficacious multivalent norovirus vaccines is essential for reducing the large-scale outbreaks that commonly occur in settings in which close human contact is unavoidable.

FUTURE WORKS

Several studies have indicated that histo-blood group antigens (HBGA) function as receptors or co-receptors for productive norovirus infection. HBGAs are complex carbohydrates linked to glycoproteins or glycolipids that are present on the red blood cells and mucosal epithelial cells or as free antigens in biological fluids such as blood, saliva, intestinal contents and milk. These antigens are synthesized by sequential additions of monosaccharides to the active portion of the antigen precursors by several glycosyltransferases that are controlled mainly by the ABO, Lewis, and secretor gene families. Noroviruses can generally be divided into two binding profile groups: (i) strains that bind A/B and or H antigens, and (ii) strains that bind Lewis and/or H antigens [35]. To investigate the binding properties of GII.14 and GII.6 strains detected in this study, binding and blocking experiments using saliva, synthetic oligosaccharides, and monoclonal antibodies to characterize norovirus binding specificities are necessary. Through these further experiments, questions would be answered whether or not theses variations result in conformational changes that alter the binding preference of the viruses to persist by targeting unique populations for infection.

Although RT-PCR has been accepted as the standard method for diagnosis of viral gastroenteritis, it requires well-trained personnel and sophisticated equipments. Recently, a immunochromatography (IC) test for rapid detection of rotaviruses and noroviruses has been developed. However, there is no commercially available IC test for sapovirus diagnosis. Therefore, the development of a rapid IC test should be made to diagnosis not only GI and GII, but also GIV sapoviruses.

ACKNOWLEDGEMENTS

At first, I would like to thank the Asian Youth Fellowship (AYF) program and the Ministry of Education, Culture, Sports, Sciences and Technology and the Ministry of Health, Labor and Welfare (MEXT) for giving me a chance to study in Japan.

I am grateful to Professor Masashi Mizuguchi, staffs, and students in the Department of Developmental Medical Sciences, School of International Health, Graduate School of Medicine, the University of Tokyo, for making my three years in Japan very fruitful and enjoyable through their strong encouragement and support. I am also grateful to Professor Hiroshi Ushijima, Dr Shoko Okitsu, and staffs in Aino University for the valuable advice, suggestions and support throughout my studies.

I also would like to thank Drs Toshimasa Kuroiwa, Kunio Kaneshi, Yuichi Ueda, Shigekazu Nakaya, Shuichi Nishimura, Kumiko Sugita, Tadashi Nishimura, Masaaki Kobayashi and Atsuko Yamamoto for fecal specimen collection.

REFERENCES

1. Parashar UD, Gibson CJ, Bresse JS, Glass RI: **Rotavirus and severe childhood diarrhea**. *Emerg Infect Dis* 2006, **12**(2):304-306.
2. Parashar UD, Hummelman EG, Bresee JS, Miller MA, Glass RI: **Global illness and deaths caused by rotavirus disease in children**. *Emerg Infect Dis* 2003, **9**(5):565-572.
3. Soriano-Gabarro M, Mrukowicz J, Vesikari T, Verstraeten T: **Burden of rotavirus disease in European Union countries**. *Pediatr Infect Dis J* 2006, **25**(1 Suppl):S7-S11.
4. Glass RI, Bresee JS, Turcios R, Fischer TK, Parashar UD, Steele AD: **Rotavirus vaccines: targeting the developing world**. *J Infect Dis* 2005, **192 Suppl 1**:S160-166.
5. Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Banyai K, Estes MK, Gentsch JR, Iturriza-Gomara M, Kirkwood CD, Martella V *et al*: **Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments**. *Arch Virol* 2008, **153**(8):1621-1629.
6. Solberg OD, Hasing ME, Trueba G, Eisenberg JN: **Characterization of novel VP7, VP4, and VP6 genotypes of a previously untypeable group A rotavirus**. *Virology* 2009, **385**(1):58-67.
7. Maes P, Matthijnssens J, Rahman M, Van Ranst M: **RotaC: a web-based tool for the complete genome classification of group A rotaviruses**. *BMC Microbiol* 2009, **9**:238.
8. Trojnar E, Otto P, Johne R: **The first complete genome sequence of a chicken group A rotavirus indicates independent evolution of mammalian and avian strains**. *Virology* 2009, **386**(2):325-333.
9. Ursu K, Kisfali P, Rigo D, Ivanics E, Erdelyi K, Dan A, Melegh B, Martella V, Banyai K: **Molecular analysis of the VP7 gene of pheasant rotaviruses identifies a new genotype, designated G23**. *Arch Virol* 2009, **154**(8):1365-1369.
10. Abe M, Ito N, Morikawa S, Takasu M, Murase T, Kawashima T, Kawai Y, Kohara J, Sugiyama M: **Molecular epidemiology of rotaviruses among healthy calves in Japan: isolation of a novel bovine rotavirus bearing new P and G genotypes**. *Virus Res* 2009, **144**(1-2):250-257.
11. Schumann T, Hotzel H, Otto P, Johne R: **Evidence of interspecies transmission and reassortment among avian group A rotaviruses**. *Virology* 2009, **386**(2):334-343.
12. 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**(1):29-56.
13. Gentsch JR, Laird AR, Bielfelt B, Griffin DD, Banyai K, Ramachandran M, Jain V, Cunliffe NA, Nakagomi O, Kirkwood CD *et al*: **Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs**. *J Infect Dis* 2005, **192 Suppl 1**:S146-159.
14. Rahman M, Matthijnssens J, Yang X, Delbeke T, Arijs I, Taniguchi K, Iturriza-Gomara M, Iftekharuddin N, Azim T, Van Ranst M: **Evolutionary history and global spread of the emerging g12 human rotaviruses**. *J Virol* 2007, **81**(5):2382-2390.

15. Glass RI, Parashar UD, Bresee JS, Turcios R, Fischer TK, Widdowson MA, Jiang B, Gentsch JR: **Rotavirus vaccines: current prospects and future challenges**. *Lancet* 2006, **368**(9532):323-332.
16. **Rotavirus vaccines**. *Wkly Epidemiol Rec* 2007, **82**(32):285-295.
17. **Meeting of the immunization Strategic Advisory Group of Experts, April 2009--conclusions and recommendations**. *Wkly Epidemiol Rec* 2009, **84**(23):220-236.
18. Linhares AC, Velazquez FR, Perez-Schael I, Saez-Llorens X, Abate H, Espinoza F, Lopez P, Macias-Parra M, Ortega-Barria E, Rivera-Medina DM *et al*: **Efficacy and safety of an oral live attenuated human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in Latin American infants: a randomised, double-blind, placebo-controlled phase III study**. *Lancet* 2008, **371**(9619):1181-1189.
19. Ruiz-Palacios GM, Perez-Schael I, Velazquez FR, Abate H, Breuer T, Clemens SC, Cheuvart B, Espinoza F, Gillard P, Innis BL *et al*: **Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis**. *N Engl J Med* 2006, **354**(1):11-22.
20. Vesikari T, Karvonen A, Prymula R, Schuster V, Tejedor JC, Cohen R, Meurice F, Han HH, Damaso S, Bouckenooghe A: **Efficacy of human rotavirus vaccine against rotavirus gastroenteritis during the first 2 years of life in European infants: randomised, double-blind controlled study**. *Lancet* 2007, **370**(9601):1757-1763.
21. Madhi SA, Cunliffe NA, Steele D, Witte D, Kirsten M, Louw C, Ngwira B, Victor JC, Gillard PH, Cheuvart BB *et al*: **Effect of human rotavirus vaccine on severe diarrhea in African infants**. *N Engl J Med*, **362**(4):289-298.
22. Vesikari T, Matson DO, Dennehy P, Van Damme P, Santosham M, Rodriguez Z, Dallas MJ, Heyse JF, Goveia MG, Black SB *et al*: **Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine**. *N Engl J Med* 2006, **354**(1):23-33.
23. Patel M, Pedreira C, De Oliveira LH, Tate J, Orozco M, Mercado J, Gonzalez A, Malespin O, Amador JJ, Umana J *et al*: **Association between pentavalent rotavirus vaccine and severe rotavirus diarrhea among children in Nicaragua**. *Jama* 2009, **301**(21):2243-2251.
24. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: **Food-related illness and death in the United States**. *Emerg Infect Dis* 1999, **5**(5):607-625.
25. Atmar RL, Estes MK: **The epidemiologic and clinical importance of norovirus infection**. *Gastroenterol Clin North Am* 2006, **35**(2):275-290, viii.
26. Rockx B, De Wit M, Vennema H, Vinje J, De Bruin E, Van Duynhoven Y, Koopmans M: **Natural history of human calicivirus infection: a prospective cohort study**. *Clin Infect Dis* 2002, **35**(3):246-253.
27. Caul EO: **Small round structured viruses: airborne transmission and hospital control**. *Lancet* 1994, **343**(8908):1240-1242.
28. Lopman BA, Reacher MH, Vipond IB, Sarangi J, Brown DW: **Clinical manifestation of norovirus gastroenteritis in health care settings**. *Clin Infect Dis* 2004, **39**(3):318-324.
29. Kapikian AZ, Wyatt RG, Dolin R, Thornhill TS, Kalica AR, Chanock RM: **Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis**. *J Virol* 1972, **10**(5):1075-1081.

30. Tan M, Huang P, Meller J, Zhong W, Farkas T, Jiang X: **Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket**. *J Virol* 2003, **77**(23):12562-12571.
31. Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, Zhang XC, Jiang X, Li X, Rao Z: **Structural basis for the recognition of blood group trisaccharides by norovirus**. *J Virol* 2007, **81**(11):5949-5957.
32. Allen DJ, Gray JJ, Gallimore CI, Xerry J, Iturriza-Gomara M: **Analysis of amino acid variation in the P2 domain of the GII-4 norovirus VP1 protein reveals putative variant-specific epitopes**. *PLoS One* 2008, **3**(1):e1485.
33. Iritani N, Vennema H, Siebenga JJ, Siezen RJ, Renckens B, Seto Y, Kaida A, Koopmans M: **Genetic analysis of the capsid gene of genotype GII.2 noroviruses**. *J Virol* 2008, **82**(15):7336-7345.
34. Lindesmith LC, Donaldson EF, Lobue AD, Cannon JL, Zheng DP, Vinje J, Baric RS: **Mechanisms of GII.4 norovirus persistence in human populations**. *PLoS Med* 2008, **5**(2):e31.
35. Donaldson EF, Lindesmith LC, Lobue AD, Baric RS: **Viral shape-shifting: norovirus evasion of the human immune system**. *Nat Rev Microbiol*, **8**(3):231-241.
36. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS: **Norovirus classification and proposed strain nomenclature**. *Virology* 2006, **346**(2):312-323.
37. Tan M, Jiang X: **The p domain of norovirus capsid protein forms a subviral particle that binds to histo-blood group antigen receptors**. *J Virol* 2005, **79**(22):14017-14030.
38. Siebenga JJ, Vennema H, Renckens B, de Bruin E, van der Veer B, Siezen RJ, Koopmans M: **Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006**. *J Virol* 2007, **81**(18):9932-9941.
39. Bull RA, Tu ET, McIver CJ, Rawlinson WD, White PA: **Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis**. *J Clin Microbiol* 2006, **44**(2):327-333.
40. Tu ET, Bull RA, Greening GE, Hewitt J, Lyon MJ, Marshall JA, McIver CJ, Rawlinson WD, White PA: **Epidemics of gastroenteritis during 2006 were associated with the spread of norovirus GII.4 variants 2006a and 2006b**. *Clin Infect Dis* 2008, **46**(3):413-420.
41. Doyle TJ, Stark L, Hammond R, Hopkins RS: **Outbreaks of noroviral gastroenteritis in Florida, 2006-2007**. *Epidemiol Infect* 2009, **137**(5):617-625.
42. Donaldson EF, Lindesmith LC, Lobue AD, Baric RS: **Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations**. *Immunol Rev* 2008, **225**:190-211.
43. Worobey M, Holmes EC: **Evolutionary aspects of recombination in RNA viruses**. *J Gen Virol* 1999, **80 (Pt 10)**:2535-2543.
44. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC: **Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination**. *J Virol* 1999, **73**(1):152-160.
45. Worobey M, Rambaut A, Holmes EC: **Widespread intra-serotype recombination in natural populations of dengue virus**. *Proc Natl Acad Sci U S A* 1999, **96**(13):7352-7357.

46. Chare ER, Gould EA, Holmes EC: **Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses**. *J Gen Virol* 2003, **84**(Pt 10):2691-2703.
47. Jiang X, Huang PW, Zhong WM, Farkas T, Cubitt DW, Matson DO: **Design and evaluation of a primer pair that detects both Norwalk- and Sapporo-like caliciviruses by RT-PCR**. *J Virol Methods* 1999, **83**(1-2):145-154.
48. Iritani N, Seto Y, Kubo H, Murakami T, Haruki K, Ayata M, Ogura H: **Prevalence of Norwalk-like virus infections in cases of viral gastroenteritis among children in Osaka City, Japan**. *J Clin Microbiol* 2003, **41**(4):1756-1759.
49. LoBue AD, Thompson JM, Lindesmith L, Johnston RE, Baric RS: **Alphavirus-adjuvanted norovirus-like particle vaccines: heterologous, humoral, and mucosal immune responses protect against murine norovirus challenge**. *J Virol* 2009, **83**(7):3212-3227.
50. LoBue AD, Lindesmith L, Yount B, Harrington PR, Thompson JM, Johnston RE, Moe CL, Baric RS: **Multivalent norovirus vaccines induce strong mucosal and systemic blocking antibodies against multiple strains**. *Vaccine* 2006, **24**(24):5220-5234.
51. Tacket CO, Sztein MB, Losonsky GA, Wasserman SS, Estes MK: **Humoral, mucosal, and cellular immune responses to oral Norwalk virus-like particles in volunteers**. *Clin Immunol* 2003, **108**(3):241-247.
52. Chiba S, Sakuma Y, Kogasaka R, Akihara M, Horino K, Nakao T, Fukui S: **An outbreak of gastroenteritis associated with calicivirus in an infant home**. *J Med Virol* 1979, **4**(4):249-254.
53. Hansman GS, Takeda N, Oka T, Oseto M, Hedlund KO, Katayama K: **Intergenogroup recombination in sapoviruses**. *Emerg Infect Dis* 2005, **11**(12):1916-1920.
54. Akihara S, Phan TG, Nguyen TA, Yagyu F, Okitsu S, Muller WE, Ushijima H: **Identification of sapovirus infection among Japanese infants in a day care center**. *J Med Virol* 2005, **77**(4):595-601.
55. Phan TG, Trinh QD, Yagyu F, Okitsu S, Ushijima H: **Emergence of rare sapovirus genotype among infants and children with acute gastroenteritis in Japan**. *Eur J Clin Microbiol Infect Dis* 2007, **26**(1):21-27.
56. Hansman GS, Takeda N, Katayama K, Tu ET, McIver CJ, Rawlinson WD, White PA: **Genetic diversity of Sapovirus in children, Australia**. *Emerg Infect Dis* 2006, **12**(1):141-143.
57. Phan TG, Khamrin P, Quang TD, Dey SK, Takanashi S, Okitsu S, Maneekarn N, Ushijima H: **Emergence of intragenotype recombinant sapovirus in Japan**. *Infect Genet Evol* 2007, **7**(4):542-546.
58. Phan TG, Okitsu S, Muller WE, Kohno H, Ushijima H: **Novel recombinant sapovirus, Japan**. *Emerg Infect Dis* 2006, **12**(5):865-867.
59. Pang XL, Honma S, Nakata S, Vesikari T: **Human caliciviruses in acute gastroenteritis of young children in the community**. *J Infect Dis* 2000, **181 Suppl 2**:S288-294.
60. Phan TG, Trinh QD, Yagyu F, Sugita K, Okitsu S, Muller WE, Ushijima H: **Outbreak of sapovirus infection among infants and children with acute gastroenteritis in Osaka City, Japan during 2004-2005**. *J Med Virol* 2006, **78**(6):839-846.

61. Nakata S, Chiba S, Terashima H, Yokoyama T, Nakao T: **Humoral immunity in infants with gastroenteritis caused by human calicivirus**. *J Infect Dis* 1985, **152**(2):274-279.
62. Phan TG, Nguyen TA, Yan H, Yagyu F, Kozlov V, Kozlov A, Okitsu S, Muller WE, Ushuijma H: **Development of a novel protocol for RT-multiplex PCR to detect diarrheal viruses among infants and children with acute gastroenteritis in Eastern Russia**. *Clin Lab* 2005, **51**(7-8):429-435.
63. Dey SK, Thongprachum A, Ota Y, Phan TG, Nishimura S, Mizuguchi M, Okitsu S, Ushijima H: **Molecular and epidemiological trend of rotavirus infection among infants and children in Japan**. *Infect Genet Evol* 2009, **9**(5):955-961.
64. Phan TG, Khamrin P, Quang TD, Dey SK, Takanashi S, Okitsu S, Maneekarn N, Ushijima H: **Detection and genetic characterization of group A rotavirus strains circulating among children with acute gastroenteritis in Japan**. *J Virol* 2007, **81**(9):4645-4653.
65. Phan TG, Khamrin P, Trinh DQ, Dey SK, Yagyu F, Okitsu S, Nishio O, Ushijima H: **Genetic characterization of group A rotavirus strains circulating among children with acute gastroenteritis in Japan in 2004-2005**. *Infect Genet Evol* 2007, **7**(2):247-253.
66. Phan TG, Trinh QD, Khamrin P, Kaneshi K, Ueda Y, Nakaya S, Nishimura S, Sugita K, Nishimura T, Yamamoto A *et al*: **Emergence of new variant rotavirus G3 among infants and children with acute gastroenteritis in Japan during 2003-2004**. *Clin Lab* 2007, **53**(1-2):41-48.
67. Phan TG, Kaneshi K, Ueda Y, Nakaya S, Nishimura S, Yamamoto A, Sugita K, Takanashi S, Okitsu S, Ushijima H: **Genetic heterogeneity, evolution, and recombination in noroviruses**. *J Med Virol* 2007, **79**(9):1388-1400.
68. Phan TG, Kuroiwa T, Kaneshi K, Ueda Y, Nakaya S, Nishimura S, Yamamoto A, Sugita K, Nishimura T, Yagyu F *et al*: **Changing distribution of norovirus genotypes and genetic analysis of recombinant GIIb among infants and children with diarrhea in Japan**. *J Med Virol* 2006, **78**(7):971-978.
69. Phan TG, Takanashi S, Kaneshi K, Ueda Y, Nakaya S, Nishimura S, Sugita K, Nishimura T, Yamamoto A, Yagyu F *et al*: **Detection and genetic characterization of norovirus strains circulating among infants and children with acute gastroenteritis in Japan during 2004-2005**. *Clin Lab* 2006, **52**(9-10):519-525.
70. Dey SK, Phan TG, Mizuguchia M, Okitsua S, Ushijima H: **Novel recombinant norovirus in Japan**. *Virus Genes*, **40**(3):362-364.
71. Dey SK, Phan GT, Nishimura S, Mizuguchi M, Okitsu S, Ushijima H: **Molecular and epidemiological trend of sapovirus, and astrovirus infection in Japan**. *J Trop Pediatr*, **56**(3):205-207.
72. Phan TG, Yan H, Khamrin P, Quang TD, Dey SK, Yagyu F, Okitsu S, Muller WE, Ushijima H: **Novel intragenotype recombination in sapovirus**. *Clin Lab* 2006, **52**(7-8):363-366.
73. Gouvea V, Glass RI, Woods P, Taniguchi K, Clark HF, Forrester B, Fang ZY: **Polymerase chain reaction amplification and typing of rotavirus nucleic acid from stool specimens**. *J Clin Microbiol* 1990, **28**(2):276-282.
74. Ushijima H, Koike H, Mukoyama A, Hasegawa A, Nishimura S, Gentsch J: **Detection and serotyping of rotaviruses in stool specimens by using**

reverse transcription and polymerase chain reaction amplification. *J Med Virol* 1992, **38**(4):292-297.
75. Sen A, Kobayashi N, Das S, Krishnan T, Bhattacharya SK, Urasawa S, Naik TN: **Amplification of various genes of human group B rotavirus from stool specimens by RT-PCR**. *J Clin Virol* 2000, **17**(3):177-181.
76. Kuzuya M, Fujii R, Hamano M, Ogura H: **[Outbreak of acute gastroenteritis caused by human group C rotavirus in a youth educational center in Okayama Prefecture]**. *Kansenshogaku Zasshi* 2003, **77**(2):53-59.
77. Xu W, McDonough MC, Erdman DD: **Species-specific identification of human adenoviruses by a multiplex PCR assay**. *J Clin Microbiol* 2000, **38**(11):4114-4120.
78. Matsui M, Ushijima H, Hachiya M, Kakizawa J, Wen L, Oseto M, Morooka K, Kurtz JB: **Determination of serotypes of astroviruses by reverse transcription-polymerase chain reaction and homologies of the types by the sequencing of Japanese isolates**. *Microbiol Immunol* 1998, **42**(8):539-547.
79. Kojima S, Kageyama T, Fukushi S, Hoshino FB, Shinohara M, Uchida K, Natori K, Takeda N, Katayama K: **Genogroup-specific PCR primers for detection of Norwalk-like viruses**. *J Virol Methods* 2002, **100**(1-2):107-114.
80. Yan H, Yagyu F, Okitsu S, Nishio O, Ushijima H: **Detection of norovirus (GI, GII), Sapovirus and astrovirus in fecal samples using reverse transcription single-round multiplex PCR**. *J Virol Methods* 2003, **114**(1):37-44.
81. Gentsch JR, Glass RI, Woods P, Gouvea V, Gorziglia M, Flores J, Das BK, Bhan MK: **Identification of group A rotavirus gene 4 types by polymerase chain reaction**. *J Clin Microbiol* 1992, **30**(6):1365-1373.
82. Wang QH, Han MG, Cheetham S, Souza M, Funk JA, Saif LJ: **Porcine noroviruses related to human noroviruses**. *Emerg Infect Dis* 2005, **11**(12):1874-1881.
83. White PA, Hansman GS, Li A, Dable J, Isaacs M, Ferson M, McIver CJ, Rawlinson WD: **Norwalk-like virus 95/96-US strain is a major cause of gastroenteritis outbreaks in Australia**. *J Med Virol* 2002, **68**(1):113-118.
84. Saito H, Saito S, Kamada K, Harata S, Sato H, Morita M, Miyajima Y: **Application of RT-PCR designed from the sequence of the local SRSV strain to the screening in viral gastroenteritis outbreaks**. *Microbiol Immunol* 1998, **42**(6):439-446.
85. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0**. *Mol Biol Evol* 2007, **24**(8):1596-1599.
86. Kimura M: **A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences**. *J Mol Evol* 1980, **16**(2):111-120.
87. Kuntal BK, Aparoy P, Reddanna P: **EasyModeller: A graphical interface to MODELLER**. *BMC Res Notes*, **3**:226.
88. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller**. *Curr Protoc Bioinformatics* 2006, **Chapter 5**:Unit 5 6.
89. Eswar N, Ramakrishnan C: **Deterministic features of side-chain main-chain hydrogen bonds in globular protein structures**. *Protein Eng* 2000, **13**(4):227-238.

90. Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A: **Comparative protein structure modeling of genes and genomes**. *Annu Rev Biophys Biomol Struct* 2000, **29**:291-325.
91. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling**. *Electrophoresis* 1997, **18**(15):2714-2723.
92. Krieger E, Koraimann G, Vriend G: **Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field**. *Proteins* 2002, **47**(3):393-402.
93. Konagurthu AS, Whisstock JC, Stuckey PJ, Lesk AM: **MUSTANG: a multiple structural alignment algorithm**. *Proteins* 2006, **64**(3):559-574.
94. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM: **AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR**. *J Biomol NMR* 1996, **8**(4):477-486.
95. Hooft RW, Vriend G, Sander C, Abola EE: **Errors in protein structures**. *Nature* 1996, **381**(6580):272.
96. Bowie JU, Luthy R, Eisenberg D: **A method to identify protein sequences that fold into a known three-dimensional structure**. *Science* 1991, **253**(5016):164-170.
97. Colovos C, Yeates TO: **Verification of protein structures: patterns of nonbonded atomic interactions**. *Protein Sci* 1993, **2**(9):1511-1519.
98. Pontius J, Richelle J, Wodak SJ: **Deviations from standard atomic volumes as a quality measure for protein crystal structures**. *J Mol Biol* 1996, **264**(1):121-136.
99. Jin M, Xie HP, Duan ZJ, Liu N, Zhang Q, Wu BS, Li HY, Cheng WX, Yang SH, Yu JM *et al*: **Emergence of the GII4/2006b variant and recombinant noroviruses in China**. *J Med Virol* 2008, **80**(11):1997-2004.
100. Kittigul L, Pombubpa K, Taweekate Y, Diraphat P, Sujirarat D, Khamrin P, Ushijima H: **Norovirus GII-4 2006b variant circulating in patients with acute gastroenteritis in Thailand during a 2006-2007 study**. *J Med Virol*, **82**(5):854-860.
101. Chhabra P, Dhongade RK, Kalrao VR, Bavdekar AR, Chitambar SD: **Epidemiological, clinical, and molecular features of norovirus infections in western India**. *J Med Virol* 2009, **81**(5):922-932.
102. Buesa J, Montava R, Abu-Mallouh R, Fos M, Ribes JM, Bartolome R, Vanaclocha H, Torner N, Dominguez A: **Sequential evolution of genotype GII.4 norovirus variants causing gastroenteritis outbreaks from 2001 to 2006 in Eastern Spain**. *J Med Virol* 2008, **80**(7):1288-1295.
103. Onishi N, Hosoya M, Matsumoto A, Imamura T, Katayose M, Kawasaki Y, Hashimoto O, Hayashi A, Ishiko H, Suzuki H: **Molecular epidemiology of norovirus gastroenteritis in Soma, Japan, 2001-2003**. *Pediatr Int* 2008, **50**(1):65-69.
104. Ho EC, Cheng PK, Lau AW, Wong AH, Lim WW: **Atypical norovirus epidemic in Hong Kong during summer of 2006 caused by a new genogroup II/4 variant**. *J Clin Microbiol* 2007, **45**(7):2205-2211.
105. Reuter G, Krisztalovics K, Vennema H, Koopmans M, Szucs G: **Evidence of the etiological predominance of norovirus in gastroenteritis outbreaks--emerging new-variant and recombinant noroviruses in Hungary**. *J Med Virol* 2005, **76**(4):598-607.

106. Motomura K, Oka T, Yokoyama M, Nakamura H, Mori H, Ode H, Hansman GS, Katayama K, Kanda T, Tanaka T *et al*: **Identification of monomorphic and divergent haplotypes in the 2006-2007 norovirus GII/4 epidemic population by genomewide tracing of evolutionary history**. *J Virol* 2008, **82**(22):11247-11262.
107. Ramirez S, Giammanco GM, De Grazia S, Colomba C, Martella V, Arista S: **Emerging GII.4 norovirus variants affect children with diarrhea in Palermo, Italy in 2006**. *J Med Virol* 2009, **81**(1):139-145.
108. Ueki Y, Shoji M, Sato C, Sato Y, Okimura Y, Saito N, Otomo H, Onodera H: **Norovirus GII/4 variants observed in outbreaks of gastroenteritis in Miyagi prefecture between November and December of 2006**. *Jpn J Infect Dis* 2007, **60**(4):240-241.
109. Bull RA, Eden JS, Rawlinson WD, White PA: **Rapid evolution of pandemic noroviruses of the GII.4 lineage**. *PLoS Pathog*, **6**(3):e1000831.
110. Chung JY, Han TH, Park SH, Kim SW, Hwang ES: **Detection of GII-4/2006b variant and recombinant noroviruses in children with acute gastroenteritis, South Korea**. *J Med Virol*, **82**(1):146-152.
111. Nakamura K, Iwai M, Zhang J, Obara M, Horimoto E, Hasegawa S, Kurata T, Takizawa T: **Detection of a novel recombinant norovirus from sewage water in toyama prefecture, Japan**. *Jpn J Infect Dis* 2009, **62**(5):394-398.
112. Bull RA, Tanaka MM, White PA: **Norovirus recombination**. *J Gen Virol* 2007, **88**(Pt 12):3347-3359.
113. Asanaka M, Atmar RL, Ruvolo V, Crawford SE, Neill FH, Estes MK: **Replication and packaging of Norwalk virus RNA in cultured mammalian cells**. *Proc Natl Acad Sci U S A* 2005, **102**(29):10327-10332.
114. Bull RA, Hansman GS, Clancy LE, Tanaka MM, Rawlinson WD, White PA: **Norovirus recombination in ORF1/ORF2 overlap**. *Emerg Infect Dis* 2005, **11**(7):1079-1085.
115. Coyne KP, Reed FC, Porter CJ, Dawson S, Gaskell RM, Radford AD: **Recombination of Feline calicivirus within an endemically infected cat colony**. *J Gen Virol* 2006, **87**(Pt 4):921-926.
116. Muller HJ: **The Relation of Recombination to Mutational Advance**. *Mutat Res* 1964, **106**:2-9.
117. Cannon JL, Lindesmith LC, Donaldson EF, Saxe L, Baric RS, Vinje J: **Herd immunity to GII.4 noroviruses is supported by outbreak patient sera**. *J Virol* 2009, **83**(11):5363-5374.
118. Suzuki H, Sakai T, Tanabe N, Okabe N: **Peak rotavirus activity shifted from winter to early spring in Japan**. *Pediatr Infect Dis J* 2005, **24**(3):257-260.
119. Khananurak K, Vutithanachot V, Simakachorn N, Theamboonlers A, Chongsrisawat V, Poovorawan Y: **Prevalence and phylogenetic analysis of rotavirus genotypes in Thailand between 2007 and 2009**. *Infect Genet Evol*, **10**(4):537-545.
120. Wang YH, Kobayashi N, Zhou X, Nagashima S, Zhu ZR, Peng JS, Liu MQ, Hu Q, Zhou DJ, Watanabe S *et al*: **Phylogenetic analysis of rotaviruses with predominant G3 and emerging G9 genotypes from adults and children in Wuhan, China**. *J Med Virol* 2009, **81**(2):382-389.
121. Urasawa S, Urasawa T, Taniguchi K, Chiba S: **Serotype determination of human rotavirus isolates and antibody prevalence in pediatric population in Hokkaido, Japan**. *Arch Virol* 1984, **81**(1-2):1-12.

122. Zhou Y, Li L, Okitsu S, Maneekarn N, Ushijima H: **Distribution of human rotaviruses, especially G9 strains, in Japan from 1996 to 2000**. *Microbiol Immunol* 2003, **47**(8):591-599.
123. Yoshinaga M, Phan TG, Nguyen TA, Yan H, Yagyu F, Okitsu S, Muller WE, Ushijima H: **Changing distribution of group A rotavirus G-types and genetic analysis of G9 circulating in Japan**. *Arch Virol* 2006, **151**(1):183-192.
124. Linhares AC, Verstraeten T, Wolleswinkel-van den Bosch J, Clemens R, Breuer T: **Rotavirus serotype G9 is associated with more-severe disease in Latin America**. *Clin Infect Dis* 2006, **43**(3):312-314.
125. Adah MI, Rohwedder A, Olaleyle OD, Werchau H: **Nigerian rotavirus serotype G8 could not be typed by PCR due to nucleotide mutation at the 3' end of the primer binding site**. *Arch Virol* 1997, **142**(9):1881-1887.
126. Banyai K, Gentsch JR, Griffin DD, Holmes JL, Glass RI, Szucs G: **Genetic variability among serotype G6 human rotaviruses: identification of a novel lineage isolated in Hungary**. *J Med Virol* 2003, **71**(1):124-134.
127. Cunliffe NA, Gondwe JS, Graham SM, Thindwa BD, Dove W, Broadhead RL, Molyneux ME, Hart CA: **Rotavirus strain diversity in Blantyre, Malawi, from 1997 to 1999**. *J Clin Microbiol* 2001, **39**(3):836-843.
128. Espinola EE, Parra GI, Russomando G, Arbiza J: **Genetic diversity of the VP4 and VP7 genes affects the genotyping of rotaviruses: analysis of Paraguayan strains**. *Infect Genet Evol* 2008, **8**(1):94-99.
129. Fischer TK, Page NA, Griffin DD, Eugen-Olsen J, Pedersen AG, Valentiner-Branth P, Molbak K, Sommerfelt H, Nielsen NM: **Characterization of incompletely typed rotavirus strains from Guinea-Bissau: identification of G8 and G9 types and a high frequency of mixed infections**. *Virology* 2003, **311**(1):125-133.
130. Iturriza-Gomara M, Kang G, Gray J: **Rotavirus genotyping: keeping up with an evolving population of human rotaviruses**. *J Clin Virol* 2004, **31**(4):259-265.
131. Martella V, Terio V, Arista S, Elia G, Corrente M, Madio A, Pratelli A, Tempesta M, Cirani A, Buonavoglia C: **Nucleotide variation in the VP7 gene affects PCR genotyping of G9 rotaviruses identified in Italy**. *J Med Virol* 2004, **72**(1):143-148.
132. Rahman M, Sultana R, Podder G, Faruque AS, Matthijnssens J, Zaman K, Breiman RF, Sack DA, Van Ranst M, Azim T: **Typing of human rotaviruses: nucleotide mismatches between the VP7 gene and primer are associated with genotyping failure**. *Virol J* 2005, **2**:24.
133. Hansman GS, Ishida S, Yoshizumi S, Miyoshi M, Ikeda T, Oka T, Takeda N: **Recombinant sapovirus gastroenteritis, Japan**. *Emerg Infect Dis* 2007, **13**(5):786-788.
134. Phan TG, Okame M, Nguyen TA, Maneekarn N, Nishio O, Okitsu S, Ushijima H: **Human astrovirus, norovirus (GI, GII), and sapovirus infections in Pakistani children with diarrhea**. *J Med Virol* 2004, **73**(2):256-261.
135. Oh DY, Gaedicke G, Schreier E: **Viral agents of acute gastroenteritis in German children: prevalence and molecular diversity**. *J Med Virol* 2003, **71**(1):82-93.
136. Pang XL, Lee BE, Tyrrell GJ, Preiksaitis JK: **Epidemiology and Genotype Analysis of Sapovirus Associated with Gastroenteritis Outbreaks in Alberta, Canada: 2004-2007**. *J Infect Dis* 2008.

137. Pankovics P, Kugler Z, Katai A, Reuter G: **[First gastroenteritis outbreak caused by sapovirus (GI2) in Hungary - part of an international epidemic?]**. *Orv Hetil* 2009, **150**(26):1223-1229.
138. Harada S, Okada M, Yahiro S, Nishimura K, Matsuo S, Miyasaka J, Nakashima R, Shimada Y, Ueno T, Ikezawa S *et al*: **Surveillance of pathogens in outpatients with gastroenteritis and characterization of sapovirus strains between 2002 and 2007 in Kumamoto Prefecture, Japan**. *J Med Virol* 2009, **81**(6):1117-1127.
139. Ootsuka Y, Yamashita Y, Ichikawa T, Kondo R, Oseto M, Katayama K, Takeda N, Oka T: **Molecular characterization of sapoviruses detected in sporadic gastroenteritis cases in 2007 in Ehime Prefecture, Japan**. *Jpn J Infect Dis* 2009, **62**(3):246-248.
140. Svraka S, Vennema H, van der Veer B, Hedlund KO, Thorhagen M, Siebenga J, Duizer E, Koopmans M: **Epidemiology and genotype analysis of emerging sapovirus-associated infections across Europe**. *J Clin Microbiol*, **48**(6):2191-2198.
141. Katayama K, Miyoshi T, Uchino K, Oka T, Tanaka T, Takeda N, Hansman GS: **Novel recombinant sapovirus**. *Emerg Infect Dis* 2004, **10**(10):1874-1876.
142. Okada M, Yamashita Y, Oseto M, Shinozaki K: **The detection of human sapoviruses with universal and genogroup-specific primers**. *Arch Virol* 2006, **151**(12):2503-2509.
143. Chan MC, Sung JJ, Lam RK, Chan PK, Lai RW, Leung WK: **Sapovirus detection by quantitative real-time RT-PCR in clinical stool specimens**. *J Virol Methods* 2006, **134**(1-2):146-153.
144. Chen R, Neill JD, Noel JS, Hutson AM, Glass RI, Estes MK, Prasad BV: **Inter- and intragenus structural variations in caliciviruses and their functional implications**. *J Virol* 2004, **78**(12):6469-6479.
145. Lindesmith LC, Donaldson E, Leon J, Moe CL, Frelinger JA, Johnston RE, Weber DJ, Baric RS: **Heterotypic humoral and cellular immune responses following Norwalk virus infection**. *J Virol*, **84**(4):1800-1815.
146. Kumar S, Ochoa W, Kobayashi S, Reddy VS: **Presence of a surface-exposed loop facilitates trypsinization of particles of Sinsiro virus, a genogroup II.3 norovirus**. *J Virol* 2007, **81**(3):1119-1128.
147. Nayak MK, Balasubramanian G, Sahoo GC, Bhattacharya R, Vinje J, Kobayashi N, Sarkar MC, Bhattacharya MK, Krishnan T: **Detection of a novel intergenogroup recombinant Norovirus from Kolkata, India**. *Virology* 2008, **377**(1):117-123.
148. Tan M, Jiang X: **Norovirus gastroenteritis, carbohydrate receptors, and animal models**. *PLoS Pathog*, **6**(8).

APPENDIX

Table 11. Results of structural validation of GII.14 and GII.6 models obtained from EasyModeller.

Sever	Property	Model	
		GII.14	**GII.6**
PROCHECK	Residues in most favoured regions	90.7%	89.2%
	Residues in allowed regions	9.3%	10%
	Residues in generously regions	0	0.4%
	Residues in disallowed regions	0	0.4%
	Main chain	6	6
	Side chain	5	5
	Bond length (in limits)	98.9%	98.5%
	Bond angle (in limits)	92.8%	91.7%
	Planar groups (in limits)	100%	100%
WHATCHECK	RMS Z-score for bond lengths	0.93	0.96
	RMS Z-score for bond angles	1.231	1.261
	Omega angle restraints	0.651	0.658
	Side chain planarity	0.314	0.318
	Improper dihedral distribution	0.97	1.009
	Outside distribution	1.087	1.063
	2nd generation packing quality	-1.906	-2.128
	Ramachandran plot appearance	-0.228	-0.47
	Chi-1/Chi-2 rotamer normality	0.946	0.181
	Backbone conformation	-4.931	-4.861
VERIFY 3D	3D-1D score>0.2	93.11%	94.01%
ERRAT	Overall quality factor	75.338	72.425
PROVE	Z-Score Mean	1.940	1.152
	Z-Score std deviation	71.736	29.396
	RMS Z-score	71.763	29.4.6

```
                10        20        30        40        50
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A TKPFTVPILTVEEMSNSRFPIPLEKLYTGPSSAFVVQPQNGRCTTDGVLL
8610   TKNFTLPVLRVSEMTNSRFPVVLDQMYTSRNENIIVQPQNGRCTTDGELL

                60        70        80        90        100
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A GTTQLSAVNICTFRGDVTHIA-GSHDYIMNLASQNWNNYDPTEEIPAPLG
8610   GTTILQSVSICNFKGTMQAKLNEEPRYQLQLTNLDGSPIDPTDDMPAPLG

                110       120       130       140       150
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A TPDFVGKIQGMLTQTTREDGSTRAHKATVSTGSVHFTPKLGSVQYTTDTN
8610   TPDFQAMLYGVASQRSSIDNATRAHDAQIDTAGDTFAPKIGQVRFKSSSN

                160       170       180       190       200
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A NDLQTGQNTKFTPVGVIQDGNNHQNEPQQWVLPNYSGRTGHNVHLAPAVA
8610   -DFDLHDPTKFTPIGVNVD-DQ--HPFRQWSLPNYGGHLALNNHLAPAVT

                210       220       230       240       250
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A PTFPGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFYQEAAPAQSDVA
8610   PLFPGEQILFFRSYIPSAGGHTDGAMDCLLPQEWVEHFYQEAAPSQSDIA

                260       270       280       290       300
     ....|....|....|....|....|....|....|....|....|....|....|
2OBS_A LLRFVNPDTGRVLFECKLHKSGYVTVAHTGPHDLVIPPNGYFRFDSWVNQ
8610   LVRFINPDTGRVLFEAKLHKQGFLTIAASGDHPIVMPTNGYFRFEAWVNP

     ....|...
2OBS_A FYTLAPMG
8610   FYTLAPVG
```

Figure 28. Sequence alignment between GII.14 norovirus P domain (strain 8610) and template VA387 GII.4 (PDB code: 2OBS chain A).

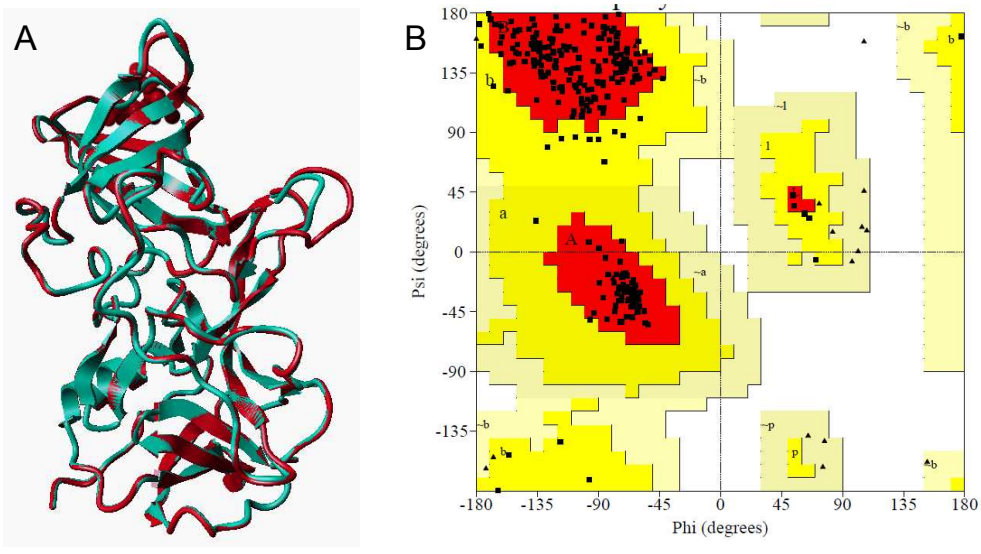


Figure 29. (A) Superimposition of GII.14 (cyan) and template VA387 GII.4 (red) structure. (B) Ramachandran plot analysis of the modeled structure.

```
                10        20        30        40        50
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    TKPFTVPILTVEEMSNSRFPIPLEKLYTGPSSAFVVQPQNGRCTTDGVLL
8915      TKPFTLPILTLGELSNSRFPAPIDMLYPDPNEGIVVQPQNGRCTLDGTLQ

                60        70        80        90        100
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    GTTQLSAVNICTFRGDVTHIAG----------SHDYIMNLASQNWNN
8915      GTTQLVPTQICAFRGTLIGQTSRSSDSTDSAPRRRDHPLHVQLKNLDGTQ

                110       120       130       140       150
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    YDPTEEIPAPLGTPDFVGKIQGMLTQTTRE---DGSTRAHKATVSTGSVH
8915      YDPTDEVPAVLGAIDFKGTVFGVASQRDVSGQQVGATRAHEVHINTTDPR

                160       170       180       190       200
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    FTPKLGSVQYTTDTNNDLQTGQNTKFTPVGVIQDGNNHQNEPQQWVLPNY
8915      YTPKLGSILMHSESD-DFVTGQPVRFTPIGMGD------NDWHQWELPDY

                210       220       230       240       250
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    SGRTGHNVHLAPAVAPTFPGEQLLFFRSTMPGCSGYPNMNLDCLLPQEWV
8915      SGHLTLNMSLAPAVAPAFPGERILFFRSIVPSAGGYGSGYIDCLIPQEWV

                260       270       280       290       300
          ....|....|....|....|....|....|....|....|....|....|
2OBT_A    QHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTVAHTGPHDLV
8915      QHFYQEAAPSQSAVALIRYVNPDTGRNIFEAKLHREGFITVANSGNNPIV

                310       320
          ....|....|....|....|...
2OBT_A    IPPNGYFRFDSWVNQFYTLAPMG
8915      VPPNGYFRFEAWVNQFYTLTPMG
```

Figure 30. Sequence alignment between GII.6 norovirus P domain (strain 8915) and template VA387 GII.4 (PDB code: 2OBT chain A).

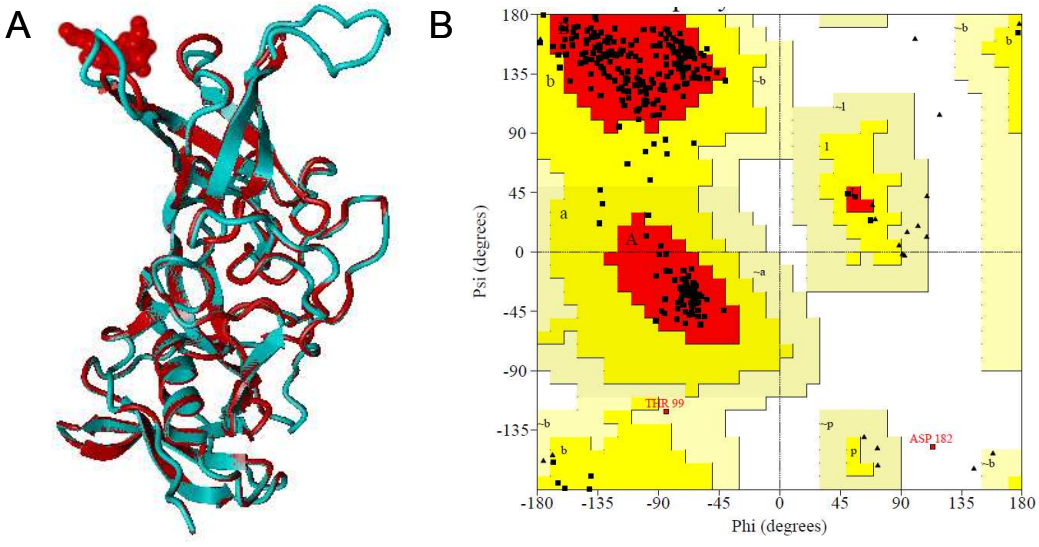


Figure 31. (A) Superimposition of GII.6 (cyan) and template VA387 GII.4 (red) structure. (B) Ramachandran plot analysis of the modeled structure.