MAJOR ARTICLE

Norovirus Illness Is a Global Problem: Emergence and Spread of Norovirus GII.4 Variants, 2001–2007

J. Joukje Siebenga,[1,2] Harry Vennema,[1] Du-Ping Zheng,[3] Jan Vinjé,[3] Bonita E. Lee,[4] Xiao-Li Pang,[4] Eric C. M. Ho,[5] Wilina Lim,[5] Avinash Choudekar,[7] Shobha Broor,[7] Tamar Halperin,[8] Nassar B. G. Rasool,[9] Joanne Hewitt,[10] Gail E. Greening,[10] Miao Jin,[6] Zhao-Jun Duan,[6] Yalda Lucero,[11] Miguel O'Ryan,[11] Marina Hoehne,[12] Eckart Schreier,[12] Rodney M. Ratcliff,[13] Peter A. White,[14] Nobuhiro Iritani,[15] Gábor Reuter,[16] and Marion Koopmans[1,2]

[1]National Institute for Public Health and the Environment, Bilthoven, and [2]Erasmus Medical Centre, Rotterdam, the Netherlands; [3]Centers for Disease Control and Prevention, Atlanta, Georgia, United States; [4]Provincial Public Health Laboratory (Microbiology), Edmonton, Alberta, Canada; [5]Centre for Health Protection, Hong Kong, and [6]Department of Viral Diarrhea, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China; [7]All India Institute of Medical Sciences, New Delhi, India; [8]Central Laboratory, Israel Defense Force Medical Corps, Israel; [9]University of Malaya, Kuala Lumpur, Malaysia; [10]Institute of Environmental Science and Research, Porirua, New Zealand; [11]Institute of Biomedical Sciences, University of Chile, Santiago, Chile; [12]Robert Koch Institute, Berlin, Germany; [13]Institute of Medical and Veterinary Science and School of Molecular and Biomedical Science, University of Adelaide, Adelaide, and [14]University of New South Wales, Sydney, Australia; [15]Osaka City Institute of Public Health and Environmental Sciences, Osaka, Japan; [16]ÁNTSZ Regional Institute of State Public Health Service, Pécs, Hungary

Background. Noroviruses (NoVs) are the most common cause of viral gastroenteritis. Their high incidence and importance in health care facilities result in a great impact on public health. Studies from around the world describing increasing prevalence have been difficult to compare because of differing nomenclatures for variants of the dominant genotype, GII.4. We studied the global patterns of GII.4 epidemiology in relation to its genetic diversity.

Methods. Data from NoV outbreaks with dates of onset from January 2001 through March 2007 were collected from 15 institutions on 5 continents. Partial genome sequences ($n = 775$) were collected, allowing phylogenetic comparison of data from different countries.

Results. The 15 institutions reported 3098 GII.4 outbreaks, 62% of all reported NoV outbreaks. Eight GII.4 variants were identified. Four had a global distribution—the 1996, 2002, 2004, and 2006b variants. The 2003Asia and 2006a variants caused epidemics, but they were geographically limited. Finally, the 2001Japan and 2001Henry variants were found across the world but at low frequencies.

Conclusions. NoV epidemics resulted from the global spread of GII.4 strains that evolved under the influence of population immunity. Lineages show notable (and currently unexplained) differences in geographic prevalence. Establishing a global NoV network by which data on strains with the potential to cause pandemics can be rapidly exchanged may lead to improved prevention and intervention strategies.

Noroviruses (NoVs) are the leading cause of acute viral gastroenteritis worldwide. People of all ages are affected, but outbreaks are most often reported in health care settings (such as nursing homes and hospitals), where infections in high-risk groups (such as elderly and immunocompromised people) can have a serious impact by causing prolonged morbidity and mortality [1, 2]. Outbreaks are difficult to control and may lead to considerable economic costs resulting from closure of wards, increased length of hospitalization, hiring of extra personnel, and use of extra supplies [3, 4].

Received 16 February 2009; accepted 10 April 2009; electronically published 23 July 2009.

Reprints or correspondence: J. Joukje Siebenga, RIVM, LIS-VIR, Postbus 1, 3720 BA Bilthoven, the Netherlands (Joukje.Siebenga@RIVM.nl).

The Journal of Infectious Diseases 2009; 200:802–12

0022-1899/2009/20005-0019$15.00
DOI: 10.1086/605127

Potential conflicts of interest: none reported.

Presented in part: Third International Calicivirus Meeting, Cancun, Mexico, 10–13 November 2007 (abstract S3-6).

Financial support: European Commission, Directorate General Research Quality of Life program (EVENT [Enteric Virus Emergence, New Tools] grant SP22-CT-2004-502571 to J.J.S.); Indian Council of Medical Research, New Delhi, India (support to S.B. and A.C.); New Zealand Ministry of Health (funding for the New Zealand studies); Hungarian Scientific Research Fund (grant F048433 to G.R.); 973 National Key Basic Research Program of China (grant 2007CB310500 to the Chinese Center for Disease Control and Prevention); Chilean National Science and Technology Research Fund (grants 10210583 and 1061079 to Y.L. and M.O.); Alberta Health and Wellness (support to B.E.L. and X.-L.P.); Ministry of Science, Technology, and Innovation, Malaysia (grant 06-02-05-001 BTK/ER/002 to N.B.G.R.).

The incidence of NoV illness is high. A population study of gastroenteritis in the Netherlands estimated that 500,000 NoV cases occurred among a population of 15.7 million in 1999 (a nonepidemic year) [5]. A study from England (1993–1996) reported that NoVs were the most commonly identified pathogen in cases of intestinal infectious disease [6]. On the basis of these and historical data, the prevalence in the United States was estimated to be 23 million cases annually [7]. In developing countries, where diarrhea is a leading cause of death in young children [8], relatively little is known about the etiological role played by NoVs, but estimates indicating that >1.1 million hospitalizations and almost 220,000 deaths occur among children <5 years old each year in such countries have been published recently [9].

NoVs are extremely contagious, owing to a very low infectious dose (estimated median infectious dose, ~18 viral particles) [10] and to levels of shedding $>1 \times 10^{10}$ RNA copies per gram of stool [11, 12]. Additionally, although patients usually recover from symptoms within 2 or 3 days, shedding may last several weeks [11]. Transmission occurs through the fecal-oral route, either by direct contact with infected individuals or via contaminated surfaces, food, or water. No or limited long-term immunity results from infection [13, 14], so one person may be repeatedly infected. Nevertheless, short-term immunity combined with the high prevalence of NoVs amounts to herd immunity [15, 16].

NoVs belong to a highly genetically and antigenically diverse genus of the family Caliciviridae. They segregate into 5 genogroups. In genogroup I, 8 genotypes are currently recognized; in genogroup II, 19 are recognized [17]. Viruses of the GII.4 genotype have been predominant during the past decade in the United States, Europe, and Oceania, causing 70%–80% of all NoV outbreaks (particularly in health care settings) [18]. In 2002, the number of reported NoV outbreaks increased sharply in many countries, followed by 4 epidemic winters in 2002–2003, 2004–2005, 2006–2007, and 2007–2008 in the Northern Hemisphere. During these seasons, incidence rates were likely much higher than reported in the previously mentioned population studies, which were all conducted before 2002. Genetic analyses showed that these epidemics coincided with the emergence of novel GII.4 variants except for the 2007–2008 epidemic, which was a continuation of the 2006–2007 epidemic [15, 19–26]. The global molecular epidemiology of GII.4 lineages has not been systematically evaluated; reports describing the molecular epidemiology of the GII.4 variants have focused on local or regional studies. Furthermore, the lack of a unified nomenclature for GII.4 variants has hampered comparison of the available data. Thus, the extent to which reported epidemics were truly global remains unclear.

The growing awareness and knowledge of both the scale and impact of NoV epidemiology raised the question as to how patterns suggested by molecular data analyses matched observations from public health surveillance across the world. We also wanted to create a common data set for NoV strains to function as a starting point for the newly established global NoV collaboration network, NoroNet. This global network aims at limiting the impact and scale of future NoV epidemics, for which monitoring of circulating strains is essential. Here, we present a global overview of the molecular epidemiology of GII.4 NoVs for the period from January 2001 through March 2007. During this period, 4 of the described GII.4 variants had a truly global distribution, whereas 4 had a different distribution or prevalence. Our results demonstrate that emergent GII.4 NoV strains spread throughout the world very rapidly and, therefore, that epidemiology needs to be assessed on a global scale in order to further the development of the basic knowledge that will underpin future research and prevention and control measures.

METHODS

Epidemiologic data. Data obtained from January 2001 through March 2007 were included in the present study. Institutions on 5 continents were contacted for participation. Selected institutions were included because they had structured NoV surveillance programs that included the molecular characterization of outbreak strains ($n = 10$) or because they provided data on geographic areas that could not otherwise be evaluated ($n = 6$).

The aggregated data set contained monthly reports describing all reported NoV outbreaks attributable to each GII.4 variant. NoV outbreaks were identified according to criteria used by each institution. Epidemiologic data describing the population covered, the number of outbreaks reported annually, the number of confirmed NoV outbreaks, and the percentage of GII.4 outbreaks were collected.

Sequence data. Each participating institution provided sequences from representative strains detected during the beginning, middle, and end of the circulation periods of each variant for genotyping; sequences obtained were from the preferred genomic regions of the strains (regions A, C, D, and E) (figure 1) [27]. These data enabled comparison of the nomenclature used by the participating institutions and in the literature (table 1), standardization of assignment of variant names, a check of the quality of the submitted data, and phylogenetic comparison of sequences detected in different geographic regions of the world. As the basis for variant assignment, we used the recently published amino acid sequences of the complete VP1 [15]. This method defines variants on the basis of phylogenetic clustering combined with epidemiologic patterns. Variant names include the first reported year of detection supplemented by a geographic region or suffix when necessary. Table 1 summarizes these names in addition to names used in other publications.

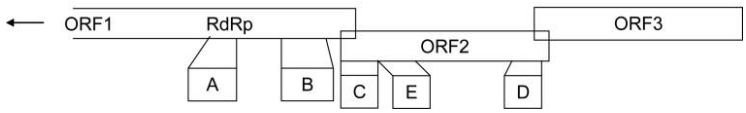

Figure 1. Schematic representation of the locations of the genomic regions of norovirus used for genotyping. Adapted from Vinjé et al [27]. Region B was not used for the analyses in the present study because it does not enable discrimination between GII.4 variants. ORF, open reading frame; RdRp, RNA-dependent RNA polymerase.

Data analyses. Variant assignment was done by the participating institutions, using either their own referencing systems, a tool offered by the Netherlands National Institute for Public Health and the Environment (RIVM) for regions A and C (Norovirus Quicktyping; available at: http://www.rivm.nl/bnwww), or a set of reference sequences compiled for this purpose that spanned the continuous genomic region covering regions A, C, E, and D.

Epidemiologic and sequence data were stored and analyzed in Microsoft Excel and BioNumerics software (version 4.61; Applied Maths). Additional reference sequences were obtained from GenBank. At RIVM, phylogenetic analysis of all sequences was performed using PhyML software (version 2.4.4) [33].

RESULTS

Participants. Table 2 lists the participants of the present study and outlines their surveillance methods. Overall, most reported outbreaks occurred in health care settings. In the countries of most participating institutions, it was mandatory to report outbreaks caused by any pathogen in health care settings. Ten institutions had an ongoing passive NoV surveillance program (table 2, top section). The Provincial Public Health Laboratory, Canada, had limited surveillance data for 2001; therefore, this year was not included. Five additional institutions, plus a targeted population study from the University of Chile, that did not have a structured NoV surveillance program provided target-study data and sequences with detection dates (table 2, bottom section). The University of Malaya, Malaysia, had different target studies mainly involving hospitalized children. The Israel Defense Force Medical Corps Central Laboratory provided data on NoV outbreaks in military units in Israel [42]. The Institute of Biomedical Sciences, University of Chile, submitted data from intermittent population surveillance and targeted population studies but usually without genotyping data [38]. The All India Institute of Medical Sciences provided data from a study comprising 4 different study groups [41].

Epidemiology of NoV outbreaks, by geographic area. In total, 4988 NoV outbreaks were reported, of which 3089 (62%) were confirmed to be GII.4 outbreaks (figure 2). Winter seasonality was seen in almost all geographic areas analyzed and was largely associated with elevated numbers of GII.4 outbreaks.

Data were comparable among the European countries, although the 2002–2003 and 2004–2005 seasons in Hungary were not as pronounced as those in Germany and the Netherlands. Germany and the Netherlands reported an off-seasonal peak in April and May 2002 that preceded the epidemic winter of 2002–2003 [32].

Similar summer peaks were seen more widely in 2006, caused by 2 emerging variants: 2006a and 2006b. These peaks were observed in the Netherlands; Germany; Hungary [23]; Australia; Alberta, Canada; Japan; and Hong Kong. In Japan and Hong Kong, the peaks were caused by the 2006b variant only. In Hong Kong, winter peaks started earlier (September–October) than in the other analyzed regions of the Northern Hemisphere (November–December), and the year 2006 in Hong Kong was marked by many outbreaks, with those occurring during the 2005–2006 winter being only partly related to GII.4 outbreaks. In Japan, the number of reported NoV outbreaks increased gradually, and, especially during the 2006–2007 winter, the proportion of GII.4 viruses was very high (almost 100%). Unlike most other countries, the United States reported many GII.4 outbreaks during the 2003–2004 winter.

The Southern Hemisphere countries Australia and New Zealand had different pictures. Interestingly, winter seasonality was absent from New Zealand. In the Adelaide area of Australia, the frequency of NoV outbreaks increased 5–10-fold during the 2006 epidemic, compared with that during the previous outbreak years of 2002 and 2004. This increase did not coincide with alterations in the surveillance scheme.

Sequence analysis and phylogeny. A total of 781 sequences were analyzed, including sequences of 19 strains from the public databases of Brazil ($n = 12$), Malawi ($n = 1$), and Ghana ($n = 6$) (to represent geographic areas that were not covered by surveillance) and 11 reference sequences. Twenty-four of the submitted sequences were too short for reliable analysis and were excluded from further analysis.

Full capsid sequences and partial sequences from regions A, C, E, and D (figure 1) were phylogenetically analyzed (figure 3*A* and 3*B*). Discriminatory power was not sufficient to distinguish different variants in region B; therefore, sequences from region B were not collected. Partial genome sequences were assigned a variant type on the basis of similarity with complete capsid sequences and associated partial polymerase sequences. Clustering into variants was consistent, as illustrated for region C sequences in figure 3*B*, except for the 2001Henry variant. Although the capsid sequence for this variant was

Table 1. Nomenclature for GII.4 Variants

GenBank reference strain	Epidemic season	Name used in present article	First citation/ full description	Other names
X86557	Before 1995	…	…	Lordsdale, GII/4 b [28]
AF145896	Before 1995	…	…	Camberwell, GII.4-1987 [16], GII/4-a [21], GII/4 c [28]
X76716	Before 1995	…	…	Bristol, GII/4-a [21], GII/4 b [28]
AJ004864	1995–1996	**1996**	[30]	Grimsby, Burwash Landing, GII.4-1997 [16], GII/4-b [21], GII/4 g [28]
AB294779	NA	**2001Japan**	[21]	GII/4-c [21], GII/4 a [28]
EU310927	NA	**2001Henry**	[31]	Houston
AY485642	2002–2003	**2002**	[32]	Farmington Hills, GII/4-d [21], GII/4 e [28]
AB220922	NA	**2003Asia**	[1]	Sakai, GII.4-2005 [16]
AY883096	2004–2005	**2004**	[19]	Hunter, GII/4 f [28]
EF126963, EF126964	2006–2007	**2006a**	[24]	Laurens, V4 [29][a]
EF126965, EF126966	2006–2007	**2006b**	[24]	Minerva, Den Haag [16], GII/4-e [21], GII/4-f [21], Kobe034, V6 [29][a]

NOTE. Reference strains were the first complete capsid sequences of each variant submitted to the public databases. NA, not applicable.

[a] Not all divisions made by Gallimore et al, which were based on motifs, fit the phylogenetic grouping.

clearly distinct, partial sequences clustered with other variants. As a result, discrimination of this strain may not have been accurate, and assessing its prevalence was difficult. Geographic trends within variant clusters were not identified. Analyses of regions A, E, and D resulted in identical clustering (data not shown).

Eight GII.4 variants were identified (table 1 and figure 3*A*); the 5 main variants were 1996, 2002, 2004, 2006a, and 2006b, and the 3 minor variants were 2003Asia, 2001Japan, and 2001Henry. The minor variants were detected at multiple geographic locations but at low frequencies. The variant 2003Asia is a recombinant, with an open reading frame 1 sequence belonging to GII.12 (Wortley-like) and open reading frame 2 and 3 sequences belonging to GII.4. This strain caused many outbreaks in Asia; however, in Hong Kong it was detected only among sporadic NoV cases, not in outbreaks. In Hong Kong, 993 sporadic NoV cases were analyzed, of which 58 (6%) were 2003Asia strains detected from late 2002 through 2005. The peak prevalence of 2003Asia in sporadic cases in Hong Kong was observed between January 2004 and January 2005 (15% [33/223] of cases; data not shown). The 2001Japan sublineage showed the most genetic resemblance to the Bristol and Lordsdale reference strains. The 2001Henry variant was identified in the United States and China (DQ364459), but prevalence was low.

All variants from the 2002 variant onward (ie, the 2001Henry, 2003Asia, 2002, 2004, 2006a, and 2006b variants) had a 1-aa insertion (at position 393) in the P2 domain of the capsid.

Quality assurance. The accuracy of variant assignment by participants was independently verified by phylogenetic analysis of submitted sequences. Of the 775 submitted sequences, 767 (99%) had been assigned to the correct variant. Furthermore, sequences of 2 separate genomic regions had been submitted for a subset of 81 outbreaks, and, with the exception of 5 (6%), they consistently belonged to the same variant in both regions. The 5 remaining strains may have been recombinants, or 2 different viruses may have been present in 1 outbreak.

Epidemiology of GII.4 variants. Epidemiologic analysis of the GII.4 variants was done using the 3089 submissions for which collection dates had been provided (figure 4). In 2001–2002, the 1996 variant caused significant numbers of outbreaks in Europe and Hong Kong but was rare in other geographic areas. The 2002, 2004, and 2006b variants caused epidemics in all areas with population surveillance, but there was a marked difference in the magnitude of the seasonal peaks. The 2002 variant heavily affected Europe, the United States, and Canada [32] but was observed less frequently in Asia and Oceania, despite significant numbers of reported outbreaks. During the same period, high numbers of GII.4 strains were reported in New Zealand, but they could not be assigned to a GII.4 variant because they were sequenced in region B only. Variants 2004 and 2006b were truly global and were prevalent in all analyzed geographic areas; 2006a is closely related phylogenetically to the 2004 variant, but both the complete capsid sequence and its distinct epidemiology (with clearly discernible epidemic peaks in 2004–2005 and 2006–2007) justify its classification as a separate variant [15]. Interestingly, 2006a was rarely reported by participants in Asia—1 outbreak was reported in Japan, and no sequences belonging to the 2006a variant were submitted by other participants in Asia (China, Japan, Hong Kong, India, and Malaysia [not all data shown]).

The minor variants were not confined to a single geographic area. For example, the 2001Japan variant has been detected in Japan, the Netherlands, Malawi (GenBank), and Chile, and the 2003Asia variant has been detected in China, Japan [1, 43], Hong Kong, New Zealand, and the United Kingdom [29]—although rarely in the latter two and, interestingly, only in sporadic cases but not in outbreaks in Hong Kong.

Pilot studies in geographic areas with no structured outbreak surveillance data. Data from participants with targeted

Table 2. Participating Institutions and Brief Description of Their Surveillance Setup

Category, institution	Geographic area	Coverage[a]	Typing region[b]	S/O[c]	Publication(s)	Note
Participants with population-level surveillance						
National Institute for Public Health and the Environment	Netherlands	Population, 100% Input mainly HC	A	O	[15, 26]	
Centers for Disease Control and Prevention	United States	Population, 100% Input mainly HC	C, D, E	O	[34, 35]	
Provincial Public Health Laboratory	Alberta, Canada	Population, 100% Input mainly HC	E	O		No data for 2001
Osaka City Institute of Public Health and Environmental Sciences	Osaka Prefecture, Japan	Population, 100%	C	O	[36]	
Centre for Health Protection	Hong Kong, China	Population, 100%	A, C since 2005	O, S	[37]	
Institute of Environmental Science and Research	New Zealand	Population, 100% Input mainly HC	B until 2004, D	O		
Institute of Biomedical Sciences, University of Chile	Santiago, Chile	100%	A, C	O	[38]	No data for 2004–2005, no genotyping for 2006
Regional Institute of State Public Health Service	Hungary	Population, 100% Input mainly HC	A	O	[22, 23]	
Robert Koch Institute	Germany	Population, 100% Input mainly HC	A	O	[39, 40]	
Institute of Medical and Veterinary Science	Adelaide, Australia	Population, 100% Input mainly HC	C, D	O		
		Target population				
Participants with targeted population studies						
University of New South Wales	New South Wales, Australia	All	C	O	[19, 24, 25]	
Department of Viral Diarrhea, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention	Various provinces and cities in China	Children <5 years old with acute gastroenteritis	A, C since 2006	S		
All India Institute of Medical Sciences	India	Various (including children's home, hospital, health center)	A	S	[41]	
University of Chile	Chile	Islands (2003) and cruise ship (2005)	A	S		
Central Laboratory, Israel Defense Force Medical Corps	Israel	Military	D	O	[42]	
University of Malaya	Malaysia	Hospitalized children (0–6 years old), 1 adult	E, D	S	Rasool et al (unpublished)	

[a] HC indicates health care settings, including hospitals, nursing homes, and long-term-care facilities.
[b] Genomic region used for typing detected norovirus strains (see figure 1).
[c] Indicates whether the institution reports sporadic cases (S), outbreaks (O), or both.

population studies—including participants in Australia, China, India, Chile, Israel, and Malaysia (table 2, bottom section), as well as GenBank sequences detected in Brazil, Malawi, and Ghana—were included to establish a more comprehensive overview of the global GII.4 variant distribution. The data thus obtained fit the overall picture of common GII.4 variants except for the data from Chile and China. Seven of 10 Chilean strains belonged to the 2001Japan variant. The sequences from China ($n = 97$) (figure 4) represented studies of children <5 years old and showed no evidence of epidemics caused by the 2002, 2004, or 2006a variants. However, a high number of 2003Asia strains were identified, especially in 2004 and 2005, and a high number of 2006b strains were identified in 2006.

Timeline of GII.4 variant emergence and spread over different continents. Figure 5 shows accumulated outbreaks per variant per continent. For the 2002 variant, the off-seasonal peak

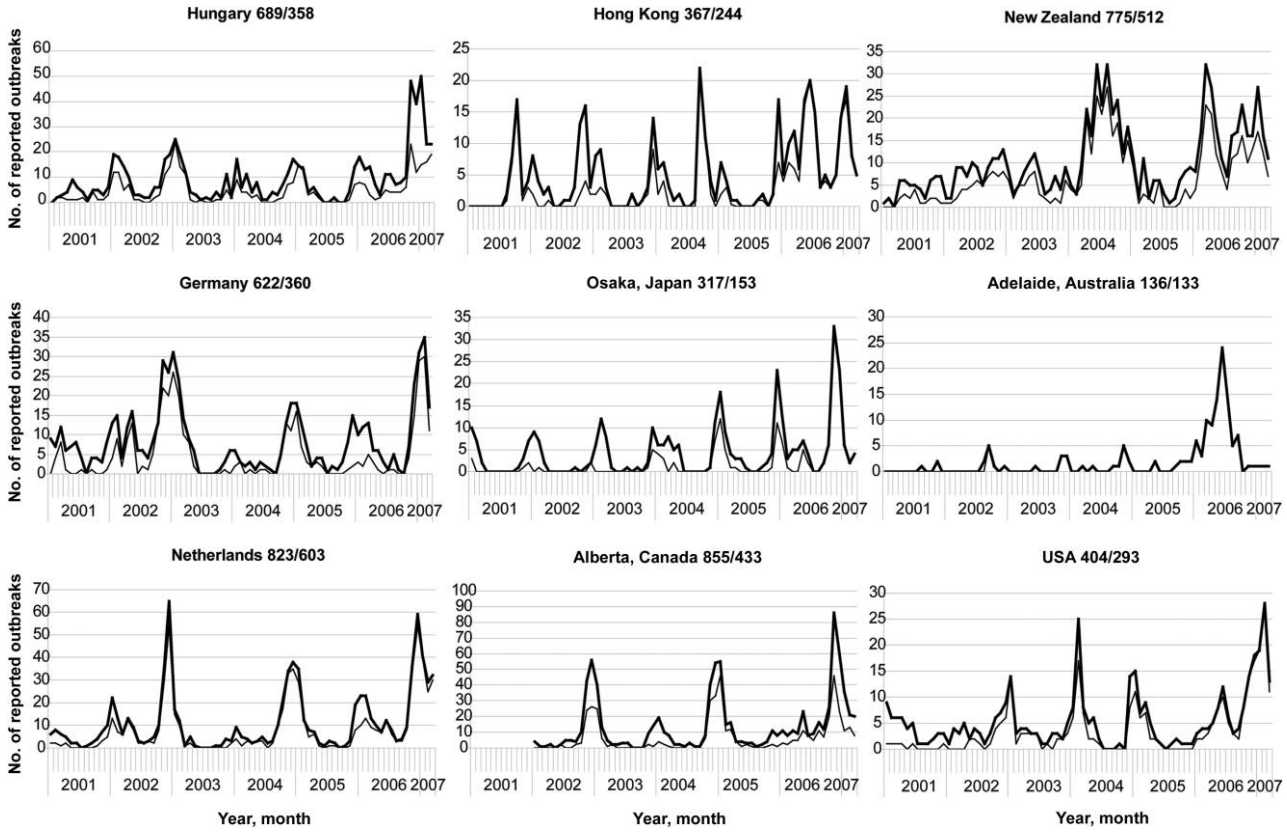

Figure 2. Monthly numbers of reported norovirus outbreaks (*thick line*) and GII.4 outbreaks (*thin line*), by geographic area. Note the different scales for the *Y*-axes. The total numbers of reported norovirus outbreaks (*left of virgule*) and the total numbers of GII.4 outbreaks (*right of virgule*) for each area are given above the graphs.

in Europe was seen first, followed by the epidemic winter in Oceania (although the latter is difficult to observe in the graphs because of the low numbers relative to those for the other continents). Three variants emerged first in Oceania or Asia—the 2004 and 2006a variants clearly first emerged in Oceania, and the 2006b lineage caused high numbers of outbreaks first in Asia (Hong Kong), followed by an off-seasonal peak in Europe and then Oceania.

DISCUSSION

This article represents the first effort to describe the molecular epidemiology of GII.4 NoVs on a global scale. We have demonstrated that GII.4 strains have been dominant all over the world during the past 7 years and were responsible for an increased number of reported NoV outbreaks globally. The impact of epidemic seasons was reflected in the outbreak-based surveillance, which focused on health care settings, although outbreaks also occurred in other places where close human contact occurs (such as military establishments, cruise ships, a hurricane Katrina refugee complex, and schools) [42, 44, 45].

Although the emergence of the 1996 variant falls outside the time range of this study, it was reported on all continents and likely had global coverage [20]. From the 1996 variant onward 8 distinct variants were identified, of which the 2002, 2004, and 2006b variants caused global epidemics. They displaced their predecessors rapidly and completely, although the 2006b variant initially cocirculated with the 2006a variant before continuing to cause the majority of outbreaks during the winter of 2007–2008, whereas the 2006a variant decreased in incidence but remained present (data not shown; personal communication with participating institutions). The 2006a and 2003Asia variants did cause epidemics but not in all analyzed geographic areas, whereas outbreaks of 2001Japan and 2001Henry occurred sporadically although at multiple geographic locations. Additional minor variants may have circulated but not been identified because of low frequencies. We have not proposed a nomenclature system for future variant naming here because a proposal for NoV nomenclature is currently being prepared.

Successive GII.4 variants differ antigenically and in their patterns of host cell binding, as shown by detailed molecular characterization of systematically collected outbreak data, the elucidated structure of the capsid protein, and binding experiments [15, 16]. The data presented here suggest that NoV GII.4 strains evolved and spread in a manner similar to that of in-

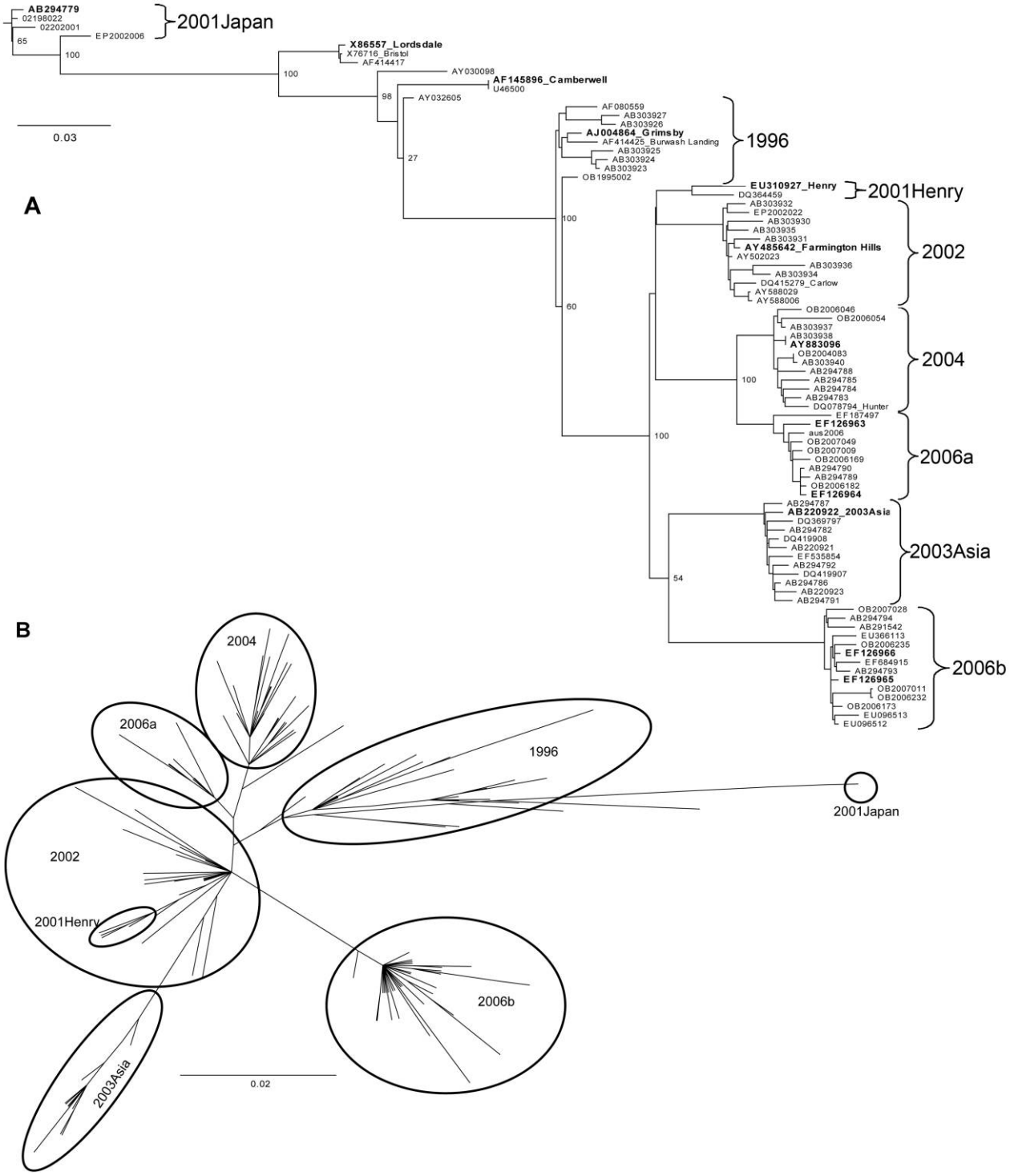

Figure 3. *A*, Phylogenetic tree of complete capsid sequences, constructed by means of PhyML software (49 sequences, 1623 nt, SYM+I+Γ_4 [symmetric, with allowance for invariant sites and gamma substitution rate variation], 1000 bootstraps). Reference strains (first complete capsid sequences submitted to the public databases) are shown in boldface type. The model was selected using MrAIC software (J. A. A. Nylander, 2004. MrAIC.pl. Program distributed by author. Evolutionary Biology Centre, Uppsala University). *B*, Phylogenetic tree of all available region C sequences, constructed by means of PhyML software (301 sequences, 266 nt, SYM+I+Γ_4). Strains detected in Australia, China, Germany, Ghana, Hong Kong, Hungary, Japan, the Netherlands, New Zealand, and the United States were included in this tree.

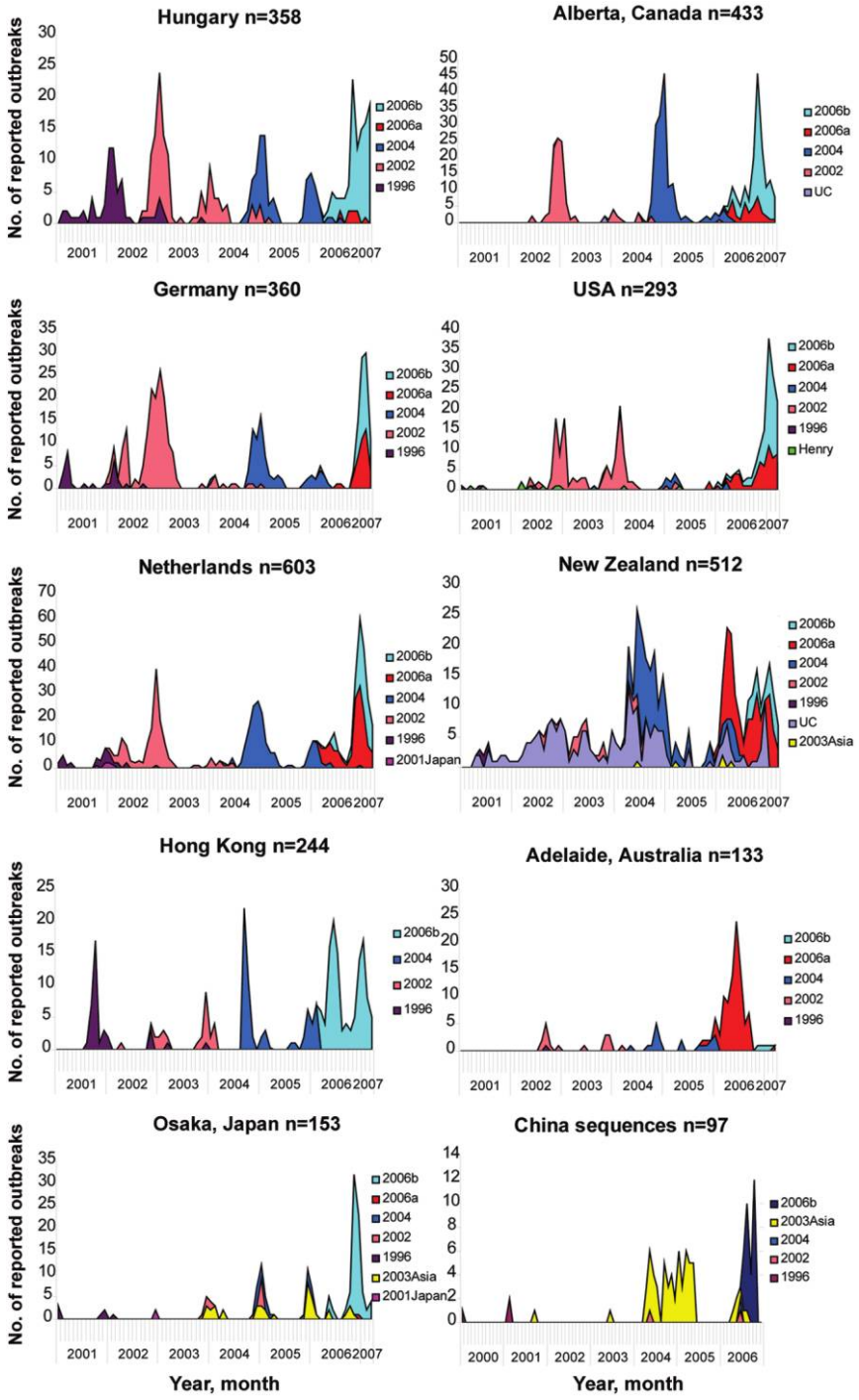

Figure 4. Prevalence of GII.4 variants, by geographic area. All data are from areas where ongoing population surveillance is being conducted; data from China are from sporadic diarrhea cases occurring in children <5 years old (derived from sequence data submitted for this study). The number of reported outbreaks is indicated for each area. Note the different scales for the *Y*-axes. The large number of unclassified (UC) strains reported by New Zealand were sequenced in region B only. In Hong Kong, 2003Asia strains were detected only in sporadic cases (especially between January 2004 and January 2005) but not in outbreaks.

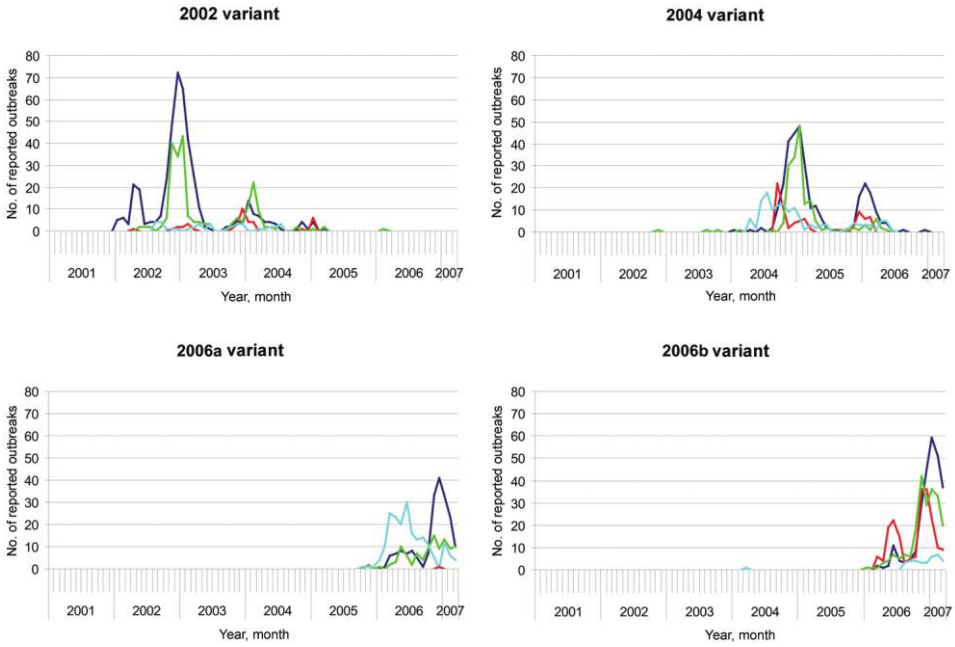

Figure 5. Timeline of reported outbreaks accumulated on each continent, by GII.4 variant. Dark blue indicates Europe (the Netherlands, Hungary, and Germany); red, Asia (Hong Kong and Japan); green, North America (United States and Canada); and light blue, Oceania (New Zealand and Australia [the Adelaide area]). Note that absolute numbers are depicted and that they vary by continent because of different sizes of reporting areas and populations.

fluenza A virus, with a rapid global spread of emerging variants. The limited information available from Africa and South America fits this general pattern, although the available data are very patchy. Clearly, more (and more-thorough) surveillance in these areas would improve our understanding of NoV epidemiology. Unlike what could be concluded from regional surveillance data alone, the present study of epidemiologic and virologic data from NoV outbreaks on 5 continents has shown that the dominant NoV genotype, GII.4, has a global impact. Additional studies, particularly in developing countries, are needed to further elucidate the burden of disease due to NoV in various communities. Although it remains possible that the increased number of reported NoV outbreaks resulted from improved diagnostics and reporting, this does not diminish our finding that the spread of the GII.4 variants was rapid and global and was unseen before the 1996 variant. Given that the incidence of NoV infection is high among young children and that diarrhea is one of the leading causes of death in children <5 years old, it is critical that studies be conducted to assess the role played by NoVs in childhood diarrhea [8, 9].

That some variants became epidemic only within a limited geographic region leads to interesting speculation as to what determines the success of different lineages (or the apparent lack thereof). The 2003Asia variant was observed in Asia but was rarely seen on other continents; conversely, the 2006a variant was rare in Asia. This was seen in our aggregated data and is substantiated in the literature, including in a recent article in which the Norovirus Surveillance Group of Japan reported a large cluster of 2003Asia and very low numbers of 2006a strains [43]. One possible explanation is that the 2006a variant shares neutralizing epitopes with the 2003Asia variant and, therefore, encountered an immune population in Asia but not in other regions. However, although the capsid of the 2003Asia variant does share 12 informative sites with that of the 2006a variant (data not shown), the same amino acid motifs are also present in the 2004 and 2006b variants, which were geographically unrestricted. Thus, prior population immunity appears to be an unlikely explanation.

A second possibility is that variants may differ in their affinity for host ligands involved in virus attachment before entry into host cells. NoVs are known to exhibit different strain-specific host-ligand-binding patterns [46]. Recently, shifts in the binding patterns of ensuing GII.4 variants were reported [16]. The lack of a receptor for the 2006a variant in Asian populations might explain why epidemics of this variant were not noted in Asia. In a similar manner, the 2003Asia variant was rarely detected outside Asia. Therefore, we compared the predicted receptor-binding interface published by Cao et al [47] in sequences of different variants, especially the 2003Asia and 2006a variants. No variation between the 2003Asia and 2006a variants could be identified at the loci involved in ligand binding other than the previously reported aa 393, which is highly polymorphic among NoV strains ([47] and X. Jiang, personal communication). Similar results indicating that more-detailed mo-

lecular characterization is needed to verify the hypotheses on binding have been reported by Lindesmith et al [16].

A third possible explanation for the differing success of GII.4 variants is the occurrence of seeding events (ie, efficient initial large-scale introduction and subsequent transmission of viruses). Diffuse, large-scale, food-related outbreaks have been described, such as the seeding of a specific strain across Europe during the winter of 2000–2001 through contaminated shellfish [48]. Such outbreaks are extremely difficult to detect given the current state of surveillance for NoVs but are estimated to be common [49].

We looked for a pattern of emergence that could be used in a global early-detection system for new epidemic strains. The date of the first reported detection is too strongly dependent on the intensity of surveillance and chance to be a reliable indicator. However, when the first epidemic peak observed in each participant's region was considered, 3 of 4 high-impact variants were first reported in Asia and Oceania. This may indicate that these antigenic variants arose in Asia and then spread throughout the world, as has been suggested for influenza A virus [50]. Additionally, the data imply that international communication of surveillance results could help prepare health care settings for "hot seasons." Oceania appeared to be ahead of countries in the Northern Hemisphere for the epidemic peaks of at least 2 variants, possibly resulting from its inverted seasons compared with the Northern Hemisphere. Epidemic peaks caused by emerging variants were preceded by off-seasonal peaks during spring and summer in the Northern Hemisphere [23, 32, 45]. As the number of outbreaks in Northern Hemisphere countries decreased during summer before becoming fully epidemic the following winter, winter started in Oceania, creating opportunities for the emerging variant there.

Our study demonstrates that NoV epidemics result from the rapid and global spread of successful GII.4 strains. These strains have been shown to evolve under the pressure of population immunity. Notable differences in the prevalence of certain lineages have been observed that cannot be explained by our current knowledge of NoVs. This highlights that joint surveillance will increase our understanding of NoV infection and epidemiology. The participants of this retrospective study will continue to collaborate and have initiated a global NoV surveillance network, NoroNet, that has already expanded to include more institutions. Through this global network, more-timely identification of unusual activity and new variants will improve efforts to limit the scale of epidemics.

Acknowledgments

We thank Dr Jutta Preiksaitits (Provincial Public Health Laboratory, Canada) for her support and Sallene Wong (Provincial Public Health Laboratory, Canada), J. K. R. Thomas, M. A. C. Hamat, H. Kamyon, and M. R. Idid (University of Malaya, Malaysia) for practical assistance. J.J.S. kindly thanks all coauthors for generously sharing their data and answering queries.

References

1. Okada M, Tanaka T, Oseto M, Takeda N, Shinozaki K. Genetic analysis of noroviruses associated with fatalities in healthcare facilities. Arch Virol **2006**; 151:1635–41.
2. Siebenga JJ, Beersma MF, Vennema H, van Biezen P, Hartwig NJ, Koopmans M. High prevalence of prolonged norovirus shedding and illness among hospitalized patients: a model for in vivo molecular evolution. J Infect Dis **2008**; 198:994–1001.
3. Hansen S, Stamm-Balderjahn S, Zuschneid I, et al. Closure of medical departments during nosocomial outbreaks: data from a systematic analysis of the literature. J Hosp Infect **2007**; 65:348–53.
4. Johnston CP, Qiu H, Ticehurst JR, et al. Outbreak management and implications of a nosocomial norovirus outbreak. Clin Infect Dis **2007**; 45:534–40.
5. De Wit MA, Koopmans MP, Kortbeek LM, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. Am J Epidemiol **2001**; 154:666–74.
6. Wheeler JG, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. BMJ **1999**; 318:1046–50.
7. Mead PS, Slutsker L, Dietz V, et al. Food-related illness and death in the United States. Emerg Infect Dis **1999**; 5:607–25.
8. Kosek M, Bern C, Guerrant RL. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull World Health Organ **2003**; 81:197–204.
9. Patel MM, Widdowson MA, Glass RI, Akazawa K, Vinje J, Parashar UD. Systematic literature review of role of noroviruses in sporadic gastroenteritis. Emerg Infect Dis **2008**; 14:1224–31.
10. Teunis PF, Moe CL, Liu P, et al. Norwalk virus: how infectious is it? J Med Virol **2008**; 80:1468–76.
11. Tu ET, Bull RA, Kim MJ, et al. Norovirus excretion in an aged-care setting. J Clin Microbiol **2008**; 46:2119–21.
12. Chan MC, Sung JJ, Lam RK, et al. Fecal viral load and norovirus-associated gastroenteritis. Emerg Infect Dis **2006**; 12:1278–80.
13. Johnson PC, Mathewson JJ, DuPont HL, Greenberg HB. Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults. J Infect Dis **1990**; 161:18–21.
14. Parrino TA, Schreiber DS, Trier JS, Kapikian AZ, Blacklow NR. Clinical immunity in acute gastroenteritis caused by Norwalk agent. N Engl J Med **1977**; 297:86–9.
15. Siebenga JJ, Vennema H, Renckens B, et al. Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J Virol **2007**; 81:9932–41.
16. Lindesmith LC, Donaldson EF, Lobue AD, et al. Mechanisms of GII.4 norovirus persistence in human populations. PLoS Med **2008**; 5:e31.
17. Green KY. Caliciviridae: the noroviruses. In: Knipe DM, Howley PM, eds. Fields virology. 5th ed. Vol 1. Philadelphia: Lippincott Williams & Wilkins, **2007**:949–80.
18. Kroneman A, Verhoef L, Harris J, et al. Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1-7-2001 to 30-6-2006. J Clin Microbiol **2008**; 46:2959–65.
19. 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**; 44:327–33.
20. Noel JS, Fankhauser RL, Ando T, Monroe SS, Glass RI. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J Infect Dis **1999**; 179:1334–44.
21. Okada M, Ogawa T, Yoshizumi H, Kubonoya H, Shinozaki K. Genetic variation of the norovirus GII-4 genotype associated with a large

number of outbreaks in Chiba prefecture, Japan. Arch Virol **2007**; 152:2249–52.
22. 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**; 76:598–607.
23. Reuter G, Pankovics P, Szucs G. Genetic drift of norovirus genotype GII-4 in seven consecutive epidemic seasons in Hungary. J Clin Virol **2008**; 42:135–40.
24. Tu ET, Bull RA, Greening GE, et al. Epidemics of gastroenteritis during 2006 were associated with the spread of norovirus GII.4 variants 2006a and 2006b. Clin Infect Dis **2008**; 46:413–20.
25. Tu ET, Nguyen T, Lee P, et al. Norovirus GII.4 strains and outbreaks, Australia. Emerg Infect Dis **2007**; 13:1128–30.
26. Siebenga JJ, Vennema H, Duizer E, Koopmans MP. Gastroenteritis caused by norovirus GGII.4, the Netherlands, 1994–2005. Emerg Infect Dis **2007**; 13:144–6.
27. Vinjé J, Hamidjaja RA, Sobsey MD. Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses. J Virol Methods **2004**; 116: 109–17.
28. Phan TG, Kaneshi K, Ueda Y, et al. Genetic heterogeneity, evolution, and recombination in noroviruses. J Med Virol **2007**; 79:1388–400.
29. Gallimore CI, Iturriza-Gomara M, Xerry J, Adigwe J, Gray JJ. Interseasonal diversity of norovirus genotypes: emergence and selection of virus variants. Arch Virol **2007**; 152:1295–303.
30. Maguire AJ, Green J, Brown DW, Desselberger U, Gray JJ. Molecular epidemiology of outbreaks of gastroenteritis associated with small round-structured viruses in East Anglia, United Kingdom, during the 1996–1997 season. J Clin Microbiol **1999**; 37:81–9.
31. Parker TD, Kitamoto N, Tanaka T, Hutson AM, Estes MK. Identification of genogroup I and genogroup II broadly reactive epitopes on the norovirus capsid. J Virol **2005**; 79:7402–9.
32. Lopman B, Vennema H, Kohli E, et al. Increase in viral gastroenteritis outbreaks in Europe and epidemic spread of new norovirus variant. Lancet **2004**; 363:682–8.
33. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **2003**; 52:696–704.
34. Fankhauser RL, Monroe SS, Noel JS, et al. Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States. J Infect Dis **2002**; 186:1–7.
35. Blanton LH, Adams SM, Beard RS, et al. Molecular and epidemiologic trends of caliciviruses associated with outbreaks of acute gastroenteritis in the United States, 2000–2004. J Infect Dis **2006**; 193:413–21.
36. Seto Y, Iritani N, Kubo H, et al. Genotyping of norovirus strains detected in outbreaks between April 2002 and March 2003 in Osaka City, Japan. Microbiol Immunol **2005**; 49:275–83.
37. Ho EC, Cheng PK, Wong DA, Lau AW, Lim WW. Correlation of norovirus variants with epidemics of acute viral gastroenteritis in Hong Kong. J Med Virol **2006**; 78:1473–9.
38. Vidal R, Solari V, Mamani N, et al. Caliciviruses and foodborne gastroenteritis, Chile. Emerg Infect Dis **2005**; 11:1134–7.
39. Hohne M, Schreier E. Detection and characterization of norovirus outbreaks in Germany: application of a one-tube RT-PCR using a fluorogenic real-time detection system. J Med Virol **2004**; 72:312–9.
40. Schreier E, Doring F, Kunkel U. Molecular epidemiology of outbreaks of gastroenteritis associated with small round structured viruses in Germany in 1997/98. Arch Virol **2000**; 145:443–53.
41. Rachakonda G, Choudekar A, Parveen S, Bhatnagar S, Patwari A, Broor S. Genetic diversity of noroviruses and sapoviruses in children with acute sporadic gastroenteritis in New Delhi, India. J Clin Virol **2008**; 43:42–8.
42. Halperin T, Yavzori M, Amitai A, et al. Molecular analysis of noroviruses involved in acute gastroenteritis outbreaks in military units in Israel, 1999–2004. Eur J Clin Microbiol Infect Dis **2005**; 24:697–700.
43. Motomura K, Oka T, Yokoyama M, et al. Identification of monomorphic and divergent haplotypes in 2006/2007 norovirus GII/4 epidemic population by genome-wide tracing of evolutionary history. J Virol **2008**; 82:11247–62.
44. Yee EL, Palacio H, Atmar RL, et al. Widespread outbreak of norovirus gastroenteritis among evacuees of Hurricane Katrina residing in a large "megashelter" in Houston, Texas: lessons learned for prevention. Clin Infect Dis **2007**; 44:1032–9.
45. Verhoef L, Depoortere E, Boxman I, et al. Emergence of new norovirus variants on spring cruise ships and prediction of winter epidemics. Emerg Infect Dis **2008**; 14:238–43.
46. Tan M, Jin M, Xie H, Duan Z, Jiang X, Fang Z. Outbreak studies of a GII-3 and a GII-4 norovirus revealed an association between HBGA phenotypes and viral infection. J Med Virol **2008**; 80:1296–301.
47. Cao S, Lou Z, Tan M, Chen, et al. Structural basis for the recognition of blood group trisaccharides by norovirus. J Virol **2007**; 81:5949–57.
48. Koopmans M, Vennema H, Heersma H, et al. Early identification of common-source foodborne virus outbreaks in Europe. Emerg Infect Dis **2003**; 9:1136–42.
49. Verhoef LP, Kroneman A, van Duynhoven Y, Boshuizen H, van Pelt W, Koopmans M. Selection tool for foodborne norovirus outbreaks. Emerg Infect Dis **2009**; 15:31–8.
50. Russell CA, Jones TC, Barr IG, et al. The global circulation of seasonal influenza A (H3N2) viruses. Science **2008**; 320:340–6.