Temporal dynamics of norovirus determined through monitoring of municipal
wastewater by pyrosequencing and virological surveillance of gastroenteritis cases

Shinobu Kazama,^a Yoshifumi Masago,^{a,b,#} Kentaro Tohma,^c Nao Souma,^c Toshifumi
Imagawa,^c Akira Suzuki,^d Xiaofang Liu,^c Mayuko Saito,^c Hitoshi Oshitani,^c Tatsuo
Omura^a

^aNew Industry Creation Hatchery Center, Tohoku University, Sendai, Miyagi, 980-8479,
Japan

^bInstitute for the Advanced Study of Sustainability, United Nations University,
Shibuya-ku, Tokyo 150-8925, Japan

^cDepartment of Virology, Tohoku Graduate School of Medicine, Tohoku University,
Sendai, Miyagi, 980-8575, Japan

^dVirus Research Center, Clinical Research Division, Sendai Medical Center, Sendai,
Miyagi, 983-8520, Japan

#Address correspondence to Yoshifumi Masago, masago@[unu.edu](mailto:masago@unu.edu)

Tel.: +81 3 5467 1212

ABSTRACT

Norovirus is a leading etiological agent of viral gastroenteritis. Because of relatively mild disease symptoms and frequent asymptomatic infections, information on the ecology of this virus is limited. Our objective was to examine the genetic diversity of norovirus circulating in the human population by means of genotyping the virus in municipal wastewater. We investigated norovirus genogroups I and II (GI and GII) in municipal wastewater in Japan by pyrosequencing and quantitative PCR (qPCR) from November 2012 to March 2013. Virological surveillance for gastroenteritis cases was concurrently conducted in the same area. A total of fourteen distinct genotypes in total (GI.1, 3, 4, 6, 7, GII.2, 4, 5, 6, 7, 12, 13, 14, and 17), with up to eight genotypes detected per sample, were observed in wastewater using pyrosequencing; only four genotypes (GI.6, GII.4, 5, and 14) were obtained from clinical samples. Seventy-eight percent of norovirus-positive stool samples contained GII.4, but this genotype was not dominant in wastewater. The norovirus GII.4 Sydney 2012 variant, which appeared and spread during our study period, was detected in both the wastewater and clinical samples. These results suggest that an environmental approach using pyrosequencing yields a more detailed distribution of norovirus genotypes/variants. Thus, wastewater monitoring by pyrosequencing is expected to provide an effective analysis of the distribution of norovirus genotypes causing symptomatic and asymptomatic infections in human populations.

Keywords

Norovirus; Wastewater; Pyrosequencing; Virological surveillance

1. INTRODUCTION

Noroviruses (family *Caliciviridae*, genus *Norovirus*) are one of the leading causative agents for acute gastroenteritis around the world. Noroviruses, which cause diarrhea and vomiting, affect individuals of all ages and are transmitted primarily through fecal-oral, aerosol-vomit, or direct-contact routes (Glass et al., 2009). Although the infection can be fatal in vulnerable populations such as infants and the elderly, asymptomatic infections are also common (Glass et al., 2009). Up to 200,000 deaths occur annually among children under the age of five in developing countries (Patel et al., 2008). Outbreaks most frequently occur during cold weather (Ahmed et al., 2013; Lopman et al., 2009; Mounts et al., 2000), and this seasonality is also observed in Japan (Mounts et al., 2000; Siebenga et al., 2009; Yoneda et al., 2014). Noroviruses are classified into five genogroups (GI-GV), which are further subdivided into at least 35 genotypes (Centers for Disease Control and Prevention, 2011; Zheng et al., 2006). Among these, human disease is caused by GI, GII, and GIV (Patel et al., 2009; Zheng et al., 2006). In recent years, most outbreaks have been caused by GII genotype 4 (GII.4) (Glass et al., 2009; Patel et al., 2009; Siebenga et al., 2009). In 2012, a new GII.4 variant called Sydney 2012 was detected in Australia and rapidly spread to many countries (Chan et al., 2014; Eden et al., 2013; Fonager et al., 2013; Giammanco et al., 2013; Leshem et al., 2013b; Mai et al., 2013; Rahman et al., 2013; Vega et al., 2014). Virological surveillance is commonly conducted for gastroenteritis cases to detect the occurrence of norovirus. However, in this approach, noroviruses can only be identified in patients who seek medical care; asymptomatic infections are not detected. Moreover, enteric viral identification at medical facilities is not mandatory, even in developed countries. Thus, the number of norovirus infections and the genotypes circulating in the

human populations and water environment are not well understood.

Because human noroviruses (GI, GII and GIV) are host-specific and cannot replicate in other organisms, the presence of the virus in municipal wastewater indicates the presence of infected individuals in the area. Based on this supposition, many studies have been conducted on norovirus occurrence and genotypes in wastewater and other environmental water (Aw and Gin, 2010; Blanco Fernandez et al., 2011; Hernandez-Morga et al., 2009; Katayama et al., 2008; Kitajima et al., 2012; 2010; La Rosa et al., 2010; Lee et al., 2012; Lee et al., 2011; Mans et al., 2013; Pérez-Sautu et al., 2012). Only a few studies have investigated norovirus occurrence and genotypes in both human populations and the water environment at the same time (Kremer et al., 2011; Rajko-Nenow et al., 2013; Ueki et al., 2005). However, because these studies only sequenced a few clones from each sample by direct-sequencing or cloning-sequencing, researchers may have just detected the predominant genotypes. Considering the wide molecular diversity and rapid evolution of the norovirus genome (Bok et al., 2009; Bull et al., 2007; Eden et al., 2014; Zheng et al., 2006), it would be highly advantageous to develop a novel approach capable of detecting multiple genotypes is desired.

Next-generation sequencing (NGS), a high-throughput sequencing technique, has recently been applied to study viral diversity, discover novel viruses, and monitor viral evolution in humans and the environment (Barzon et al., 2011; Cantalupo et al., 2011). This technique allows a large number of nucleotide sequences to be obtained from a single sample. Therefore, we expected that NGS could be used to determine the detailed distribution of norovirus genotypes in wastewater.

The present study aimed to investigate the genomic dynamics of norovirus in an urbanized part of Japan. Specifically, we performed virological monitoring of municipal

wastewater using pyrosequencing, an NGS technology. Noroviruses GI and GII in wastewater samples were genotyped using pyrosequencing and also quantified using quantitative PCR (qPCR). In parallel, we conducted virological surveillance for gastroenteritis cases in the catchment area to investigate the number of patients admitted to the hospital with norovirus infection and to determine norovirus genotypes in patient stools. This study was conducted from 2012 to 2013, during an interval when the norovirus GII.4 Sydney 2012 variant was spreading in the study area, permitting us to simultaneously detect the shift in predominant variants in both wastewater and patients. To the best of our knowledge, this is the first report concerning molecular genotyping of noroviruses in wastewater using pyrosequencing.

2. MATERIALS AND METHODS

2.1. Study design

This study was conducted from November 2012 to March 2013 in the town of Matsushima in northeastern Japan's Miyagi Prefecture. As of March 2013, the town had 15,141 inhabitants and one municipal wastewater treatment plant, which received wastewater from 67.1% of the population. There are four internal medicine clinics in the town, one of which participated in our virological surveillance for gastroenteritis cases.

2.2. Analysis of noroviruses in wastewater samples

2.2.1. Sample collection

Primary effluents were collected at the wastewater treatment plant every week from November 14, 2012 (the 46th week of 2012), to March 13, 2013 (the 11th week of 2013), except for the 1st week of 2013; thus a total of 17 weekly samples were collected and

analyzed. The grab samples (250 mL each) were collected around 10 a.m., transported to the laboratory on ice, and stored in a deep freezer (−80 °C) until analyzed.

2.2.2. Virus concentration and nucleic acid extraction

To process each weekly specimen, noroviruses were recovered from 40 mL of the primary effluent sample and concentrated using the polyethylene glycol (PEG) precipitation method. Prior to the recovery and concentration process, a known amount (approximately 10^8 genome copies) of the murine norovirus strain S7-PP3 (MNV), provided by Prof. Yukinobu Tohya (Nihon University, Japan), was spiked into each sample as a processing control. Each sample was gently mixed with 3.2 g of PEG 6000 (Wako Pure Chemical Industries, Osaka, Japan) and 0.92 g of NaCl (Kanto Kagaku, Tokyo, Japan) using a magnetic stirrer for 12 h at 4 °C, then centrifuged at $9,000\times g$ for 30 min at 4 °C. Each pellet was resuspended in 1 mL of sterilized, deionized water and mixed using a vortex mixer for 1 min. The resuspended pellets were centrifuged at $10,000\times g$ for 10 min at 4 °C and the supernatants were collected as virus concentrates.

Ribonucleic acid was extracted using QIAamp Viral RNA mini Kit (Qiagen, Hilden, Germany) with QIAcube (Qiagen), and cDNA was synthesized using iScript Advanced cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) with a Veriti 96-well Thermal Cycler (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instructions.

2.2.3. Norovirus quantification using qPCR assay

Norovirus GI and GII were detected by qPCR using the CFX96 Real-Time PCR Detection System (Bio-Rad). The qPCR was performed using SsoFast Probes Supermix

(Bio-Rad) and the following primers and probes: COG1F, COG1R, RING1(a)-TP, and RING1(b)-TP for GI; and COG2F, COG2R, and RING2AL-TP (5'-FAMTGG GAG GGS GAT CGC RAT CT-TAMRA-3') for GII (Aoki et al., 2010; Ministry of Health , Labour and Welfare, 2007; Kageyama et al., 2003). Each 20-μL reaction mixture contained 5 μL of cDNA, 10 μL of SsoFast Probes Supermix (Bio-Rad), and the primers and probes specified in the references. The PCR cycling conditions were 10 min at 95 °C, followed by 50 cycles of 15 s at 95 °C, 60 s at 56 °C, and 30 s at 72 °C. The number of norovirus genome copies was determined using a standard curve generated from a 10-fold serial dilution of standard DNA (10^8 – 10^1 copies/well). Synthesized oligonucleotide of the target region (Nihon Gene Research Laboratories, Sendai, Japan) was used as the standard DNA. A limit of quantification cycle value (LOQ) was set at 40 to follow the MIQE guidelines (Bustin et al., 2009). Each sample was measured in triplicate. The geometric mean copy number in the sample was converted to concentration in the wastewater sample when all replicated quantification cycle (Cq) values were less than 40. The concentrations of norovirus genome in the wastewater samples were calculated without recovery rates. When one or more replicated Cq value(s) was more than 40, the corresponding concentration was recorded as below LOQ. If all replicates were negative, the corresponding concentration was recorded as “not detected”. In order to evaluate overall recovery rates, the MNV spiked into each sample as a processing control was also quantified using qPCR (Hata et al., 2011; Kitajima et al., 2008). The overall recovery rate was calculated by dividing the copy number of the MNV detected in the virus concentrate by the copy number of the MNV spiked into the sample. The recovery rates ranged from 2% to 19% (geometric mean: 8%, n = 17).

2.2.4. Amplification for pyrosequencing

Semi-nested PCR was performed using primers as follows. The primers for the first and second PCR were COG1F/G1SKR and G1SKF/G1SKR for GI, and COG2F/G2SKR and G2SKF/G2SKR for GII (Kageyama et al., 2003; Kojima et al., 2002). For the first PCR, the 50- μ L reaction mixture contained 15 μ L of cDNA, 25 μ L of Q5 Hot Start High-Fidelity 2 $\times$ Master Mix (New England Biolabs, Ipswich, MA, USA), and 25 pmol of both forward and reverse primers. For the second PCR, the 100- μ L reaction mixture contained 2 μ L of the first PCR product, 50 μ L of Q5 Hot Start High-Fidelity 2 $\times$ Master Mix (New England Biolabs), and 50 pmol of both forward and reverse primers. Amplification by PCR for both the first and second PCR was performed using a Veriti 96-well Thermal Cycler (Thermo Fisher Scientific), with reactions initiated by incubation for 30 s at 98 °C, followed by 25 cycles of 10 s at 98 °C, 30 s at 50 °C, and 30 s at 72 °C, with a final extension for 30 s at 72 °C. The semi-nested PCR product sizes for GI and GII were approximately 330 bp and 340 bp, respectively.

The semi-nested PCR products were visualized using agarose gel electrophoresis, and samples with a band at the expected position were purified and submitted for pyrosequencing. Purification consisted of concentrating 90 μ L of the nested-PCR product to 30 μ L using the QIAquick PCR Purification Kit (Qiagen) with QIAcube (Qiagen).

2.2.5. Pyrosequencing

In order to perform pyrosequencing on the norovirus amplicons using the GS Junior system (Roche Applied Science, Penzberg, Germany), unique adaptors were ligated to

both the 5' and 3' ends of the amplicons by fusion PCR. The fusion primers consisted of FLX Titanium Primer A (25-mer sequence used for the sequencing), Multiplex Identifier (10-mer sequence for barcoding each sample), G1SKF or G2SKF primer sequences as forward primers, and FLX Titanium Primer B (25-mer sequence used for the sequencing) and G1SKR or G2SKR primer sequences as reverse primers. The 100- μ L reaction mixture contained 10 μ L of the purified nested-PCR products, 50 μ L of Q5 Hot Start High-Fidelity 2 $\times$ Master Mix (New England Biolabs), and 50 pmol of both forward and reverse primers. The PCR reactions were performed using a Veriti 96-well Thermal Cycler (Thermo Fisher Scientific), and started with 30 s at 98 °C, followed by 5 cycles of 10 s at 98 °C, 30 s at 50 °C, and 30 s at 72 °C, with a final extension for 30 s at 72 °C.

Ninety microliters of the fusion PCR products were purified and concentrated to 30 μ L using QIAquick PCR Purification Kit (Qiagen) with QIAcube (Qiagen). DNA concentrations of the nested PCR products were measured using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher Scientific) with infinite M1000 PRO (TECAN, Männedorf, Switzerland). Three to five samples with identical Multiplex Identifier sequences were mixed for the individual pyrosequencing runs using the GS Junior system (Roche Applied Science) with the Titanium emPCR Kit (Lib-L) and the GS Junior Titanium Sequencing Kit (Roche Applied Science) following the manufacturer's instructions.

2.2.6. Bioinformatic analysis

Although NGS provides large numbers of sequence reads, the sequence data contains non-target sequences such as ambiguous reads with noise and chimeric sequences.

These artifacts are produced during nested-PCR and pyrosequencing, and could result in overestimation of genetic diversity (Quince et al., 2011). Therefore, the bioinformatic analysis was performed as follows using QIIME 1.8.0 software (Caporaso et al., 2010). Quality filtering and primer sequence removal were performed using the software package `split_library.py` with a minimum quality score parameter of 25. Sequences with incorrect nucleotides produced in the nested-PCR and the pyrosequencing steps were corrected using the `denoiser.py` package. Chimeric sequences (sequence formed by two or more sequences) were removed using the Perseus software (Quince et al., 2011) after removing reverse primers using the `split_library.py` package. Sequences were then clustered into operational taxonomic units (OTUs) based on a minimum 97% similarity in nucleotide sequence using the `pick_otus.py` package, and a representative sequence of each OTU was selected using the `pick_rep_set.py` package.

Genotypes and variants of the representative sequences were identified using the Norovirus Genotyping Tool Version 1.0 (Kroneman et al., 2011). When not assigned to any genotypes by the tool, sequences were subjected to homology search using BLASTn, and the genotype or variant of top-hit sequences with the highest similarity (exceeding a minimum of 97%) was assigned. If a given sequence could not be assigned using either the Norovirus Genotyping Tool or BLASTN, that sequence was categorized as “not assigned”.

The diversity of the norovirus strains (OTUs) in each wastewater sample was evaluated using a rarefaction curve generated by the Analytic Rarefaction 2.0 software (<http://strata.uga.edu/software/>).

2.3. Analysis of noroviruses in clinical samples

2.3.1. Sample collection and norovirus screening

Virological surveillance for gastroenteritis cases was conducted at an outpatient Internal Medicine and Pediatrics clinic in Matsushima. Patients with diarrhea who visited the clinic from the 43rd week in 2012 to the 13th week in 2013 (96 patients) were included in this study. Rectal swabs were collected from the patients after obtaining informed consent. The samples were stored at 2-8 °C and transferred to the authors' laboratory weekly.

The samples were tested for norovirus using a real-time PCR assay. Each rectal swab was moistened in 1 mL of phosphate-buffered saline, and ribonucleic acid extraction was performed using the QIAamp Viral RNA Mini Kit (Qiagen) with QIAcube (Qiagen). cDNA was synthesized using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) following the manufacturer's instructions. Noroviruses were detected by real-time PCR (Kageyama et al., 2003) using the TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific) and the Applied Biosystems 7500 Real-Time PCR System (Thermo Fisher Scientific) following the manufacturer's instructions. The PCR cycling conditions were 2 min at 50 °C and 20 s at 95 °C, followed by 45 cycles of 3 s at 95 °C and 35 s at 60 °C. In addition to norovirus, the levels of sapovirus, astrovirus and rotavirus were also tested (Heim et al., 2003; Oka et al., 2006; Pang et al., 2004) for reference.

2.3.2. Analysis of genotypes and variants

The norovirus-positive stools were assessed using Sanger sequencing. The capsid N/S-encoding domain of the norovirus genome in the norovirus-positive stool samples was amplified by single-round PCR or nested PCR using *Ex Taq* (Takara Bio, Otsu,

Japan) and the primers p290, COG1F, COG2F, G1SKR, and G2SKR, to determine the GI and GII strain genotypes (Jiang et al., 1999; Kageyama et al., 2003; Kojima et al., 2002). The PCR products were purified using QIAquick PCR Purification Kit (Qiagen) with QIAcube (Qiagen), and then subjected to sequencing reactions using the BigDye Terminator v1.1 Sequencing Kit (Thermo Fisher Scientific) following the manufacturer's instructions. The reaction products were purified using BigDye TXTerminator Purification Kit (Thermo Fisher Scientific) followed by sequencing using the Applied Biosystems 3130 Genetic Analyzer or 3730xl DNA Analyzer (Thermo Fisher Scientific) according to the manufacturer's instructions. Multiple alignment and data cleaning were performed using MEGA5 software (Tamura et al., 2011). The genotyping was performed using the Norovirus Typing Tool Version 1.0 (Kroneman et al., 2011).

2.4. Phylogenetic analysis

In order to compare GII.4 sequences obtained from wastewater samples to those obtained from clinical samples, phylogenetic analysis was carried out. The sequences were analyzed using ClustalW; bootstrapped phylogenetic trees were then constructed by the maximum likelihood method with 1,000 bootstrap replications using MEGA5 software (Tamura et al., 2011). The genetic distances were calculated using the Kimura 2-parameter method.

2.5. Nucleotide sequence accession numbers

Nucleotide sequence data from wastewater samples and clinical samples has been deposited in the DDBJ/EMBL/GenBank databases under the accession numbers

DRA002609 and LC060872-LC060903, respectively.

3. RESULTS

3.1. Noroviruses in wastewater samples by qPCR

The concentrations of noroviruses GI and GII in wastewater samples, as determined by qPCR, are shown in Fig. 1. Norovirus GI was detected in all samples throughout the study period, with the concentrations ranging up to 8.7×10^4 copies/mL. The concentrations were relatively high between December 26, 2012 and January 16, 2013 and between February 27, 2013 and March 13, 2013. GII was first detected on December 12, and the concentrations were relatively high (up to 1.3×10^3 copies/mL) from January 9, 2013, to February 20, 2013, except on January 23, 2013.

3.2. Noroviruses in wastewater samples by pyrosequencing

In our study, 17 samples were analyzed by pyrosequencing for GI and GII, and a total of 999,097 reads (with means of 17,183 reads for GI and 41,587 reads for GII per sample) were obtained. Pyrosequencing produces a high numbers of reads, but these reads typically contain artifacts such as sequences with incorrect nucleotides and chimeric sequences generated during nested-PCR and pyrosequencing. These artifacts must be addressed through correction (denoising) or removal (chimera removal). On average, chimera removal eliminated 0.7% of GI reads and 3.7% of GII reads. Following denoising and chimera removal, the mean numbers of OTUs per sample was 5 for GI and 32 for GII. On the basis of the number of OTUs, GII exhibited a higher diversity than GI in wastewater samples. All rarefaction curves except one reached or

almost reached a plateau. The rarefaction curve generated from GI strains in the sample collected on November 14, 2012 (the lowest concentration of GI in wastewater samples) did not reach a plateau. This result indicates that the sequencing was nearly deep enough to understand the diversity of norovirus.

The representative sequences of the OTUs were genotyped using the Norovirus Typing Tool and BLASTn search. Figure 2 shows ratios of the number of reads assigned to each genotype/variant divided by the number of reads identified as norovirus in each sample. Because sequences were obtained through nested-PCR and pyrosequencing, genotype distributions (number of reads for each genotype) obtained by pyrosequencing were not expected to be identical with the actual values in the wastewater samples. Therefore, ratios exceeding 10% were denoted as high ratio. Fourteen genotypes in total and up to eight genotypes per wastewater sample (collected on January 9, 2013) were detected. No norovirus sequence was obtained on several dates (November 21, 28, and December 12, 2012, for GI; November 21 and 28, 2012, for GII), although noroviruses were detected by qPCR and nested-PCR products were observed at target length in samples from these dates. The numbers of sequences obtained from the samples was very low (less than 100 reads), and these sequences were removed during the quality filtering and chimera removal steps following pyrosequencing. Thus the nested-PCR products were inferred as non-specific products.

Norovirus GI was identified in 14 samples by pyrosequencing, including five genotypes (GI.1, 3, 4, 6, and 7). GI.4 was the most frequently detected genotype in 53% (9/17) of the samples and occurred at high ratios (>10%) in eight samples. GI.7 was detected in 47% (8/17) of the samples but occurred at high ratios only on December 28, 2012, and March 13, 2013. GI.3 was intermittently detected in 29% (5/17) of the samples in 2013

at high ratios, whereas GI.1 was detected in 12% (2/17) of the samples but only observed in 2012. GI.6 was detected in 12% (2/17) of the samples on December 28, 2012 and January 9, 2013.

Nine GII genotypes (GII.2, 4, 5, 6, 7, 12, 13, 14, and 17) were detected in 14 wastewater samples. Single genotypes (GII.2 or GII.14) per sample were detected at the beginning of the monitoring period (November 14, 2012 and December 5, 2012, respectively); multiple genotypes per sample were detected starting from December 12, 2012. GII.2 and GII.4 were the most frequently detected genotypes (11/17 samples, 65% for both genotypes), followed by GII.14 (10/17 samples, 59%). All of these genotypes appeared at high ratios, except for GII.4 on January 9, 2013. Other genotypes were less frequently detected, namely GII.5 (4/17, 24%), GII.7 (4/17, 24%), GII.17 (4/17, 24%), GII.6 (2/17, 12%), GII.12 (1/17, 6%), and GII.13 (1/17, 6%); these genotypes appeared only in 2013.

Three GII.4 variants were detected, namely Den Haag 2006b (3/17, 18%), New Orleans 2009 (1/17, 5.9%), and Sydney 2012 (10/17, 59%). All three variants were detected until the third week of 2013, but only the Sydney 2012 variant was detected thereafter, implying that Sydney 2012 became the predominant variant from the fourth week of 2013.

3.3. Norovirus detection and clinical sample characteristics

Figure 3 shows the number of outpatients identified with gastroenteritis and norovirus infection. In total, 96 stool samples were collected from patients with gastroenteritis from the 43rd week in 2012 to the 13th week in 2013. Among these, 32 samples (33%) contained norovirus, including 4 cases of GI (13%) and 28 cases of GII (87%). Other

viruses were detected in 21 of 63 norovirus-negative samples, including sapovirus (12 samples), astrovirus (7 samples), and rotavirus (2 samples). Although the number of samples was low to show an epidemic curve, the number of norovirus-positive stool samples peaked between the 48th and 50th week of 2012, with the highest number of cases detected during the 50th week (5 cases).

Norovirus GI.6 was the only GI genotype detected from patients (n = 4), while three GII genotypes [GII.4 (n = 25), GII.5 (n = 1), and GII.14 (n = 2)] were detected. Norovirus GII.4, the predominant genotype (25 samples, 78%), included three variants: Sydney 2012 (18 samples, 72%), Den Haag 2006b (6 samples, 24%), and New Orleans 2009 (1 sample, 4%) (Fig. 4). At the beginning of the study period, the Sydney 2012 and Den Haag 2006b variants appeared to be co-circulating; only a single case of New Orleans 2009 was detected in this interval. After the 52nd week of 2012, only the Sydney 2012 variant was detected (Fig. 3). Although the number of clinical samples was low, our results suggest a shift occurred towards predominance by the GII.4 variant.

3.4. Comparison between wastewater samples and clinical samples

Fourteen genotypes were detected in wastewater samples, whereas only four genotypes were detected in stool samples (Figs. 2-5). All genotypes detected from stool samples (GI.6, GII.4, GII.5, and GII.14) also were detected in wastewater. Noroviruses GII.4 and GII.14 were detected in 65% and 59% of the wastewater samples, respectively. Genotype GII.5, which was detected in one stool sample collected during the 8th week of 2013, also was detected in wastewater during an overlapping interval (the 7th, 9th, and 10th weeks of 2013). In contrast, GII.2 was detected at high ratios throughout the study period in 65% of wastewater samples but never in stool samples. Furthermore, GI.6 was

exclusively detected in four GI-positive stool samples, whereas GI.6 was detected in only two wastewater samples (12%). The wastewater contained various genotypes that were not observed in the virological surveillance of gastroenteritis cases.

The prevalence of GII.4 variants showed similar temporal trends between wastewater (Fig. 2) and clinical samples (Fig. 3) in our study period. Three variants (Den Haag 2006b, New Orleans 2009, and Sydney 2012) were detected; the Sydney 2012 variant was the most frequently detected in both types of samples. All three variants were detected at the beginning of the study period, both in wastewater and stool samples, but only Sydney 2012 was detected later in the study period. However, there were some notable differences in the timing of this shift. The Sydney 2012 variant was exclusively detected in stool samples during and after the 52nd week of 2012, whereas other variants were detected from wastewater until the 3rd week of 2013.

Fig. 5 shows a phylogenetic tree derived using GII.4 sequences obtained from the wastewater and stool samples. Given that the wastewater samples yielded large numbers of OTUs, representative sequences of OTUs containing more than 1000 reads were used for the phylogenetic analysis. All the sequences identified as the same variant clustered on distinct branches. These results indicated that genotypes/variants detected in wastewater using pyrosequencing indeed reflected gastroenteritis cases in human populations.

4. DISCUSSION

In this study, we used pyrosequencing and qPCR to assess the genomic diversity of noroviruses using virological monitoring of both wastewater and clinical samples, including analysis of genotypes and concentration in wastewater. Wastewater

monitoring allowed us to identify in detail the diversity of norovirus genotypes. This included a total of 14 different genotypes, as well as on samples, which by itself contained eight genotypes (January 9, 2013). Some previous attempts have been made to show the relationship between noroviruses in feces and their presence in the water environment (Kremer et al., 2011; Rajko-Nenow et al., 2013). However, because those previous studies used direct-sequencing or cloning-sequencing methods to genotype noroviruses from environmental samples, the respective laboratories detected only one or two genotypes in a given sample. Thus, detection may have been limited to predominant genotypes only. Although it is possible to detect multiple genotypes using cloning-sequencing, this requires large amounts of cloning and considerable cost and effort.

We additionally note that the use of NGS, as employed in the present work, could also be used to evaluate wastewater treatment processing and water quality. A recent study reported that distinct norovirus genotypes may exhibit different tolerances in wastewater treatment processing and different accumulation rates in oysters (Rajko-Nenow et al., 2013). Detection of multiple genotypes using NGS would enhance investigations such as genotype-dependent tolerance in wastewater treatment processes and oyster environments.

Norovirus GII.4 was detected most frequently in stool samples (78%), but this variant did not dominate in wastewater samples. One possible explanation for this seemingly contradictory finding is that GII.4 is more likely to cause symptomatic infection as compared with other genotypes (Barreira et al., 2010; Bucardo et al., 2010; Okabayashi et al., 2008). Because the clinical samples were obtained from symptomatic patients who sought medical care, GII.4 is expected to be detected more frequently. In contrast,

all genotypes causing both symptomatic and asymptomatic infection are deposited in wastewater. Similarly, previous studies reported that the GI concentration in wastewater was comparable to that of GII in wastewater (Katayama et al., 2008; Kitajima et al., 2012; Nordgren et al., 2009), but GII genotypes were most frequently detected in stool samples (Chan et al., 2006; Fukuda et al., 2009; Ozawa et al., 2007). Consistent with these results, other studies have reported that several genotypes that were not detected in patient stool samples were observed in wastewater or polluted river water (Aw et al., 2009; Rajko-Nenow et al., 2013).

Due to the low number of samples, the epidemic trend of GII.4 variants may not have been well represented by our study. Nonetheless, we noted that three GII.4 variants (Den Haag 2006b, New Orleans 2009, and Sydney 2012) were detected early in our study period, and the Sydney 2012 variant was detected in the late study period in both environmental and clinical samples. The Sydney 2012 variant is known to differ antigenically from other GII.4 variants (Debbink et al., 2013) and to cause more severe clinical symptoms compared to other genotypes and variants (Leshem et al., 2013b; Mai et al., 2013). This property has resulted in its emergence as the predominant strain in Japan and other countries (Fioretti et al., 2014; Fonager et al., 2013; Giammanco et al., 2013; Leshem et al., 2013b; Rahman et al., 2013).

Norovirus-positive patients were detected from the 45th week of 2012, but GII noroviruses were first detected in wastewater from the 50th week of 2012. This difference may, in part, reflect the fact that virological surveillance is based on symptom onset, whereas wastewater concentration reflects the number of people shedding norovirus at the time that the samples were collected. Therefore, the detection of norovirus GII in wastewater may have been delayed in relation to its incidence in

clinical samples. Indeed, norovirus shedding is known to continue for 9-56 days after infection (Aoki et al., 2010; Atmar et al., 2008; Tu et al., 2008).

In 2013, GII.17 was detected in wastewater samples on January 16, 23, 30, and February 6, but this variant was not detected in stool samples or in the national surveillance system for gastroenteritis in Japan during our study period. Representative nucleotide sequences of the largest GII.17 OTUs (including >99% of GII.17 reads in each sample) detected in three wastewater samples (January 23, 30, and February 6, 2013) were 100% identical. This sequence was 99% identical (100% coverage) to the strain detected in Okinawa, Japan, in 2012 (Accession No.: AB901276) and 99%-100% identical (68%-94% coverage) to strains detected in asymptomatic food handlers (Accession Nos. KF773972-3. KF773977, KF773989, and KF774001) and groundwater samples in the Republic of Korea in 2012 (Accession Nos. KC413399-403; Cho et al., 2014). However, there were no other sequences in the DDBJ/EMBL/GenBank database that were more than a 97% match with our sequences. These results suggest that the GII.17 strain caused local outbreaks across East Asia during 2012.

Moreover, GI.6 was the only GI genotype detected in four of the stool samples. Emergence of this genotype, which peaks during the summer, was reported in 2010 in the United States (Leshem et al., 2013a). Although GI.6 also was detected in two wastewater samples, other GI genotypes (GI.3, 4) were detected more frequently, perhaps because our study, which was conducted from November to March, did not include summer months.

Even after pre-processing, many OTUs were produced, and up to eight genotypes were detected in one wastewater sample, demonstrating a high genetic diversity of noroviruses in wastewater. In contrast, only four genotypes were detected in the stool

samples. This discrepancy likely reflects the fact that only one outpatient clinic participated in our surveillance efforts. Moreover, a previous report indicated that 90% of patients with norovirus infections do not seek medical care (Hall et al., 2013). These factors limited the number of samples that could be analyzed ($n = 32$). According to a nationwide Japanese database, 19 genotypes were reported in Japan from September 2012 to August 2013 (National Institute of Infectious Diseases, 2014). Thus, more stool samples would have been needed for a detailed analysis of infectious gastroenteritis cases. However, considering the cost and effort involved in clinical surveillance, wastewater monitoring could more easily provide a detailed distribution of genotypes circulating in the population. Another possible explanation for the discrepancy is that pyrosequencing was used for environmental samples but not for the clinical samples. Since wastewater receives noroviruses from all patients in the served area, multiple genotypes were expected. This is why NGS is beneficial for detection of genotypes in wastewater samples. In contrast, the aim of virological surveillance of gastroenteritis cases is to identify a genotype/variant causing clinical symptom of each patient. Because infection by multiple genotypes/variants is rare (Saito et al., 2014; Thongprachum et al., 2014; Chan et al., 2015), the conventional Sanger sequencing method is preferable both in terms of cost and experimental simplicity including bioinformatics analysis. We have applied both pyrosequencing and Sanger sequencing to three of the stool samples, and the results were identical: two samples contained a single genotype/variant and the other contained two genotypes/variants, possibly caused by co-infection (data not shown). To identify co-infection, we always examined the chromatograms obtained by Sanger sequencing, and confirmed that there were no samples with ambiguous nucleotides (e.g. peaks of multiple bases in the same position).

5. CONCLUSIONS

Our study showed that norovirus monitoring in municipal wastewater by pyrosequencing could provide similar or better information regarding temporal and genomic dynamics of norovirus than the virological surveillance of gastroenteritis cases. We demonstrated that ten more genotypes were detected in wastewater samples as compared with stool samples. Because novel genotypes or variants caused by genome recombination are frequently reported, monitoring of multiple genotypes is important, even if the genotypes are not dominant. This approach would also be useful in regions where nationwide clinical investigation is difficult and gastroenteritis surveillance systems do not exist. Applying this approach in such regions would enhance understanding of global movement of noroviruses. Further investigation, including screening for other pathogens, is necessary to permit virological surveillance based on municipal wastewater monitoring using NGS.

ACKNOWLEDGMENTS

We thank Mr. Yoshimitsu Konta for technical assistance on wastewater analysis and Dr. Hitoshi Onodera for his support with virological surveillance of gastroenteritis cases. This study was supported by CREST from the Japan Science and Technology Agency.

REFERENCES

- Ahmed, S.M., Lopman, B.A., Levy, K., 2013. A Systematic Review and Meta-Analysis of the Global Seasonality of Norovirus. *PLoS ONE* 8, e75922. doi:10.1371/journal.pone.0075922.s002
- Aoki, Y., Suto, A., Mizuta, K., Ahiko, T., Osamu, Matsuzaki, Y., 2010. Duration of norovirus excretion and the longitudinal course of viral load in norovirus-infected elderly patients. *J. Hosp. Infect.* 75, 42–46. doi:10.1016/j.jhin.2009.12.016
- Atmar, R.L., Opekun, A.R., Gilger, M.A., Estes, M.K., Crawford, S.E., Neill, F.H., Graham, D.Y., 2008. Norwalk virus shedding after experimental human infection. *Emerg. Infect. Dis.* 14, 1553–1557. doi:10.3201/eid1410.080117
- Aw, T.G., Gin, K.Y.H., 2010. Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *J. Appl. Microbiol.* 109, 716–730. doi:10.1111/j.1365-2672.2010.04701.x
- Aw, T.G., Gin, K.Y.H., Ean Oon, L.L., Chen, E.X., Woo, C.H., 2009. Prevalence and Genotypes of Human Noroviruses in Tropical Urban Surface Waters and Clinical Samples in Singapore. *Appl. Environ. Microbiol.* 75, 4984–4992. doi:10.1128/AEM.00489-09
- Barreira, D.M.P.G., Ferreira, M.S.R., Fumian, T.M., Checon, R., de Sadovsky, A.D.I., Leite, J.P.G., Miagostovich, M.P., Spano, L.C., 2010. Viral load and genotypes of noroviruses in symptomatic and asymptomatic children in Southeastern Brazil. *J. Clin. Virol.* 47, 60–64. doi:10.1016/j.jcv.2009.11.012
- Barzon, L., Lavezzo, E., Militello, V., Toppo, S., Palù, G., 2011. Applications of next-generation sequencing technologies to diagnostic virology. *Int. J. Mol. Sci.* 12, 7861–7884. doi:10.3390/ijms12117861

544 Blanco Fernandez, M.D., Torres, C., Martinez, L.C., Giordano, M.O., Masachessi, G.,
 545 Barril, P.A., Isa, M.B., Campos, R.H., Nates, S.V., Mbayed, V.A., 2011. Genetic
 546 and evolutionary characterization of norovirus from sewage and surface waters in
 547 Córdoba City, Argentina. 11, 1631–1637. doi:10.1016/j.meegid.2011.06.005
 548 Bucardo, F., Nordgren, J., Carlsson, B., Kindberg, E., Paniagua, M., Mollby, R.,
 549 Svensson, L., 2010. Asymptomatic norovirus infections in Nicaraguan children and
 550 its association with viral properties and histo-blood group antigens. *Pediatr. Infect.*
 551 *Dis. J.* 29, 934–939. doi:10.1097/INF.0b013e3181ed9f2f
 552 Bull, R.A., Tanaka, M.M., White, P.A., 2007. Norovirus recombination. *J. Gen. Virol.*
 553 88, 3347–3359. doi:10.1099/vir.0.83321-0
 554 Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller,
 555 R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009.
 556 The MIQE guidelines: minimum information for publication of quantitative
 557 real-time PCR experiments. *Clinical Chemistry*. doi:10.1373/clinchem.2008.112797
 558 Cantalupo, P.G., Calgua, B., Zhao, G., Hundesa, A., Wier, A.D., Katz, J.P., Grabe, M.,
 559 Hendrix, R.W., Girones, R., Wang, D., Pipas, J.M., 2011. Raw sewage harbors
 560 diverse viral populations. *mBio* 2, e00180–11–e00180–11.
 561 doi:10.1128/mBio.00180-11
 562 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello,
 563 E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley,
 564 S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge,
 565 B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A.,
 566 Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows

567 analysis of high-throughput community sequencing data. *Nat. Meth.* 7, 335–336.
 568 doi:10.1038/nmeth.f.303
 569 Centers for Disease Control and Prevention, 2011. Updated norovirus outbreak
 570 management and disease prevention guidelines. *MMWR. Recomm. Rep.* 60 (3),
 571 1–15.
 572 Chan, M.C.W., Leung, T.F., Chung, T.W.S., Kwok, A.K., Nelson, E.A.S., Lee, N.,
 573 Chan, P.K.S., 2015. Virus Genotype Distribution and Virus Burden in Children and
 574 Adults Hospitalized for Norovirus Gastroenteritis, 2012–2014, Hong Kong. *Sci Rep*
 575 5, 11507. doi:10.1038/srep11507
 576 Chan, M.C.W., Leung, T.F., Kwok, A.K., Lee, N., Chan, P.K.S., 2014. Characteristics
 577 of Patients Infected with Norovirus GII.4 Sydney 2012, Hong Kong, China. *Emerg.*
 578 *Infect. Dis.* 20, 664–667. doi:10.3201/eid2004.131457
 579 Chan, M.C.W., Sung, J.J.Y., Lam, R.K.Y., Chan, P.K.S., Lee, N.L.S., Lai, R.W.M.,
 580 Leung, W.K., 2006. Fecal viral load and norovirus-associated gastroenteritis. *Emerg.*
 581 *Infect. Dis.* 12, 1278–1280. doi:10.3201/eid1208.060081
 582 Cho, H.G., LEE, S.G., Kim, H.S., LEE, J.S., Park, G.W., Cheon, D.S., Jheong, W.H.,
 583 Jho, E.H., LEE, J.B., Paik, S.Y., 2014. Acute gastroenteritis outbreaks associated
 584 with ground-waterborne norovirus in South Korea during 2008–2012. *Epidemiol.*
 585 *Infect.* 142, 1–6. doi:10.1017/S0950268814000247
 586 Debbink, K., Lindesmith, L.C., Donaldson, E.F., Costantini, V., Beltramello, M., Corti,
 587 D., Swanstrom, J., Lanzavecchia, A., Vinjé, J., Baric, R.S., 2013. Emergence of
 588 new pandemic GII.4 Sydney norovirus strain correlates with escape from herd
 589 immunity. *J. Infect. Dis.* 208, 1877–1887. doi:10.1093/infdis/jit370

590 Eden, J.-S., Hewitt, J., Lim, K.L., Boni, M.F., Merif, J., Greening, G., Ratcliff, R.M.,
 591 Holmes, E.C., Tanaka, M.M., Rawlinson, W.D., White, P.A., 2014. The emergence
 592 and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. *Virology*
 593 450-451, 106–113. doi:10.1016/j.virol.2013.12.005
 594 Eden, J.-S., Tanaka, M.M., Boni, M.F., Rawlinson, W.D., White, P.A., 2013.
 595 Recombination within the pandemic norovirus GII.4 lineage. *J. Virol.* 87, 6270–
 596 6282. doi:10.1128/JVI.03464-12
 597 Fioretti, J.M., Bello, G., Rocha, M.S., Victoria, M., Leite, J., 2014. Temporal Dynamics
 598 of Norovirus GII. 4 Variants in Brazil between 2004 and 2012. *PLoS ONE* 9,
 599 e92988. doi:10.1371/journal.pone.0092988.t001
 600 Fonager, J., Hindbaek, L.S., Fischer, T.K., 2013. Rapid emergence and antigenic
 601 diversification of the norovirus 2012 Sydney variant in Denmark, October to
 602 December, 2012. *Eurosurveillance* 18, 2–5.
 603 Fukuda, S., Takao, S., Shigemoto, N., Tanizawa, Y., Seno, M., 2009. Transition of
 604 genotypes associated with norovirus gastroenteritis outbreaks in a limited area of
 605 Japan, Hiroshima Prefecture, during eight epidemic seasons. *Arch. Virol.* 155, 111–
 606 115. doi:10.1007/s00705-009-0528-0
 607 Giammanco, G.M., De Grazia, S., Tummo, F., Bonura, F., Calderaro, A., Buonavoglia,
 608 A., Martella, V., Medici, M.C., 2013. Norovirus GII.4/Sydney/2012 in Italy, winter
 609 2012-2013. *Emerg. Infect. Dis.* 19, 1348–1349. doi:10.3201/eid1908.130619
 610 Glass, R.I., Parashar, U.D., Estes, M.K., 2009. Norovirus gastroenteritis. *N. Engl. J.*
 611 *Med.* 361, 1776–1785. doi:10.1056/NEJMra0804575

612 Hall, A.J., Lopman, B.A., Payne, D.C., Patel, M.M., Gastañaduy, P.A., Vinjé, J.,
 613 Parashar, U.D., 2013. Norovirus Disease in the United States. *Emerg. Infect. Dis.*
 614 19, 1198–1205. doi:10.3201/eid1908.130465
 615 Hata, A., Katayama, H., Kitajima, M., Visvanathan, C., Nol, C., Furumai, H., 2011.
 616 Validation of Internal Controls for Extraction and Amplification of Nucleic Acids
 617 from Enteric Viruses in Water Samples. *Appl. Environ. Microbiol.* 77, 4336–4343.
 618 doi:10.1128/AEM.00077-11
 619 Heim, A., Ebnet, C., Harste, G., Pring-Åkerblom, P., 2003. Rapid and quantitative
 620 detection of human adenovirus DNA by real-time PCR. *J. Med. Virol.* 70, 228–239.
 621 doi:10.1002/jmv.10382
 622 Hernandez-Morga, J., Leon-Felix, J., Peraza-Garay, F., Gil-Salas, B.G., Chaidez, C.,
 623 2009. Detection and characterization of hepatitis A virus and Norovirus in estuarine
 624 water samples using ultrafiltration--RT-PCR integrated methods. *J. Appl. Microbiol.*
 625 106, 1579–1590. doi:10.1111/j.1365-2672.2008.04125.x
 626 Jiang, X., Huang, P.W., Zhong, W.M., Farkas, T., Cubitt, D.W., Matson, D.O., 1999.
 627 Design and evaluation of a primer pair that detects both Norwalk- and Sapporo-like
 628 caliciviruses by RT-PCR. *J. Virol. Meth.* 83, 145–154.
 629 Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B.,
 630 Takeda, N., Katayama, K., 2003. Broadly reactive and highly sensitive assay for
 631 Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J.*
 632 *Clin. Microbiol.* 41, 1548–1557. doi:10.1128/JCM.41.4.1548-1557.2003
 633 Katayama, H., Haramoto, E., Oguma, K., Yamashita, H., Tajima, A., Nakajima, H.,
 634 Ohgaki, S., 2008. One-year monthly quantitative survey of noroviruses,

635 enteroviruses, and adenoviruses in wastewater collected from six plants in Japan.
 636 Water Res. 42, 1441–1448. doi:10.1016/j.watres.2007.10.029
 637 Kitajima, M., Haramoto, E., Phanuwat, C., Katayama, H., Furumai, H., 2012.
 638 Molecular detection and genotyping of human noroviruses in influent and effluent
 639 water at a wastewater treatment plant in Japan. J. Appl. Microbiol. 112, 605–613.
 640 doi:10.1111/j.1365-2672.2012.05231.x
 641 Kitajima, M., Oka, T., Haramoto, E., Phanuwat, C., Takeda, N., Katayama, K.,
 642 Katayama, H., 2010. Genetic diversity of genogroup IV noroviruses in wastewater
 643 in Japan. Lett. Appl. Microbiol. 52, 181–184.
 644 doi:10.1111/j.1472-765X.2010.02980.x
 645 Kitajima, M., Tohya, Y., Matsubara, K., Haramoto, E., Utagawa, E., Katayama, H.,
 646 Ohgaki, S., 2008. Use of murine norovirus as a novel surrogate to evaluate
 647 resistance of human norovirus to free chlorine disinfection in drinking water supply
 648 system. Environ. Eng. Res. 45, 361–370. (in Japanese)
 649 Kojima, S., Kageyama, T., Fukushi, S., Hoshino, F.B., Shinohara, M., Uchida, K.,
 650 Natori, K., Takeda, N., Katayama, K., 2002. Genogroup-specific PCR primers for
 651 detection of Norwalk-like viruses. J. Virol. Meth. 100, 107–114.
 652 Kremer, J.R., Langlet, J., Skrabber, S., Weicherding, P., Weber, B., Cauchie, H.M., De
 653 Landsheer, S., Even, J., Muller, C.P., Hoffmann, L., Mossong, J., 2011. Genetic
 654 diversity of noroviruses from outbreaks, sporadic cases and wastewater in
 655 Luxembourg 2008-2009. Clin. Microbiol. Infect. 17, 1173–1176.
 656 doi:10.1111/j.1469-0691.2010.03407.x
 657 Kroneman, A., Vennema, H., Deforche, K., d Avoort, von, H., Peñaranda, S., Oberste,
 658 M.S., Vinjé, J., Koopmans, M., 2011. An automated genotyping tool for

659 enteroviruses and noroviruses. *J. Clin. Virol.* 51, 121–125.
660 doi:10.1016/j.jcv.2011.03.006

661 La Rosa, G., Iaconelli, M., Pourshaban, M., Muscillo, M., 2010. Detection and
662 molecular characterization of noroviruses from five sewage treatment plants in
663 central Italy. *Water Res.* 44, 1777–1784. doi:10.1016/j.watres.2009.11.055

664 Lee, G.-C., Jheong, W.-H., Jung, G.S., Oh, S.-A., Kim, M.-J., Rhee, O.-J., Park, S., Lee,
665 C.H., 2012. Detection and molecular characterization of human noroviruses in
666 Korean groundwater between 2008 and 2010. *Food. Environ. Virol.* 4, 115–123.
667 doi:10.1007/s12560-012-9084-y

668 Lee, H., Kim, M., Lee, J.E., Lim, M., Kim, M., Kim, J.-M., Jheong, W.-H., Kim, J., Ko,
669 G., 2011. Investigation of norovirus occurrence in groundwater in metropolitan
670 Seoul, Korea. *Sci. Total Environ.* 409, 2078–2084.
671 doi:10.1016/j.scitotenv.2011.01.059

672 Leshem, E., Barclay, L., Wikswo, M., Vega, E., Gregoricus, N., Parashar, U.D., Vinjé,
673 J., Hall, A.J., 2013a. Genotype GI.6 norovirus, United States, 2010–2012. *Emerg.*
674 *Infect. Dis.* 19, 1317–1320. doi:10.3201/eid1908.130445

675 Leshem, E., Wikswo, M., Barclay, L., Brandt, E., Storm, W., Salehi, E., DeSalvo, T.,
676 Davis, T., Saupe, A., Dobbins, G., Booth, H.A., Biggs, C., Garman, K., Woron,
677 A.M., Parashar, U.D., Vinjé, J., Hall, A.J., 2013b. Effects and clinical significance
678 of GII.4 Sydney norovirus, United States, 2012–2013. *Emerg. Infect. Dis.* 19, 1231–
679 1238. doi:10.3201/eid1908.130458

680 Lopman, B., Armstrong, B., Atchison, C., Gray, J.J., 2009. Host, Weather and
681 Virological Factors Drive Norovirus Epidemiology: Time-Series Analysis of

682 Laboratory Surveillance Data in England and Wales. PLoS ONE 4, e6671.
 683 doi:10.1371/journal.pone.0006671.s002
 684 Mai, H., Jin, M., Guo, X., Liu, J., Liu, N., Cong, X., Gao, Y., Wei, L., 2013. Clinical
 685 and Epidemiologic Characteristics of Norovirus GII.4 Sydney during Winter 2012–
 686 13 in Beijing, China following Its Global Emergence. PLoS ONE 8, e71483.
 687 doi:10.1371/journal.pone.0071483.t002
 688 Mans, J., NetshiIKWETA, R., Magwalivha, M., van ZYL, W.B., TAYLOR, M.B., 2013.
 689 Diverse norovirus genotypes identified in sewage-polluted river water in South
 690 Africa. Epidemiol. Infect. 141, 303–313. doi:10.1017/S0950268812000490
 691 Ministry of Health, Labour and Welfare, 2007. Detection protocol of norovirus. Notice
 692 no. 0514004, Safety Division, Pharmaceutical and Food Safety Bureau, Ministry of
 693 Health, Labour and Welfare.
 694 [<http://www.mhlw.go.jp/topics/syokuchu/kanren/kanshi/dl/031105-1a.pdf> (in
 695 Japanese)]
 696 Mounts, A.W., Ando, T., Koopmans, M., Bresee, J.S., Noel, J., Glass, R.I., 2000. Cold
 697 Weather Seasonality of Gastroenteritis Associated with Norwalk-like Viruses. J.
 698 Infect. Dis. 5, S284–287.
 699 National Institute of Infectious Diseases, 2014. Gastroenteritis viruses detected from
 700 patients in Japan, 2004-2014. Infectious Agents Surveillance Report, Sep. 29th
 701 2014. [<https://nesid3g.mhlw.go.jp/Byogentai/Pdf/data96j.pdf>. (in Japanese)]
 702 Nordgren, J., Matussek, A., Mattsson, A., Svensson, L., Lindgren, P.-E., 2009.
 703 Prevalence of norovirus and factors influencing virus concentrations during one
 704 year in a full-scale wastewater treatment plant. Water Res. 43, 1117–1125.
 705 doi:10.1016/j.watres.2008.11.053

706 Oka, T., Katayama, K., Hansman, G.S., Kageyama, T., Ogawa, S., Wu, F.-T., White,
 707 P.A., Takeda, N., 2006. Detection of human sapovirus by real-time reverse
 708 transcription-polymerase chain reaction. *J. Med. Virol.* 78, 1347–1353.
 709 doi:10.1002/jmv.20699

710 Okabayashi, T., Yokota, S.-I., Ohkoshi, Y., Ohuchi, H., Yoshida, Y., Kikuchi, M., Yano,
 711 K., Fujii, N., 2008. Occurrence of norovirus infections unrelated to norovirus
 712 outbreaks in an asymptomatic food handler population. *J. Clin. Microbiol.* 46,
 713 1985–1988. doi:10.1128/JCM.00305-08

714 Ozawa, K., Oka, T., Takeda, N., Hansman, G.S., 2007. Norovirus infections in
 715 symptomatic and asymptomatic food handlers in Japan. *J. Clin. Microbiol.* 45,
 716 3996–4005. doi:10.1128/JCM.01516-07

717 Pang, X.L., Lee, B., Boroumand, N., Leblanc, B., Preiksaitis, J.K., Yu Ip, C.C., 2004.
 718 Increased detection of rotavirus using a real time reverse transcription-polymerase
 719 chain reaction (RT-PCR) assay in stool specimens from children with diarrhea. *J.*
 720 *Med. Virol.* 72, 496–501. doi:10.1002/jmv.20009

721 Patel, M.M., Hall, A.J., Vinjé, J., Parashar, U.D., 2009. Noroviruses: a comprehensive
 722 review. *J. Clin. Virol.* 44, 1–8. doi:10.1016/j.jcv.2008.10.009

723 Patel, M.M., Widdowson, M.-A., Glass, R.I., Akazawa, K., Vinjé, J., Parashar, U.D.,
 724 2008. Systematic Literature Review of Role of Noroviruses in Sporadic
 725 Gastroenteritis. *Emerg. Infect. Dis.* 14, 1224–1231. doi:10.3201/eid1408.071114

726 Pérez-Sautu, U., Sano, D., Guix, S., Kasimir, G., Pintó, R.M., Bosch, A., 2012. Human
 727 norovirus occurrence and diversity in the Llobregat river catchment, Spain. *Environ.*
 728 *Microbiol.* 14, 494–502. doi:10.1111/j.1462-2920.2011.02642.x

729 Quince, C., Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing noise from
 730 pyrosequenced amplicons. *BMC Bioinformatics* 12, 38.
 731 doi:10.1186/1471-2105-12-38

732 Rahman, M., Nahar, S., Afrad, M.H., Faruque, A.S.G., Azim, T., 2013. Norovirus
 733 Variant GII. 4/Sydney/2012, Bangladesh. *Emerg. Infect. Dis.* 19, 1347–1348.
 734 doi:10.3201/eid1908.130227

735 Rajko-Nenow, P., Waters, A., Keaveney, S., Flannery, J., Tuite, G., Coughlan, S.,
 736 O'Flaherty, V., Doré, W., 2013. Norovirus genotypes present in oysters and in
 737 effluent from a wastewater treatment plant during the seasonal peak of infections in
 738 Ireland in 2010. *Appl. Environ. Microbiol.* 79, 2578–2587.
 739 doi:10.1128/AEM.03557-12

740 Saito, M., Goel-Apaza, S., Espetia, S., Velasquez, D., Cabrera, L., Loli, S., Crabtree,
 741 J.E., Black, R.E., Kosek, M., Checkley, W., Zimic, M., Bern, C., Cama, V., Gilman,
 742 R.H., Norovirus Working Group in Peru, 2014. Multiple norovirus infections in a
 743 birth cohort in a Peruvian Periurban community. *Clin. Infect. Dis.* 58, 483–491.
 744 doi:10.1093/cid/cit763

745 Siebenga, J.J., Vennema, H., Zheng, D.P., Vinjé, J., Lee, B.E., Pang, X.L., Ho, E.C.M.,
 746 Lim, W., Choudekar, A., Broor, S., Halperin, T., Rasool, N.B.G., Hewitt, J.,
 747 Greening, G.E., Jin, M., Duan, Z.J., Lucero, Y., O'Ryan, M., Hoehne, M., Schreier,
 748 E., Ratcliff, R.M., White, P.A., Iritani, N., Reuter, G., Koopmans, M., 2009.
 749 Norovirus illness is a global problem: emergence and spread of norovirus GII.4
 750 variants, 2001-2007. *J. Infect. Dis.* 200, 802–812. doi:10.1086/605127

751 Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5:
 752 molecular evolutionary genetics analysis using maximum likelihood, evolutionary

distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.
doi:10.1093/molbev/msr121

Thongprachum, A., Chan-it, W., Khamrin, P., Saparpakorn, P., Okitsu, S., Takanashi, S., Mizuguchi, M., Hayakawa, S., Maneekarn, N., Ushijima, H., 2014. Molecular epidemiology of norovirus associated with gastroenteritis and emergence of norovirus GII.4 variant 2012 in Japanese pediatric patients. *Infection, Genetics and Evolution* 23, 65–73. doi:10.1016/j.meegid.2014.01.030

Tu, E.T.V., Bull, R.A., Greening, G.E., Hewitt, J., Lyon, M.J., Marshall, J.A., McIver, C.J., Rawlinson, W.D., White, P.A., 2008. Epidemics of Gastroenteritis during 2006 Were Associated with the Spread of Norovirus GII.4 Variants 2006a and 2006b. *Clin. Infect. Dis.* 46, 413–420. doi:10.1086/525259

Ueki, Y., Sano, D., Watanabe, T., Akiyama, K., Omura, T., 2005. Norovirus pathway in water environment estimated by genetic analysis of strains from patients of gastroenteritis, sewage, treated wastewater, river water and oysters. *Water Res.* 39, 4271–4280. doi:10.1016/j.watres.2005.06.035

Vega, E., Barclay, L., Gregoricus, N., Shirley, S.H., Lee, D., Vinjé, J., 2014. Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *J. Clin. Microbiol.* 52, 147–155. doi:10.1128/JCM.02680-13

Yoneda, M., Okayama, A., Kitahori, Y., 2014. Epidemiological Characteristics of Norovirus Associated with Sporadic Gastroenteritis among Children from the 2006/2007 to 2011/2012 Season in Nara Prefecture, Japan. *Intervirology* 57, 31–35. doi:10.1159/000353852

Zheng, D.P., Ando, T., Fankhauser, R.L., Beard, R.S., Glass, R.I., Monroe, S.S., 2006. Norovirus classification and proposed strain nomenclature. *Virology* 346, 312–323. doi:10.1016/j.virol.2005.11.015

FIGURE LEGENDS

Fig. 1 Concentrations (as determined by qPCR) of norovirus genotypes I and II (GI and GII) in wastewater samples collected at the indicated intervals.

< LOQ, below the limit of quantification; ND, not detected.

Fig. 2 Genotypes and GII.4 variants identified from wastewater samples.

The numbers show ratios calculated as the number of reads assigned to each genotype/variant divided by the number of reads identified as norovirus in each sample.

Heat map (color coding) is as indicated.

Fig. 3 Number of infectious gastroenteritis cases and norovirus-positive/negative stool samples collected each week.

For a given interval, the number of infectious gastroenteritis cases in the clinic is plotted as solid line (asterisk (*)); the numbers and variants of norovirus are plotted as a bar graph with color coding as indicated. Discordance between the number of the stool samples and number of the infectious gastroenteritis cases in a given week reflects the number of stool samples that could not be obtained for analysis.

Fig. 4 Genotypes and GII.4 variants detected in stool samples.

Fig. 5 Phylogenetic tree of GII.4 sequences obtained from wastewater samples and clinical samples. The number on each branch shows the bootstrap value. Because many OTUs were obtained from the wastewater samples, representative sequences of OTUs

802 containing more than 1000 reads were used for the analysis. The sequences obtained
803 from the wastewater samples are designated with names starting with “W”, followed by
804 date, variant (2006b, Den Haag 2006b; 2009, New Orleans 2009; 2012, Sydney 2012),
805 and identification number of OTU (e.g., W_Dec28_2012_OTU1). The sequences
806 obtained from the clinical samples are designated with names starting with “P”,
807 followed by variant and date of onset (e.g., P_2006b_Nov29). The reference sequences
808 are designated with names starting with “II.4” followed by accession number (e.g.,
809 II.4|2012|JX459908).

Fig. 1 Concentrations (as determined by qPCR) of norovirus genotypes I and II (GI and GII) in wastewater samples collected at the indicated intervals.

< LOQ, below the limit of quantification; ND, not detected.

year		2012											2013										
week		46	47	48	49	50	51	52	1	2	3	4	5	6	7	8	9	10	11				
mo/day		11/14	11/21	11/28	12/5	12/12	12/19	12/28		1/9	1/16	1/23	1/30	2/6	2/13	2/20	2/27	3/6	3/13				
GI	GI.1	100.0						13.3															
	GI.3									19.0	99.7		98.9			96.9		100.0					
	GI.4				100.0		100.0	10.0		69.5		100.0	0.0	99.9	99.7		99.6						
	GI.6							47.7		11.5													
	GI.7							29.0			0.3		1.1	0.1	0.3	3.1	0.4		100.0				
GII	GII.2	100.0				31.3		31.2		21.0	21.7	45.6	23.5	19.5	32.4	23.6	18.2						
	GII.4					22.1		29.9		6.4	19.9		13.1	15.5	27.8	23.7	18.5	20.1	39.8				
	GII.5														39.8		18.2	14.3					
	GII.6													1.6				4.4					
	GII.7									3.1							27.1		60.2				
	GII.12									32.7													
	GII.13																	43.5					
	GII.14				100.0	46.3		38.9		36.8	58.0		32.5	26.5		52.7	17.1	16.9					
	GII.17										0.3	54.4	30.2	36.9									
	GII (not assigned)					0.2							0.7	0.1			0.9	0.8					
GII.4 variants	Den Haag 2006b					92.3		23.7			30.0												
	New Orleans 2009							9.1															
	Sydney 2012							62.6		76.5	61.5		95.2	84.8	96.6	94.2	87.7	90.5	98.6				
	GII.4 (not assigned)					7.7		4.6		23.5	8.6		4.8	15.2	3.4	5.8	12.3	9.5	1.4				

Fig. 2 Genotypes and GII.4 variants identified from wastewater samples.

The numbers show ratios calculated as the number of reads assigned to each genotype/variant divided by the number of reads identified as norovirus in each sample. Heat map (color coding) is as indicated.

Fig. 3 Number of infectious gastroenteritis cases and norovirus-positive/negative stool samples collected each week.

For a given interval, the number of infectious gastroenteritis cases in the clinic is plotted as solid line (asterisk (*)); the numbers and variants of norovirus are plotted as a bar graph with color coding as indicated. Discordance between the number of the stool samples and number of the infectious gastroenteritis cases in a given week reflects the number of stool samples that could not be obtained for analysis.

FIG 4 Genotypes and GII.4 variants detected in stool samples.

Fig. 5 Phylogenetic tree of GII.4 sequences obtained from wastewater samples and clinical samples. The number on each branch shows the bootstrap value. Because many OTUs were obtained from the wastewater samples, representative sequences of OTUs containing more than 1000 reads were used for the analysis. The sequences obtained from the wastewater samples are designated with names starting with “W”, followed by date, variant (2006b, Den Haag 2006b; 2009, New Orleans 2009; 2012, Sydney 2012), and identification number of OTU (e.g., W_Dec28_2012_OTU1). The sequences obtained from the clinical samples are designated with names starting with “P”, followed by variant and date of onset (e.g., P_2006b_Nov29). The reference sequences are designated with names starting with “H.4” followed by accession number (e.g., H.4|2012|JX459908).

Noroviruses circulating in human population