Phylogenetic Analyses Suggest that Factors Other Than the Capsid Protein Play a Role in the Epidemic Potential of GII.2 Norovirus

Kentaro Tohma, Cara J. Lepore, Lauren A. Ford-Siltz, (iD) Gabriel I. Parra

Division of Viral Products, Food and Drug Administration, Silver Spring, Maryland, USA

ABSTRACT Norovirus is the leading cause of acute gastroenteritis worldwide. For over two decades, a single genotype (GII.4) has been responsible for most norovirus-associated cases. However, during the winter of 2014 to 2015, the GII.4 strains were displaced by a rarely detected genotype (GII.17) in several countries of the Asian continent. Moreover, during the winter of 2016 to 2017, the GII.2 strain reemerged as predominant in different countries worldwide. This reemerging GII.2 strain is a recombinant virus that presents a GII.P16 polymerase genotype. In this study, we investigated the evolutionary dynamics of GII.2 to determine the mechanism of this sudden emergence in the human population. The phylogenetic analyses indicated strong linear evolution of the VP1-encoding sequence, albeit with minor changes in the amino acid sequence over time. Without major genetic differences among the strains, a clustering based on the polymerase genotype was observed in the tree. This association did not affect the substitution rate of the VP1. Phylogenetic analyses of the polymerase region showed that reemerging GII.P16-GII.2 strains diverged into a new cluster, with a small number of amino acid substitutions detected on the surface of the associated polymerase. Thus, besides recombination or antigenic shift, point mutations in nonstructural proteins could also lead to novel properties with epidemic potential in different norovirus genotypes.

IMPORTANCE Noroviruses are a major cause of gastroenteritis worldwide. Currently, there is no vaccine or specific antiviral available to treat norovirus disease. Multiple norovirus strains infect humans, but a single genotype (GII.4) has been regarded as the most important cause of viral gastroenteritis outbreaks worldwide. Its persistence and predominance have been explained by the continuous replacement of variants that present new antigenic properties on their capsid protein, thus evading the herd immunity acquired to the previous variants. Over the last three seasons, minor genotypes have displaced the GII.4 viruses as the predominant strains. One of these genotypes, GII.2, reemerged as predominant during 2016 to 2017. Here we show that factors such as minor changes in the polymerase may have driven the reemergence of GII.2 during the last season. A better understanding of norovirus diversity is important for the development of effective treatments against noroviruses.

KEYWORDS calicivirus, noroviruses, phylogenetic analysis, transmissible gastroenteritis virus

Norovirus is an important cause of acute gastroenteritis worldwide, affecting people in all age groups. Norovirus infections are mainly associated with outbreaks occurring during the winter season in enclosed communities such as schools, nursing homes, or hospitals (1–4). Although norovirus infection is self-limiting in healthy individuals, it can cause severe symptoms in vulnerable populations (i.e., young children, the elderly, or immunocompromised individuals). Thus, norovirus has been

associated with up to 200,000 deaths per year, mostly in children in developing countries (5, 6).

Norovirus has an ~7.5-kb single-stranded positive-sense RNA genome that is organized into three open reading frames (ORFs). The ORF1 encodes a nonstructural polyprotein that includes the RNA-dependent RNA polymerase (RdRp). The ORF2 encodes the major capsid protein (VP1), and the ORF3 encodes the minor capsid protein (VP2). VP1 is structurally divided into a shell (S) and protruding (P) domain, with the major antigenic sites located in the P domain.

Noroviruses are a highly diverse group of viruses. They are grouped into seven genogroups (GI to GVII) and more than 30 genotypes, based on the sequence differences of their VP1 proteins (7). Norovirus typing is currently based on a dual system, where sequence information from the RdRp (encoded by ORF1) and VP1 (encoded by ORF2) is required for precise molecular identification of norovirus strains (8). In several viruses, phylogenetic discrepancies have been reported in analyzing sequences from the ORF1 and ORF2 regions; thus, the junction between ORF1 and ORF2 is considered a hot spot site for norovirus recombination (9). Survey studies have shown that humans are primarily infected by GI and GII strains; however, the GII.4 genotype has been responsible for most of the norovirus-associated cases of acute gastroenteritis worldwide for almost two decades (10–12). The persistence and predominance of GII.4 strains have been linked to the constant evolution of their VP1, which results in the replacement of antigenically distinct variants every 3 to 8 years (13, 14). During the winter of 2014 to 2015, many countries in Asia reported a sudden increase in the detection of GII.17 strains, which displaced the GII.4 strains as the major cause of norovirus-associated disease (15–17). GII.17 has been rarely reported for over 30 years, but studies suggested that the fast-evolving nature of the novel GII.17 strains enabled an antigenic shift in VP1 (16, 18). The emergence of this GII.17 was also associated with the report of a new polymerase genotype, GII.P17 (15, 19). Although the acquisition of this new polymerase could have been associated with the faster evolution of the novel GII.17 strains, the mechanism of this emergence is not completely understood (18).

Just recently, Niendorf et al. reported the predominance of a recombinant GII.P16-GII.2 strain in Germany (20). The predominance of this recombinant strain was also reported in China and Japan during the 2016–2017 season, mainly in childcare centers (21–23). Although the GII.2 genotype is considered one of the four most common norovirus genotypes associated with sporadic cases in children, global analyses found that it only accounts for <2% of all strains (24). Outbreaks and sporadic infections caused by GII.2 strains have been reported since 1979 (25–29). GII.2 noroviruses have been associated with multiple different polymerases (i.e., GII.P2, GII.P12, GII.P16, GII.P21, GII.P22, GII.Pc, and GII.Ph); however, GII.P16-GII.2 noroviruses have been detected only in the last decade (30). Here, we investigated the evolutionary dynamics of GII.2 strains to elucidate the mechanisms that led to the predominance of the GII.P16-GII.2 strains during the 2016–2017 season.

RESULTS

Clock-like evolution of GII.2 VP1-encoding region. To investigate the evolutionary dynamics of GII.2 VP1, sequence data were collected for GII.2 strains with a nearly complete ORF2 (1,337 nucleotides [nt]) available in the GenBank database ($n = 151$; see Table S1 in the supplemental material). The 151 VP1 sequences from the included GII.2 noroviruses were associated with seven different polymerase genotypes (GII.P2, GII.P12, GII.P16, GII.P21, GII.P22, GII.Pc, and GII.Ph), with GII.P2 (51/151, 34%) and GII.P16 (55/151, 36%) being the most commonly reported. The maximum likelihood (ML) tree showed that the GII.2 strains were mostly, but not always, grouped according to their polymerase genotypes (Fig. 1a). The reemerging GII.P16-GII.2 strains clustered together but were separated from pre-2016 GII.2 strains. Despite this apparent clustering, the pairwise p-distance of the GII.2 sequences indicated a small variation at the nt level, with no clear partition to define variants within GII.2 strains (see Fig. S1 in the supplemental material). The exceptions to this were three GII.P21-GII.2 and three
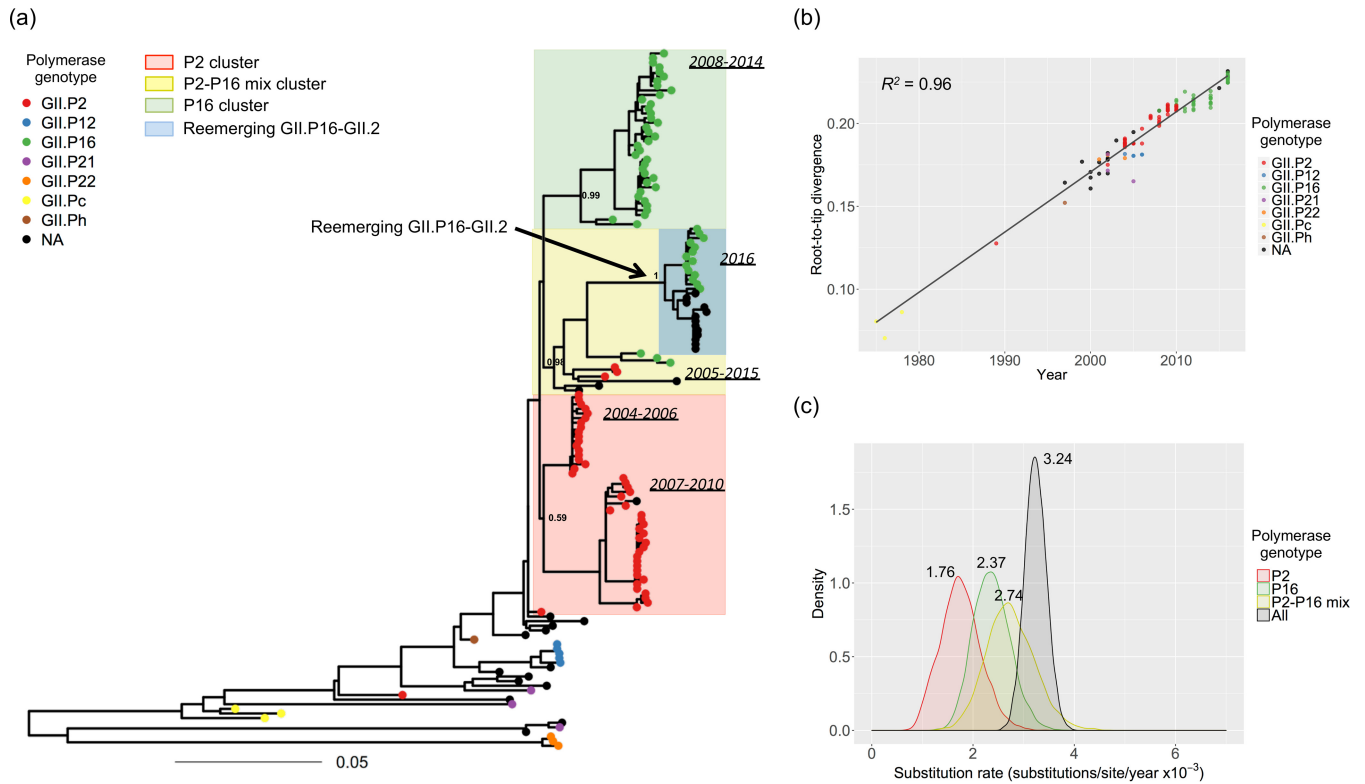
FIG 1 Evolutionary dynamics of the major capsid protein of GII.2 noroviruses. (a) Maximum likelihood phylogenetic tree of VP1-encoding (nt) sequences from GII.2 strains. The numbers on the ancestral nodes of the clusters represent the node support value calculated by the approximate likelihood-ratio test using PhyML. (b) Root-to-tip divergence plot of VP1 nt sequences of GII.2 strains. The x axis indicates isolation year, and the y axis shows the root-to-tip divergence on the maximum likelihood phylogenetic tree. The black line indicates a linear regression line of the root-to-tip divergence and isolation year. A total of 151 sequences were included for both analyses. Each strain is represented by a circle and colored according to its polymerase genotype. Strains whose polymerase genotype was not available are indicated in black. The genomic region used in the analyses spanned nt 5085 to 6419 relative to Snow Mountain virus (GenBank accession number AY134748). (c) The kernel density plot indicates the posterior estimates of substitution rate (substitutions/site/year) of VP1 nt sequences. To calculate the substitution rate of the VP1 associated with different polymerases, strains were clustered using the VP1 phylogenetic tree. Thus, the P2 cluster is shaded in red, the P16 cluster in green, and the P2-P16 mix in yellow. The overall rate is indicated by black shading.

GII.P22-GII.2 strains detected during 2002 to 2004 (Fig. 1). The root-to-tip divergence plot of VP1-encoding sequences showed strong clock-like evolution of the VP1, with a coefficient of determination (R^2) value of 0.96 (Fig. 1b). The GII.P2-GII.2 strains were detected from 1989 to 2010 and were then replaced by the GII.P16-GII.2 strains that started to circulate in 2008. This replacement did not alter the evolutionary rate of VP1, which kept diverging on the linear regression line regardless of the polymerase genotypes (Fig. 1b). To determine whether the acquisition of a different polymerase affected the substitution rate, we grouped the strains based on their polymerase genotype and VP1 phylogenetic clustering. Thus, four data sets were used for these analyses: one that included all sequences, one that included strains from a P2 cluster (red shaded, Fig. 1a), one that contained strains from a P16 cluster (green shaded, Fig. 1a), and another one that contained strains with P2 or P16 but that clustered together in the phylogenetic tree (yellow shaded, Fig. 1a). Of note is that the recently reported reemerging strains, the recombinant GII.P16-GII.2 strains, were located in the P2-P16 mix cluster (blue shaded, Fig. 1a). The mean substitution rate of the overall VP1 sequences was 3.24×10^{-3} substitutions/site/year (95% highest posterior density [HPD] interval: 2.84×10^{-3} to 3.67×10^{-3}). The mean substitution rates of strains bearing GII.P2, GII.P16, or GII.P2-P16 were similar to each other (i.e., 1.75×10^{-3}, 2.37×10^{-3}, and 2.74×10^{-3} substitutions/site/year, respectively [Fig. 1c]); however, the substitution rate of the GII.P2 strains was slightly lower than the overall rate. These substitution rates were also observed when the complete VP1 sequences were used in the analyses (Fig. S2).
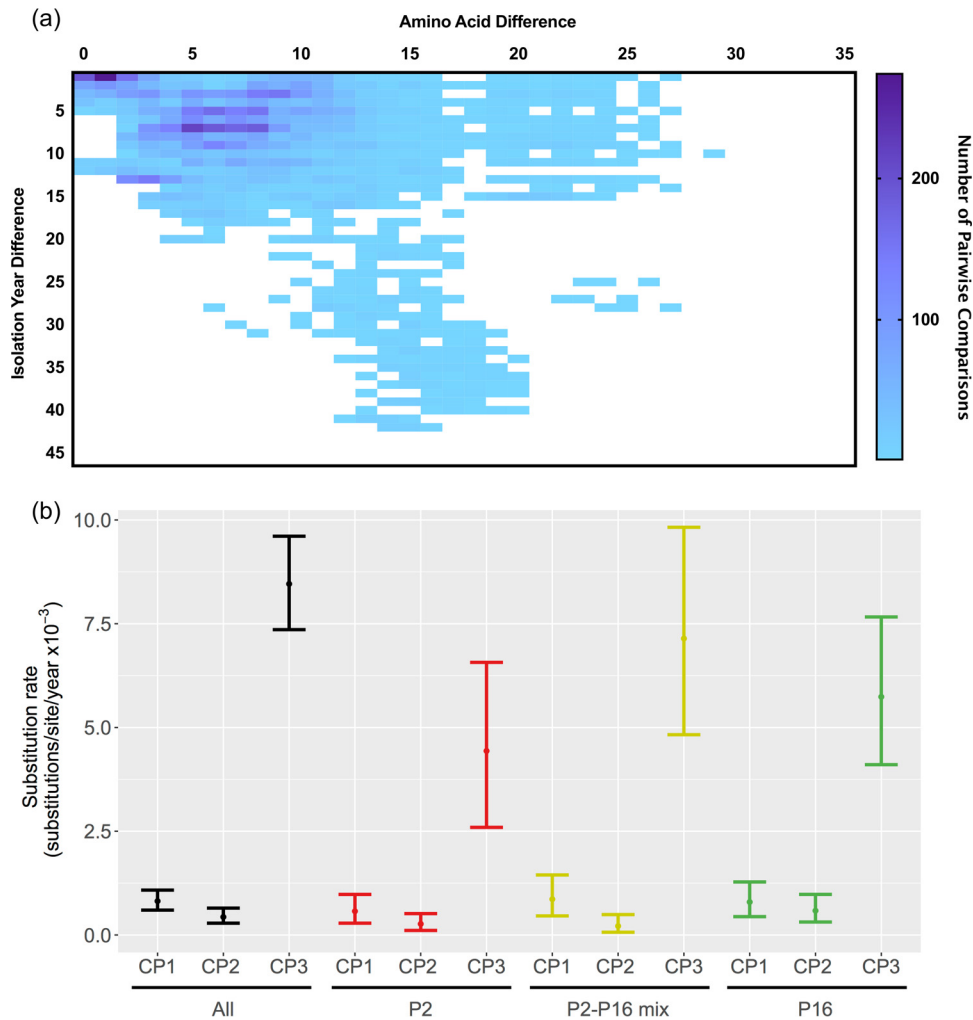
FIG 2 Amino acid diversification of the major capsid protein of GII.2 noroviruses. (a) Heat map indicating the accumulation of aa mutations in the partial VP1 sequences of GII.2 strains (445 amino acids) over time. The x axis indicates the pairwise aa mutations, and the y axis indicates isolation year. The number of pairwise comparisons is represented by the color gradient in the heat map. (b) The plot indicates the substitution rates (substitutions/site/year) of VP1 nucleotide sequences at each codon position. The mean and the 95% highest posterior density intervals are indicated. The substitution rates were calculated using the same sequence matrix and groups as those represented in Fig. 1.

Small divergence of amino acid (aa) sequences during a 42-year period. Although nt sequences showed clock-like linear evolution, the aa substitutions did not accumulate over time (Fig. 2a). The plot of aa differences over time indicates minor divergence (≤5%, a cutoff value to define intragenotype variants) at the VP1 for almost four decades, in concordance with the static nature of non-GII.4 noroviruses (14). Only three GII.2 strains detected in Japan during 2001 to 2004 and associated with the GII.P22 polymerase showed >5% aa divergence when the complete VP1 sequence was used (Fig. S3). This discrepancy in the clock-like evolution of the nt sequences and the limited diversity at the aa level were mostly driven by the high rate of evolution presented at the third codon positions, which led to synonymous substitutions (Fig. 2b). This pattern was observed regardless of the polymerase genotypes associated with the strain (Fig. 2b). The substitution rates at the first and second codon positions were comparable among all data sets, suggesting similar rates of nonsynonymous substitution. In addition, the phylogenetic tree of VP1 inferred using aa sequences showed no clear divergence among the P2 and P2-P16 clusters, including the reemerging GII.P16-GII.2 strains (Fig. S4). Of note, the aa sequences from the reemerging

TABLE 1 Amino acid substitution observed in the consensus VP1 sequence among phylogenetic clusters

Strain or cluster	No. of sequences	Amino acid substitution at codon:								
		71	130	303	335	344	354	364	386	400
Reemerging GII.P16-GII.2 (2016)	27	A	V	V	I (1 V)[a]	S (1G)[a]	G	A	N	E
P2-P16 mix cluster[b] (2005–2015) (without reemerging strains)	9	A	V	V	I (2 V)[a]	S	G	A	N	E
P16 cluster[b] (2008–2014)	38	S	V	V	V	T (1 I, 2 A)[a]	A	A	N	D (2 E)[a]
P2 cluster[b] (2007–2010)	29	A	I (4 V)[a]	I (1 V)[a]	I	S	G	A	S	E
P2 cluster[b] (2004–2006)	18	A	V	V	V	S	G	S (1P)[a]	N	E

[a]Data in parentheses represent minor variants within the cluster.
[b]Phylogenetic clustering as shown in Fig. 1, assigned based on the phylogenetic clustering of the VP1-encoding sequence and the associated polymerase genotype.

GII.P16-GII.2 strains were almost identical (≤2 aa) to those from some of the GII.2 strains in the P2 and P2-P16 mix clusters detected in 2004 to 2009 (Table 1; Fig. S4). Of note is that 11 of 27 reemerging GII.2 strains, including GII.P16-GII.2/CUHK-NS-1082/2016/HKG (KY771081; full-length VP1), GII.2/CUHK-NS-1231/2016/HKG (KY421044; full-length VP1), and strains isolated in Germany (KY357450, KY357451, KY357452, KY357453, KY357454, KY357456, KY357457, KY357458, and KY357462; partial VP1), presented 100% identical capsid protein sequences compared with GII.P2-GII.2/OH06023/2006/JPN (AB662863) and GII.2/Vaals87/2005/NLD (AB281090) strains.

Diversifying selection through the evolution of GII.2 VP1. To determine whether positive selection on the VP1 led to the reemergence and posterior predominance of the GII.P16-GII.2 strains, we performed selection analyses using internal fixed-effect likelihood (iFEL) and mixed-effect model of evolution (MEME) methods (31, 32). The iFEL test estimates the site-by-site positive selection along the internal branches of the phylogenetic tree, assuming the same nonsynonymous/synonymous substitution ratio on the branches. The MEME test, on the other hand, estimates the branch-to-branch positive selections at individual sites to identify the episodic selections (episodes of diversifying selection in a portion of branches). As a result, significant positive selection ($P < 0.05$) was detected at eight codon sites, among which codon site 345 was detected by both methods (Table 2). Five of those eight sites were located on the P2 domain of the virus particle. All of the codons identified using the partial VP1 sequences were also detected as positively selected using the full-length VP1 data set, which showed three additional positively selected sites on the P2 domain (Table S2). Of note, only a few positive sites were observed on the branches diversifying the clusters, and empirical Bayes factor (EBF) data indicated no evidence of positive selection on the branches of reemerging GII.P16-GII.2 strains (Fig. 3).

Taking the data together, the reemergence of GII.P16-GII.2 strains during 2016 did not involve changes in the substitution rate or acquisition of aa mutations in the major capsid protein.

Phylogeny of RdRp-encoding region. We also performed phylogenetic analyses of the RdRp region of GII.P2 and GII.P16, the most commonly detected polymerase genotypes of GII.2. A total of 72 GII.P2 sequences and 131 GII.P16 sequences were collected for the analyses (Table S1). GII.P2 was associated with the GII.2 capsid, and

TABLE 2 Codon position of VP1 sequences (1,337 nucleotides) with positive selection

Positively selected site	Method(s) (P value)	Domain
24	MEME[a] (0.015)	N terminus
78	iFEL[b] (0.015)	Shell
275	MEME (0.016)	P1
344	MEME (0.039)	P2 (surface)
345	iFEL (0.01), MEME (0.020)	P2 (surface)
384	MEME (0.01)	P2 (surface)
385	MEME (0.024)	P2 (surface)
397	iFEL (0.045)	P2 (surface)

[a]MEME, mixed-effect model of evolution.
[b]iFEL, internal fixed-effect likelihood methods.

(a) Codon 344

P2 cluster
P2-P16 mix cluster
P16 cluster
Reemerging GII.P16-GII.2

Reemerging
GII.P16-GII.2

(b) Codon 345

Reemerging
GII.P16-GII.2

(c) Codon 384

Reemerging
GII.P16-GII.2

(d) Codon 385

Reemerging
GII.P16-GII.2
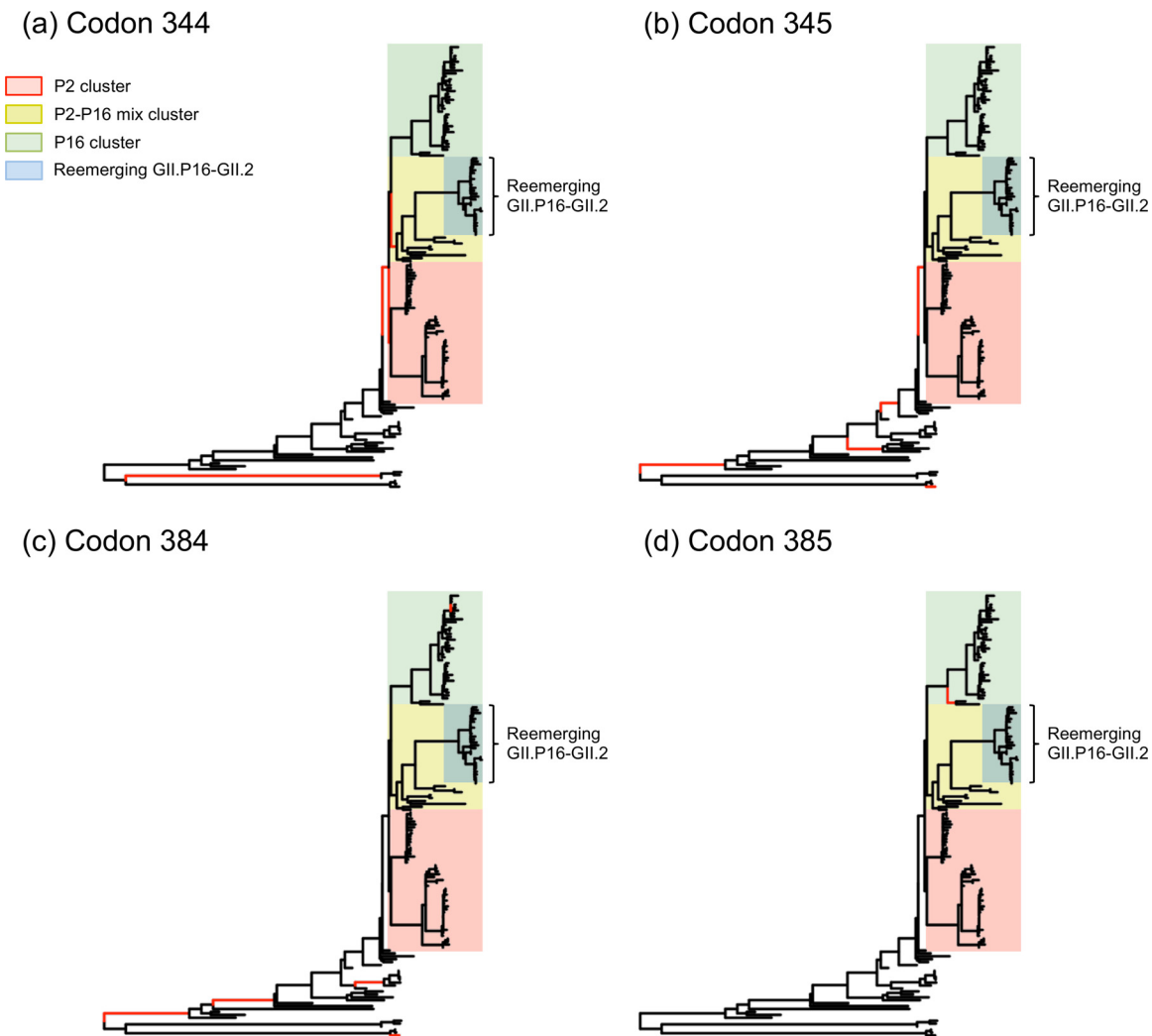
FIG 3 Positive evolution of the major capsid protein of GII.2 noroviruses. Maximum likelihood trees of VP1 nucleotide (nt) sequences of GII.2 indicate the branches under possible positive selection. The branches with evidence of positive selection (empirical Bayes factor of >10, $P < 0.05$; estimated using the mixed-effect model of evolution [MEME] method) at codon positions 344 (a), 345 (b), 384 (c), and 385 (d) are represented by red branch lines. The color shade indicates the phylogenetic clustering based on the polymerase genotype as indicated in Fig. 1. Reemerging GII.P16-GII.2 strains are indicated in blue.

GII.P16 was associated with seven different capsid genotypes (GII.P16-GII.2, GII.P16-GII.3, GII.P16-GII.4, GII.P16-GII.10, GII.P16-GII.13, GII.P16-GII.16, and GII.P16-GII.17). The phylogenetic clustering shown on the VP1 nt tree was also observed in the tree of RdRp nt sequences of GII.P2 and GII.P16, although the support for the P2 cluster in the tree of GII.P2 was weak (Fig. 4a). The root-to-tip divergence plot showed moderate clock-like evolution over time ($R^2 = 0.60$ and $R^2 = 0.65$ for GII.P2 and GII.P16, respectively) (Fig. 4b). The reemerging GII.P16-GII.2 strains were clustered together with the emerging recombinant GII.P16-GII.4_Sydney strains detected in 2015 to 2016 (Fig. 4a and b). It is difficult to determine whether these two strains came from the same parental strain or whether one originated from the other. However, the phylogenetically similar GII.P16-GII.2 strains have been circulating since 2011, and the regression line in the root-to-tip analyses of both RdRp and VP1 shows a linear evolution from 2011 to the 2016–2017 GII.P16-GII.2 strains. This suggests that the GII.4_Sydney strain acquired GII.P16 from GII.P16-GII.2 strains. The substitution rate of all GII.P16 strains was 2.68×10^{-3} substitutions/site/year (95% HPD interval: 2.12×10^{-3} to 3.21×10^{-3}). It slightly decreased when reemerging GII.P16 strains detected during 2015 to 2016 (GII.P16-GII.2 and GII.P16-GII.4_Sydney) were removed from the analyses ($2.17E-3$ substitutions/site/
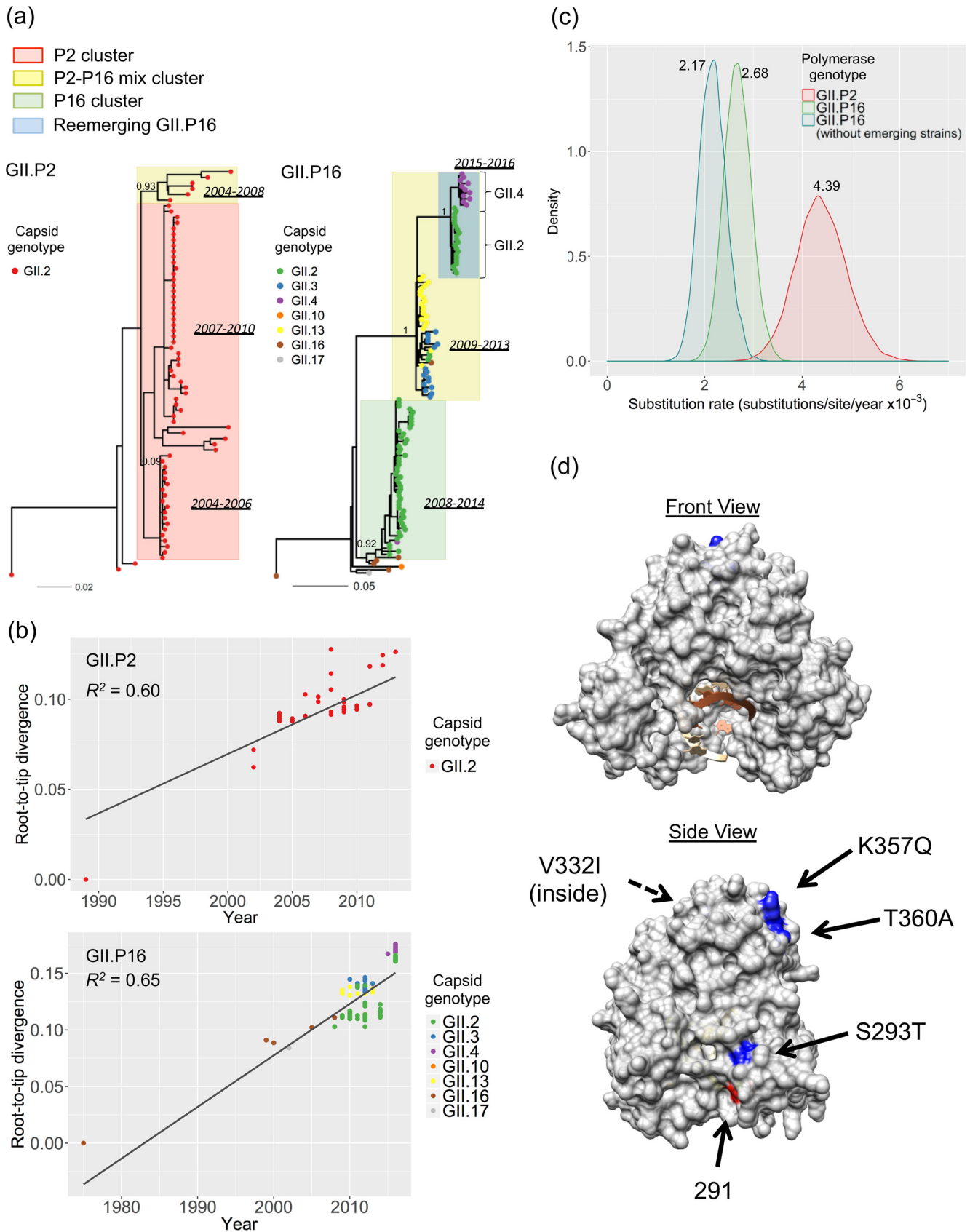
FIG 4 Evolutionary dynamics of the GII.2 polymerase. (a) Maximum likelihood trees of RdRp-encoding (nt) sequences of strains GII.P2 ($n = 72$) and GII.P16 ($n = 131$). Phylogenetic clusters based on VP1 nt sequences (Fig. 1) are shown here; the P2 cluster is represented in red, the P16 cluster in green, and the

year [95% HPD interval: 1.64×10^{-3} to 2.72×10^{-3}]), but their 95% HPD intervals overlapped with each other, suggesting no significant difference (Fig. 4c). The substitution rate of the RdRp-encoding sequence of GII.P2 ($4.39E-3$ substitutions/site/year [95% HPD interval: 3.41×10^{-3} to 5.44×10^{-3}]) was higher than that of GII.P16 (Fig. 4c), but the aa tree showed no variation in GII.P2 (Fig. S5a). On the other hand, the aa tree of GII.P16 showed phylogenetic clustering similar to what was observed in the nt tree. The P2-P16 mix cluster and the P16 cluster were clearly separated in the aa tree of GII.P16 (Fig. S5b). The reemerging GII.P16 strains circulating in 2015 to 2016 clustered separately from the ones detected pre-2013, and their RdRp consensus sequences presented 4 aa substitutions compared with the other GII.P16 strains detected pre-2013 (Table S3). Three of the aa substitutions were located on the surface of the RdRp structure (Fig. 4d). This suggests that modification in the RdRp may play some role in the predominance of GII.P16-GII.2 strains.

Divergence history of VP1 and RdRp. The time of the most recent common ancestor (tMRCA) was determined and summarized in the time-measured maximum clade credibility (MCC) trees for both VP1-encoding and RdRp-encoding sequences (Fig. S6). The times of divergence of the phylogenetic clusters among VP1 and RdRp MCC trees were comparable, indicating the coevolution of both regions. The tMRCA of the recently reemerged GII.P16-GII.2 strains, based on partial VP1 and partial RdRp sequences of GII.P16, dated back to 2012 to 2013. Both MCC trees clearly showed that these strains evolved from the GII.P16-GII.2 strains reported in 2011 to 2012.

DISCUSSION

For over two decades, GII.4 noroviruses have been the predominant genotype circulating in humans. This dominance in human population has been attributed to the continuous replacement of antigenically distinct intragenotype variants, which can escape from herd immunity (13, 14). In contrast, none of the other genotypes have been shown to possess this advantage, as their VP1 aa sequences remain almost invariable (or static) after decades of circulation in the human population (14). During the 2014–2015 season, the GII.4 genotype was unexpectedly replaced by the GII.17 genotype in some Asian countries (15, 17, 33–35). The predominance of GII.17 strains did not last long, as during the winter of 2016 to 2017, the GII.2 genotype was shown to predominate in Germany, China, Japan, Taiwan, and Hong Kong (20–23, 36–38).

During the last 2 years, extensive research has been done to understand the driving forces that led GII.17 to become predominant in various Asian countries (16, 18). The major hypothesis is that acquisition of a novel polymerase could have driven rapid evolution of the VP1 protein that resulted in modifications of the capsid protein that increased its infectivity potential. The support for this hypothesis is based on the following: (i) the GII.P17 polymerase was described for the first time with the novel strains (15, 19); (ii) the VP1 protein from the novel GII.17 presents multiple differences (>10%), including aa insertions and deletions, compared with previous GII.17 viruses (15, 19); (iii) the people infected with GII.P17-GII.17 viruses were significantly older than the ones infected with GII.4 viruses (34); (iv) the novel GII.17 presented a broad pattern of binding to histoblood group antigens (susceptibility factors for human noroviruses) (39); and (v) despite being detected worldwide, the novel GII.17 strain predominated

FIG 4 Legend (Continued)

P2-P16 mix cluster in yellow. Reemerging GII.P16 strains, including GII.P16-GII.2 and GII.P16-GII.4, are indicated in blue. Each strain is represented by a circle, with different colors according to the associated capsid genotypes; GII.P2-GII.2 and GII.P16-GII.2 strains are indicated by red and green, respectively. The numbers on the ancestral nodes of the clusters represent the node support value calculated by the approximate likelihood-ratio test using PhyML. (b) The root-to-tip divergence plot of RdRp-encoding (nt) sequences of GII.P2 and GII.P16 strains. The x axis indicates isolation year, and the y axis shows the root-to-tip divergence on the maximum likelihood phylogenetic tree. The black lines indicate a linear regression line of the root-to-tip divergence and isolation year. Each strain is colored according to its capsid genotype. The genomic regions used spanned nt 4385 to 5104 relative to Snow Mountain virus (GenBank accession number AY134748). (c) The kernel density plot indicates the posterior estimates of substitution rate (substitutions/site/year) among GII.P2, GII.P16, and GII.P16 without reemerging strains. (d) Structural mapping of the amino acid (aa) substitutions in the reemerging GII.P16-GII.2 strains. Conservative mutations in the RdRp aa sequence from the reemerging GII.P16 strains compared to pre-2016 GII.P16 strains (P16 cluster and P2-P16 mix cluster) are mapped on the structural model of a GII.P4 RdRp (PDB number 4QPX). The molecular model was visualized using Chimera v.1.11. aa substitutions are indicated in blue. Residue 291, a residue that was shown by Bull et al. (46) to alter the incorporation rate of GII.4 polymerases, is indicated in red.

only in Asian countries, suggesting race-related susceptibility. Of note, despite all these observations, the GII.17 strains present overall a very low number of aa changes in their VP1 (14), and these "novel GII.17 strains" have been found to have been circulating as far back as the 1970s (40).

Although a large number of complete VP1 sequences are available in the public databases for GII.17, most of them (136/143) correspond to the novel GII.P17-GII.17 strain (14) and extensive analyses of the differences in the rate of evolution among the different variants of GII.17 were not possible (18). Since a better (time-ordered) sequence database is available for GII.2 in GenBank, we investigated the evolutionary history of this genotype to better understand the recent predominance of the GII.P16-GII.2 strains. Our analyses show that the VP1 nt sequences of GII.2 viruses have been evolving linearly for decades, regardless of the polymerase genotypes associated with that VP1, and with similar substitution rates among phylogenetic clusters. GII.P2-GII.2 strains have been shown to be circulating at high frequencies in different areas of Japan since 2004 (27), and GII.P16-GII.2 strains have been detected since 2008 (30). Comparisons of the substitution rates in each of the different strains have shown very little variation in the VP1 evolution of the GII.2 strains.

The VP1 protein from GII.2 strains has been stable at the aa level for more than 40 years. Although few aa substitutions were noted in the VP1 sequences, and some of them mapped in locations that would resemble the antigenic sites described for GII.4 viruses, the VP1 aa sequences of the reemerging GII.P16-GII.2 strain were almost (>99.5%) identical with those of previously detected GII.P2-GII.2 strains. In addition, we found no evidence that episodic diversifications (estimated throughout the internal branches of the VP1 nt tree) acted positively on the reemergence of the GII.P16-GII.2 viruses. Using virus-like particles (VLPs) from time-ordered GII.2 variants collected from 1975 to 2010, Swanstrom et al. have shown that the cross-blockade titers remained the same (41). This suggests that the aa differences found among different GII.2 strains have minimal impact on the antigenicity of GII.2 viruses. Taken together, these data suggest that neither immune pressure nor the presence of different polymerase genotypes altered the evolutionary pattern of the VP1.

Using large-scale genomics, it has been recently shown that non-GII.4 noroviruses present minor variations in their VP1 over decades (static genotypes), while GII.4 noroviruses evolve by continuous replacement of antigenically distinct variants, thus infecting a larger population (14). A concordance between the divergence times seen using VP1 or RdRp data indicates that the GII.P16-GII.2 strain, which is predominating during the 2016–2017 season, is not a novel recombinant and that this strain had evolved following a linear root-to-tip regression line from the GII.P16-GII.2 strains circulating back in 2011 to 2012. Thus, if the GII.P16-GII.2 strains were circulating years before their predominance, and if no differences were found in the evolutionary patterns of VP1 and its antigenicity, then what are the forces behind their sudden predominance over the evolving GII.4 genotype in certain countries?

Our analyses of the RdRp region suggest that GII.P2 strains evolve at a higher substitution rate at the RdRp-encoding region than any GII.P16 strains; however, sequence and epidemiological data suggest that GII.P16-GII.2 strains replaced GII.P2-GII.2 strains. This seems contradictory; however, the predominance of GII.P16-GII.2 viruses in certain countries (20, 21) coincided with the detection of a novel GII.4 strain, GII.P16-GII.4_Sydney (20, 42–44), suggesting that this GII.P16 polymerase could have a positive impact on the fitness of the virus. Indeed, using the murine norovirus model, Arias et al. have shown that the fidelity of the polymerase could influence the transmissibility of the virus (45). Analysis of the partial RdRp sequences indicated that GII.P2 strains are very stable in their aa sequence; however, reemerging GII.P16 strains presented 4 conservative aa substitutions compared with the pre-2016–2017 strains. These mutations in the RdRp region were located on the surface of the polymerase and could have altered the polymerase kinetics or fidelity of the reemerging GII.P16 strains. Although this is a very low number of aa differences, Bull et al. have shown that single-point mutations can affect the biological properties of the RdRp from different

GII.4 strains (46). Interestingly, one of the four conservative aa substitutions, S293T, is very close to residue 291, which was shown by Bull et al. (46) to alter the kinetic activity of GII.4 polymerases. Thus, further studies on the phenotype of the different mutations within the RdRp are needed for better understanding of polymerase kinetics and/or fidelity and of their link to transmissibility and pandemic impact.

A phylogenetic clustering of norovirus genotypes (immunotypes) has been recently used to explain the pattern of norovirus reinfection recorded in children and adults (14). Of note, GII.2, GII.4, and GII.17 all classify into different immunotypes. Thus, an additional possible explanation for the reemergence of GII.P16-GII.2 is that GII.4 might have exhausted all possible VP1 variants and that the herd immunity incidentally selected genotypes from other immunotypes to prevail in the human community. Most of the GII.P16-GII.2 outbreaks have been reported to occur in childcare facilities or elementary schools in Germany (39 of 69 norovirus-associated outbreaks), Japan (30 of 74 reported outbreaks), and Taiwan (67% of 108 cluster cases) (20, 21, 36). Thus, "GII.2-naïve" populations mostly seem to be affected by the reemerging GII.P16-GII.2 strains without changes in the VP1. As the non-GII.4 are static genotypes and are unable to evolve antigenically, a non-GII.4 genotype would prevail for only a short period, possibly in limited areas and populations, before shifting to another genotype. This pattern is largely represented by the norovirus epidemiological landscape seen in the last 2 to 3 seasons.

Our results need careful interpretation because bioinformatics analysis reflects the use of different models to explain the evolutionary process, which could not be evidenced by real phenotypic differences. *In vitro* assays such as neutralization assays and replication assays are required to determine the phenotypic differences in both the antigenicity and polymerase kinetics of the reemerging GII.P16-GII.2 viruses. In addition, only a partial sequence of RdRp region was analyzed whereas other nonstructural proteins and VP2 may also play a role in viral replication and capsid assembly of norovirus (47, 48). Comprehensive analysis using full-length sequences may shed light on the evolutionary dynamics of the emerging/reemerging norovirus.

Our analyses indicated that the evolution of GII.2 VP1 was not influenced by acquisition of a different RdRp. The GII.P16-GII.2 strains have been circulating in the human population since (at least) 2008, and the current GII.P16-GII.2 strains did not show any differences in VP1 aa sequences (and thus antigenicity) that could explain their sudden predominance. Only a slight difference was found in RdRp aa sequences between reemerging and pre-2016 GII.P16 strains that could have led to their predominance. Continuing surveillance alongside complete genome information could help in the understanding of norovirus evolutionary mechanisms and of genotype replacement and, ultimately, in the control and development of effective vaccines against norovirus disease.

MATERIALS AND METHODS

Dataset. Sequence data of GII.2 strains with a nearly complete ORF2 (nucleotide [nt] position 5085 to 6419 relative to Snow Mountain virus [GenBank accession number AY134748]) region that were available in the GenBank database were retrieved and aligned (n = 151 as of 8 February 2017; see Table S1 in the supplemental material). The data set included strains detected from 1976 to 2016 and the reemerging GII.P16-GII.2 strains detected during the 2016-2017 season (20, 37). The capsid and polymerase genotypes (when available) were confirmed using the online-based norovirus Typing Tool (8). In addition, we generated another data set (n = 134) using complete ORF2 sequences that did not include the reemerging GII.P16-GII.2 strains (nt positions 5085 to 6713; Table S1). All sequences were subjected to multiple alignment using ClustalW as implemented in MEGA7 (49).

Phylogenetic analysis. A ML phylogenetic tree of ORF2 sequences was reconstructed using the PhyML online tool (http://www.atgc-montpellier.fr/phyml/). The best substitution model (general time-reversible model with rate variation among sites and a proportion of invariable sites [GTR+G+I]) was selected according to the corrected Akaike information criterion (AICc) implemented in MEGA7. Node support was evaluated by the approximate likelihood-ratio test (50). The reconstructed ML tree was visualized using R v3.3.2 and the ape package (51, 52).

The aa sequences were translated from nt sequences using MEGA7. The ML trees of aa sequences were also estimated by using THE PhyML online tool.

Time-measured phylogenetic analysis. To observe the clock-like nature of the evolution of VP1 (encoded by ORF2), the root-to-tip divergence was calculated using the inferred ML tree and isolation

years of each strain using TempEst v1.5 (53). The best-fitting root option was used to obtain the best correlation of the root-to-tip divergence. Calculated root-to-tip divergence was plotted using R v3.3.2 and the ggplot2 package. To reconstruct the evolutionary history of GII.2 VP1-encoding sequences, time-measured phylogenetic analysis was performed using the Bayesian Markov Chain Monte Carlo (MCMC) framework as implemented in BEAST v1.8.3 (54). The clock models (strict or relaxed lognormal clock) and tree priors (constant population size, exponential growth, or skyline) were tested, and the best models were selected on the basis of the model selection procedure using path-sampling/stepping stone-sampling marginal-likelihood estimation (55). The models used are summarized in Table S4. The MCMC runs were performed until all the parameters reached convergence (effective sample size, >200). MCMC runs were analyzed using Tracer v1.6 (http://tree.bio.ed.ac.uk/software/tracer/). The initial 10% of the logs from the MCMC run were removed before summarizing the posterior values. The density of the posterior values of the substitution rates was plotted using R and ggplot2. The MCC tree was reconstructed using the posterior set of trees and TreeAnnotator v1.8.3 and was visualized using R v3.3.2.

Accumulation of aa substitutions over time. To visualize the accumulation of aa substitutions within GII.2 VP1 over time, the pairwise aa differences and the timespan of isolation were calculated as indicated elsewhere (14). Heat map plots were calculated for all GII.2 VP1 sequences using GraphPad Prism version 7 (GraphPad Software, Inc., La Jolla, CA, USA), with the values from each cell representing the number of strains compared. The aa substitutions in the P domain were mapped on the structural model of the GII.2 Snow Mountain virus (Protein Data Bank [PDB] number 4RPB [56]) using UCSF Chimera v 1.11 (57).

Selection analysis. Diversifying selection of the VP1-encoding sequence through its evolutionary history was analyzed by using the iFEL model and MEME methods (31, 32). We aimed to detect codon sites under positive selection (i.e., more nonsynonymous substitutions than synonymous substitutions) during the evolution. We focused on the internal branches under positive selection, especially the branches diversifying into the clusters by the polymerase genotypes, to assess the influence of recombination on the adaptive evolution of VP1. Significant positive selection was indicated by P values of <0.05 in both methods. The evidence of positive selection on the branches was indicated by EBF values of >10 in MEME.

Phylogenetic analysis of RdRp region. Partial RdRp sequences of GII.P2 and GII.P16 (nt positions 4385 to 5104, relative to Snow Mountain virus) were retrieved from GenBank, with other capsid genotypes added to make a total of 72 GII.P2-GII.2 sequences (detected from 1994 to 2013) and 131 GII.P16 sequences (including GII.P16-GII.2, GII.P16-GII.3, GII.P16-GII.4, GII.P16-GII.10, GII.P16-GII.13, GII.P16-GII.16, and GII.P16-GII.17 and those detected from 1975 to 2016; Table S1). Data corresponding to the ML phylogenetic tree, root-to-tip divergence, and time-measured phylogenetic tree of each polymerase genotype were estimated as indicated above for the VP1-encoding sequences.

The aa substitutions on the RdRp of reemerging GII.P16 strains were mapped on the structural model of the GII.4 virus (PDB number 4QPX [58]) using UCSF Chimera v 1.11 (57).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSphereDirect.00187-17.

FIG S1, TIF file, 1 MB.
FIG S2, TIF file, 1.9 MB.
FIG S3, TIF file, 0.3 MB.
FIG S4, TIF file, 1.1 MB.
FIG S5, TIF file, 1.5 MB.
FIG S6, TIF file, 2.3 MB.
TABLE S1, PDF file, 0.04 MB.
TABLE S2, PDF file, 0.03 MB.
TABLE S3, PDF file, 0.03 MB.
TABLE S4, PDF file, 0.03 MB.

REFERENCES

1. Glass RI, Parashar UD, Estes MK. 2009. Norovirus gastroenteritis. N Engl J Med 361:1776–1785. https://doi.org/10.1056/NEJMra0804575.

2. Lopman B, Gastañaduy P, Park GW, Hall AJ, Parashar UD, Vinjé J. 2012.

Environmental transmission of norovirus gastroenteritis. Curr Opin Virol 2:96–102. https://doi.org/10.1016/j.coviro.2011.11.005.

3. Ahmed SM, Lopman BA, Levy K. 2013. A systematic review and meta-analysis of the global seasonality of norovirus. PLoS One 8:e75922. https://doi.org/10.1371/journal.pone.0075922.

4. Mounts AW, Ando T, Koopmans M, Bresee JS, Noel J, Glass RI. 2000. Cold weather seasonality of gastroenteritis associated with Norwalk-like viruses. J Infect Dis 181(Suppl 2):S284–S287. https://doi.org/10.1086/315586.

5. Patel MM, Widdowson MA, Glass RI, Akazawa K, Vinjé J, Parashar UD. 2008. Systematic literature review of role of noroviruses in sporadic gastroenteritis. Emerg Infect Dis 14:1224–1231. https://doi.org/10.3201/eid1408.071114.

6. Pires SM, Fischer-Walker CL, Lanata CF, Devleesschauwer B, Hall AJ, Kirk MD, Duarte AS, Black RE, Angulo FJ. 2015. Aetiology-specific estimates of the global and regional incidence and mortality of diarrhoeal diseases commonly transmitted through food. PLoS One 10:e0142927. https://doi.org/10.1371/journal.pone.0142927.

7. Vinjé J. 2015. Advances in laboratory methods for detection and typing of norovirus. J Clin Microbiol 53:373–381. https://doi.org/10.1128/JCM.01535-14.

8. Kroneman A, Vennema H, Deforche K, van der Avoort H, Peñaranda S, Oberste MS, Vinjé J, Koopmans M. 2011. An automated genotyping tool for enteroviruses and noroviruses. J Clin Virol 51:121–125. https://doi.org/10.1016/j.jcv.2011.03.006.

9. Bull RA, Tanaka MM, White PA. 2007. Norovirus recombination. J Gen Virol 88:3347–3359. https://doi.org/10.1099/vir.0.83321-0.

10. Siebenga JJ, Vennema H, Zheng DP, Vinje J, Lee BE, Pang XL, Ho EC, Lim W, Choudekar A, Broor S, Halperin T, Rasool NB, Hewitt J, Greening GE, Jin M, Duan ZJ, Lucero Y, O'Ryan M, Hoehne M, Schreier E, Ratcliff RM, White PA, Iritani N, Reuter G, Koopmans M. 2009. Norovirus illness is a global problem: emergence and spread of norovirus GII 4 variants, 2001-2007. J Infect Dis 200:802–812.

11. Vinjé J, Altena SA, Koopmans MP. 1997. The incidence and genetic variability of small round-structured viruses in outbreaks of gastroenteritis in The Netherlands. J Infect Dis 176:1374–1378. https://doi.org/10.1086/517325.

12. Noel JS, Fankhauser RL, Ando T, Monroe SS, Glass RI. 1999. Identification of a distinct common strain of "Norwalk-like viruses" having a global distribution. J Infect Dis 179:1334–1344. https://doi.org/10.1086/314783.

13. Lindesmith LC, Beltramello M, Donaldson EF, Corti D, Swanstrom J, Debbink K, Lanzavecchia A, Baric RS. 2012. Immunogenetic mechanisms driving norovirus GII.4 antigenic variation. PLoS Pathog 8:e1002705. https://doi.org/10.1371/journal.ppat.1002705.

14. Parra GI, Squires RB, Karangwa CK, Johnson JA, Lepore CJ, Sosnovtsev SV, Green KY. 2017. Static and evolving norovirus genotypes: implications for epidemiology and immunity. PLoS Pathog 13:e1006136. https://doi.org/10.1371/journal.ppat.1006136.

15. Matsushima Y, Ishikawa M, Shimizu T, Komane A, Kasuo S, Shinohara M, Nagasawa K, Kimura H, Ryo A, Okabe N, Haga K, Doan YH, Katayama K, Shimizu H. 2015. Genetic analyses of GII.17 norovirus strains in diarrheal disease outbreaks from December 2014 to March 2015 in Japan reveal a novel polymerase sequence and amino acid substitutions in the capsid region. Euro Surveill 20. https://doi.org/10.2807/1560-7917.ES2015.20.26.21173.

16. Chan MC, Lee N, Hung TN, Kwok K, Cheung K, Tin EK, Lai RW, Nelson EA, Leung TF, Chan PK. 2015. Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. Nat Commun 6:10061. https://doi.org/10.1038/ncomms10061.

17. Lu J, Sun L, Fang L, Yang F, Mo Y, Lao J, Zheng H, Tan X, Lin H, Rutherford S, Guo L, Ke C, Hui L. 2015. Gastroenteritis outbreaks caused by norovirus GII.17, Guangdong Province, China, 2014–2015. Emerg Infect Dis 21:1240–1242. https://doi.org/10.3201/eid2107.150226.

18. Lu J, Fang L, Zheng H, Lao J, Yang F, Sun L, Xiao J, Lin J, Song T, Ni T, Raghwani J, Ke C, Faria NR, Bowden TA, Pybus OG, Li H. 2016. The evolution and transmission of epidemic GII.17 noroviruses. J Infect Dis 214:556–564. https://doi.org/10.1093/infdis/jiw208.

19. Parra GI, Green KY. 2015. Genome of emerging norovirus GII.17, United States, 2014. Emerg Infect Dis 21:1477–1479. https://doi.org/10.3201/eid2108.150652.

20. Niendorf S, Jacobsen S, Faber M, Eis-Hübinger AM, Hofmann J, Zimmermann O, Höhne M, Bock CT. 2017. Steep rise in norovirus cases and emergence of a new recombinant strain GII.P16-GII.2, Germany, winter 2016. Euro Surveill 22. https://doi.org/10.2807/1560-7917.ES.2017.22.4.30447.

21. Infectious Disease Surveillance Center. 2017. Norovirus trends in Japan, 2015/16 season. National Institute of Infectious Diseases, Japan.

22. Ao Y, Wang J, Ling H, He Y, Dong X, Wang X, Peng J, Zhang H, Jin M, Duan Z. 15 July 2017. Norovirus GII.P16/GII.2-associated gastroenteritis, China. 2016. Emerg Infect Dis https://doi.org/10.3201/eid2307.

23. Lu J, Fang L, Sun L, Zeng H, Li Y, Zheng H, Wu S, Yang F, Song T, Lin J, Ke C, Zhang Y, Vinjé J, Li H. 2017. Association of GII.P16-GII.2 recombinant norovirus strain with increased norovirus outbreaks, Guangdong, China, 2016. Emerg Infect Dis 23. https://doi.org/10.3201/eid2307.170333.

24. Hoa Tran TN, Trainor E, Nakagomi T, Cunliffe NA, Nakagomi O. 2013. Molecular epidemiology of noroviruses associated with acute sporadic gastroenteritis in children: global distribution of genogroups, genotypes and GII.4 variants. J Clin Virol 56:185–193. https://doi.org/10.1016/j.jcv.2012.11.011.

25. Fankhauser RL, Noel JS, Monroe SS, Ando T, Glass RI. 1998. Molecular epidemiology of "Norwalk-like viruses" in outbreaks of gastroenteritis in the United States. J Infect Dis 178:1571–1578. https://doi.org/10.1086/314525.

26. Morens DM, Zweighaft RM, Vernon TM, Gary GW, Eslien JJ, Wood BT, Holman RC, Dolin R. 1979. A waterborne outbreak of gastroenteritis with secondary person-to-person spread. Association with a viral agent. Lancet i:964–966.

27. Iritani N, Kaida A, Kubo H, Abe N, Murakami T, Vennema H, Koopmans M, Takeda N, Ogura H, Seto Y. 2008. Epidemic of genotype GII.2 noroviruses during spring 2004 in Osaka City, Japan. J Clin Microbiol 46:2406–2409. https://doi.org/10.1128/JCM.01993-07.

28. Iritani N, Vennema H, Siebenga JJ, Siezen RJ, Renckens B, Seto Y, Kaida A, Koopmans M. 2008. Genetic analysis of the capsid gene of genotype GII.2 noroviruses. J Virol 82:7336–7345. https://doi.org/10.1128/JVI.02371-07.

29. Sakon N, Yamazaki K, Nakata K, Kanbayashi D, Yoda T, Mantani M, Kase T, Takahashi K, Komano J. 2015. Impact of genotype-specific herd immunity on the circulatory dynamism of norovirus: a 10-year longitudinal study of viral acute gastroenteritis. J Infect Dis 211:879–888. https://doi.org/10.1093/infdis/jiu496.

30. Iritani N, Kaida A, Abe N, Sekiguchi J, Kubo H, Takakura K, Goto K, Ogura H, Seto Y. 2012. Increase of GII.2 norovirus infections during the 2009–2010 season in Osaka City, Japan. J Med Virol 84:517–525. https://doi.org/10.1002/jmv.23211.

31. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8:e1002764. https://doi.org/10.1371/journal.pgen.1002764.

32. Pond SL, Frost SD, Grossman Z, Gravenor MB, Richman DD, Brown AJ. 2006. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput Biol 2:e62. https://doi.org/10.1371/journal.pcbi.0020062.

33. Pan L, Xue C, Fu H, Liu D, Zhu L, Cui C, Zhu W, Fu Y, Qiao S. 2016. The novel norovirus genotype GII.17 is the predominant strain in diarrheal patients in Shanghai, China. Gut Pathog 8:49. https://doi.org/10.1186/s13099-016-0131-3.

34. Chen H, Qian F, Xu J, Chan M, Shen Z, Zai S, Shan M, Cai J, Zhang W, He J, Liu Y, Zhang J, Yuan Z, Zhu Z, Hu Y. 2015. A novel norovirus GII.17 lineage contributed to adult gastroenteritis in Shanghai, China, during the winter of 2014–2015. Emerg Microbes Infect 4:e67. https://doi.org/10.1038/emi.2015.67.

35. Gao Z, Liu B, Huo D, Yan H, Jia L, Du Y, Qian H, Yang Y, Wang X, Li J, Wang Q. 2015. Increased norovirus activity was associated with a novel norovirus GII.17 variant in Beijing, China during winter 2014–2015. BMC Infect Dis 15:574. https://doi.org/10.1186/s12879-015-1315-z.

36. I-chia L. 2017. Diarrhea cases reach record high. Taipei Times. http://www.taipeitimes.com/News/taiwan/archives/2017/03/15/2003666806.

37. Chan MCW, Kwok K, Hung T-N, Chan L-Y, Chan PKS. 2017. Complete genome sequence of an emergent recombinant GII.P16-GII.2 norovirus strain associated with an epidemic spread in the winter of 2016-2017 in Hong Kong, China. Genome Announc 1:e00343-17. https://doi.org/10.1128/genomeA.00343-17.

38. Thongprachum A, Okitsu S, Khamrin P, Maneekarn N, Hayakawa S, Ushijima H. 2017. Emergence of norovirus GII.2 and its novel recombination during the gastroenteritis outbreak in Japanese children in mid-

2016. Infect Genet Evol 51:86–88. https://doi.org/10.1016/j.meegid.2017.03.020.

39. Zhang XF, Huang Q, Long Y, Jiang X, Zhang T, Tan M, Zhang QL, Huang ZY, Li YH, Ding YQ, Hu GF, Tang S, Dai YC. 2015. An outbreak caused by GII.17 norovirus with a wide spectrum of HBGA-associated susceptibility. Sci Rep 5:17687. https://doi.org/10.1038/srep17687.

40. Mori K, Motomura K, Somura Y, Kimoto K, Akiba T, Sadamasu K. 14 March 2017. Comparison of genetic characteristics in the evolution of norovirus GII.4 and GII. 17. J Med Virol https://doi.org/10.1002/jmv.24791.

41. Swanstrom J, Lindesmith LC, Donaldson EF, Yount B, Baric RS. 2014. Characterization of blockade antibody responses in GII.2.1976 Snow Mountain virus-infected subjects. J Virol 88:829–837. https://doi.org/10.1128/JVI.02793-13.

42. Matsushima Y, Shimizu T, Ishikawa M, Komane A, Okabe N, Ryo A, Kimura H, Katayama K, Shimizu H. 2016. Complete genome sequence of a recombinant GII.P16-GII.4 norovirus detected in Kawasaki City, Japan, in 2016. Genome Announc 4:e01099-16. https://doi.org/10.1128/genomeA.01099-16.

43. Choi YS, Koo ES, Kim MS, Choi JD, Shin Y, Jeong YS. 24 January 017. Re-emergence of a GII.4 norovirus Sydney 2012 variant equipped with GII.P16 RdRp and its predominance over novel variants of GII.17 in South Korea in 2016. Food Environ Virol https://doi.org/10.1007/s12560-017-9278-4.

44. Centers for Disease Control and Prevention. 2017. CaliciNet data. https://www.cdc.gov/norovirus/reporting/calicinet/data.html. Accessed 10 March 2017.

45. Arias A, Thorne L, Ghurburrun E, Bailey D, Goodfellow I. 2016. Norovirus polymerase fidelity contributes to viral transmission in vivo. mSphere 1:e00279-16. https://doi.org/10.1128/mSphere.00279-16.

46. Bull RA, Eden JS, Rawlinson WD, White PA. 2010. Rapid evolution of pandemic noroviruses of the GII.4 lineage. PLoS Pathog 6:e1000831. https://doi.org/10.1371/journal.ppat.1000831.

47. Hardy ME. 2005. Norovirus protein structure and function. FEMS Microbiol Lett 253:1–8. https://doi.org/10.1016/j.femsle.2005.08.031.

48. Vongpunsawad S, Venkataram Prasad BV, Estes MK. 2013. Norwalk virus minor capsid protein VP2 associates within the VP1 shell domain. J Virol 87:4818–4825. https://doi.org/10.1128/JVI.03508-12.

49. Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.

50. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010.

51. R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

52. Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290. https://doi.org/10.1093/bioinformatics/btg412.

53. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly path-O-Gen). Virus Evol 2:vew007. https://doi.org/10.1093/ve/vew007.

54. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973. https://doi.org/10.1093/molbev/mss075.

55. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol Biol Evol 29:2157–2167. https://doi.org/10.1093/molbev/mss084.

56. Singh BK, Leuthold MM, Hansman GS. 2016. Structural constraints on human norovirus binding to histo-blood group antigens. mSphere 1:e00049-16. https://doi.org/10.1128/mSphere.00049-16.

57. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.

58. Zamyatkin D, Rao C, Hoffarth E, Jurca G, Rho H, Parra F, Grochulski P, Ng KK. 2014. Structure of a backtracked state reveals conformational changes similar to the state following nucleotide incorporation in human norovirus polymerase. Acta Crystallogr D Biol Crystallogr 70:3099–3109. https://doi.org/10.1107/S1399004714021518.