JOURNAL OF CLINICAL MICROBIOLOGY, Oct. 2005, p. 5179–5186 Vol. 43, No. 10
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.10.5179–5186.2005

Characterization of New Recombinant Noroviruses

K. Ambert-Balay,[1,2]* F. Bon,[1,2] F. Le Guyader,[3] P. Pothier,[1,2] and E. Kohli[1,2]

Laboratoire de Virologie, Reference Laboratory for Enteric Viruses, Centre Hospitalier Universitaire, 21079 Dijon,[1] Microbiologie Médicale et Moléculaire, UFRs Médecine et Pharmacie, 21079 Dijon,[2] and Laboratoire de Microbiologie, IFREMER, BP 21105, 44311 Nantes Cedex 03,[3] France

Received 16 December 2004/Returned for modification 31 January 2005/Accepted 20 June 2005

Noroviruses are important etiologic agents of acute gastroenteritis and show great genetic diversity. To characterize more fully previously detected strains that could not be assigned unequivocally to one particular genotype based on the RNA polymerase, we have sequenced a region in the capsid gene and, in some cases, in the junction between open reading frame 1 (ORF1) and ORF2. The results allowed us to identify several recombinant noroviruses: GGIIb viruses were detected for the first time in France in August 2000 and then spread through France and to Europe during the following winter. Here we present the characterization of three other probable GII recombinants which showed different phylogenetic positions depending on their ORF1 and ORF2 sequences. Analysis of the region located between ORF1 and ORF2 by a nucleotide identity window search showed a sudden shift in similarities. Moreover, recombination breakpoints were identified upstream and downstream of the beginning of ORF2 by using a statistical test, thus confirming the involvement of this region in recombination. Unlike GGIIb, the three recombinants described here do not seem to have diffused widely in the community: one was found in a waterborne outbreak, and the other two were found in sporadic cases. Recombination is important for the evolution of RNA viruses and has already been described for noroviruses. Our results suggest that recombination is not a rare phenomenon among noroviruses, but not all these presumed recombinants that formed during RNA replication are able to spread widely.

Recombination is an important mechanism in the evolution of RNA viruses since it can create changes in virus genomes by exchanging sequences, thus generating genetic variation and producing new viruses (43). RNA recombination has been described as a common event for some RNA viruses such as human immunodeficiency virus (18), picornaviruses (9, 34), and coronaviruses (26). For other RNA viruses, for example, astroviruses (42) and rotaviruses (37), recombination has rarely been observed.

Human calciviruses (HuCVs) have been shown to be the most common cause of viral gastroenteritis outbreaks and sporadic community cases in all age groups in industrialized countries (5–8, 31). HuCVs are single-stranded, positive-sense RNA viruses of the family *Caliciviridae*. Complete nucleotide sequencing of several caliciviruses has shown that their genome is organized in three open reading frames (ORFs) (20, 27). ORF1 encodes a large nonstructural polyprotein that is cleaved into several viral proteins, among which is an RNA-dependent RNA polymerase. ORF2 encodes the major capsid protein VP1, and ORF3 encodes a minor structural protein VP2 (13). Complete genomic sequences are available for only a few strains, but numerous partial sequencings of either part of the RNA polymerase gene or part of the capsid gene have been done independently.

HuCVs are phylogenetically divided into two genera, *Norovirus* and *Sapovirus* (13). Noroviruses are divided into five distinct genogroups, and each genogroup can be divided into several genotypes. The classification of caliciviruses in genogroups and genotypes is based on the sequence of the capsid, partial (2, 39) or complete (12). According to a classification adapted from Green et al. (12) and Ando et al. (2), genogroup I (GI), GII, and GIII are composed of seven, eight, and two genotypes, respectively. GIV is represented by one strain (41), as is genogroup V (23). This classification is in constant evolution with the discovery of new strains. Recently, Kageyama et al. proposed a classification with 31 genotypes (14 for GI and 17 for GII) (22). At the same time, Vinje et al. identified 15 genotypes in the GII (40). Also, Zheng et al. (44) defined 8 and 17 genotypes in the GI and GII, respectively.

For most strains, analysis of the ORF1 region, used by most laboratories for diagnostic reverse transcription (RT)-PCR, correlates with the clustering based on the ORF2. However, several studies have shown that strains group within different clusters, depending on whether the polymerase or capsid gene is sequenced, suggesting that recombination may have occurred. This is the case for the strains Snow Mountain (16) and Wortley and Seacroft (39). Jiang et al. (19) characterized a first naturally occurring recombinant, Arg320, by cloning and sequencing a long fragment of the genome from the RNA polymerase region to the 3′ end of the genome. This recombinant would derive from two parent strains of genogroup II, Lordsdale virus for the RNA polymerase and Mexico virus for the capsid. Katayama et al. (24) described two recombinants, WUG1 of genogroup I and Saitama U1 of genogroup II. Recently, Snow Mountain virus was confirmed to be a recombinant norovirus (28). Also, Hansman et al. (15) characterized Mc37 as a new recombinant by complete genome sequencing.

We have isolated and characterized by cloning and sequencing several strains of noroviruses that correspond to different recombinants. GGIIb variants were detected for the first time in France in August 2000 (F. Bon, P. Pothier, C. Hemery, M.

* Corresponding author. Mailing address: Laboratoire de Virologie, 2 boulevard Maréchal de Lattre de Tassigny, 21079 Dijon, France. Phone: 33 3 80 29 34 37. Fax: 33 3 80 29 36 04. E-mail: katia.balay@chu-dijon.fr.

Cournot, C. Castor, H. de Valk, P. Gourier-Frery, P. Benhamida, R. Roques, L. Le Coustumier, F. Villeneuve, P. Megraud, P. Le Cann, E. Kohli, and A. Gallay, 21st Réunion Interdisciplinaire de Chimiothérapie Anti-infectieuse, Paris, France, 2001) and then in several European countries (4, 17, 33, 38). Their RNA polymerases do not correspond to any known genotype, but their capsid genes are close to the capsid gene of viruses of three different genotypes, Hawaii, Toronto, and Snow Mountain, indicating that recombinations occurred. This paper reports on the characterization of three other strains of HuCV that are probable recombinants. The sequencing of long genomic fragments containing a portion of the RNA polymerase gene and a portion of the capsid gene shows that these two portions of the genome present homologies with distinct genotypes.

MATERIALS AND METHODS

Stool samples. From a collection of 173 calicivirus strains previously characterized by RT-PCR and nucleotide sequence analysis of the RNA polymerase gene (3, 5), we selected 55 samples that all presented a nucleotide identity in the RNA polymerase region between 77 and 88% with known genotypes; and we further characterized them by sequencing a region in the capsid gene and, for some of them, the junction between ORF1 and ORF2. The stool samples were stored at −40°C.

RNA extraction. Viral RNA was extracted from 10% stool suspensions in phosphate-buffered saline (pH 7.5) with a QIAamp viral RNA kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions, and was either reverse transcribed immediately or stored at −40°C.

RT-PCR. RT-PCR was performed to amplify sequences of the capsid gene in the 55 stool samples by using murine leukemia virus reverse transcriptase and AmpliTaq DNA polymerase purchased from Applied Biosystems (Foster City, Calif.). Both RT and PCR were carried out with the reaction mixture preparations recommended by the manufacturer. Primer set SRI1 and SRI2 (14) and primer set Mon381 and Mon383 (32) were used to amplify the sequences of the capsid gene of noroviruses genogroup I and II, respectively, giving amplicons of 316 and 322 bp, respectively. The cycling conditions were as follows: one cycle of reverse transcription at 42°C for 15 min and, for PCR, denaturation for 2 min at 94°C; 40 amplification cycles with denaturation for 30 s at 94°C, annealing for 1 min at 50°C, and extension for 1 min at 72°C; and a final cycle of incubation at 72°C for 15 min.

For some strains, another RT-PCR was performed with primers NI (11) and Mon383 to amplify a long genomic fragment that included the junction between the RNA polymerase and the capsid genes and with the Superscript II RNase H-reverse transcriptase (Invitrogen Corporation, Carlsbad, Calif.) and the Expand High Fidelity PCR system (Roche Molecular Biochemicals, Mannheim, Germany), according to the manufacturer's instructions for both reaction mixture preparations. The PCR cycle conditions used were the same as those used for amplification of the capsid gene.

Cloning and sequencing of the RT-PCR products. RT-PCR-amplified DNA fragments of the expected sizes, as determined by agarose gel electrophoresis, were excised from the gel, extracted, and purified with a Qiaex II gel extraction kit (QIAGEN). Capsid gene fragments were sequenced directly after purification.

The PCR products obtained with primer set NI and Mon383 (1,189 bp) were cloned into the pGEM-T Easy Vector System (Promega Corporation, Madison, Wis.), according to the manufacturer's instructions.

Sequencing of the PCR products or plasmids was performed with the same primers used for amplification by using the ABI Prism Big Dye Terminator cycle sequencing ready reaction kit on an automated sequencer (model 373A DNA sequencing system) (both from Applera Corporation, Foster City, Calif.). For the long fragments cloned into the plasmid, if necessary the sequence was determined by using multiple steps: after determination of the end sequences of the cDNA with the primers NI and Mon383, the remaining sequences were determined by genomic walking with primers based upon newly obtained sequences adjacent to those regions. All sequences were confirmed by sequencing at least twice.

Consensus sequences were produced by overlapping the sequences of the RNA polymerase genes and the cloned genes.

Characterization and phylogenetic analyses. Alignments of the sequences with the sequences in the GenBank library were carried out by using Fasta version 3.3t06 and GCG software, available in the national service Infobiogen (http://www.infobiogen.fr). For the RNA polymerase, alignments were also performed with reference strains available in the database of the European Food-Borne Viruses Network (25, 30).

For the phylogenetic analyses, sequences were first aligned by using the CLUSTAL-W program in the PHYLIP format. The neighbor-joining method was used to create phylogenetic trees by using the different alignments.

Detection of recombination. Detection of recombination was achieved by using similarity plots in SimPlot, as described previously (29), as well as a statistical analysis based on Sawyer's test (S. Sawyer, 1999, http://www.math.wustl.edu/sawyer). SimPlot calculates and plots the percent identity of a query sequence to a panel of reference sequences versus the position in a sliding window, which is moved across the alignment in steps. Sawyer's test looks for aligned segments for which a pair of sequences is sufficiently similar to be suggestive of gene conversion.

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been deposited in GenBank under accession numbers AY682550, AY682551, and AY682552 for samples S63, L23, and E3, respectively.

RESULTS

Amplification of the partial capsid sequences of particular strains. Among the 173 samples previously characterized by sequencing of an RNA polymerase gene fragment, we selected for this study 55 samples that were impossible to assign to one or another cluster because the percentage of nucleotide identity with reference strains was between 77 and 88%. In fact, according to Vinje et al. (39), strains belonging to the same genotype present more than 87% and 91% nucleotide identities for genogroups I and II, respectively, in this RNA polymerase fragment. Forty-one of the 55 samples appeared to belong to the new group GGIIb; seven samples presented between 77 and 81% nucleotide identity with Desert Shield virus, and seven samples presented low nucleotide identity percentages with various genotypes (Table 1).

For these 55 strains, a region of the capsid gene was amplified and sequenced. We finally obtained 47 samples for which both RNA polymerase and capsid gene sequences were available; 8 amplifications of the capsid failed. According to Green et al. (13), strains should present more than 80% amino acid identity in the capsid to be assigned to the same genotype. Thirty-three samples out of the 47 which were sequenced both in the RNA polymerase and in the capsid were of the GGIIb group. Among the 14 remaining samples, the 7 samples that presented a low nucleotide identity percentage with the Desert Shield virus in the RNA polymerase sequence could be assigned to the Desert Shield genotype according to their capsid sequences (90 to 99% amino acid identity with the Desert Shield virus). These Desert Shield-like strains are very close to a strain described by Johansson et al. (21) that presents 82% nucleotide identity in the polymerase region with the Desert Shield virus. Also, sample GEA264, which presented a low percentage of nucleotide identity with the Hawaii virus in the RNA polymerase (84%), can be assigned to the Hawaii genotype since the amino acid identity with the Hawaii virus in the capsid is 95%. For each of the six remaining strains (strains S63, L23, E3, E23, GEA471, and S55), sequence analysis of the capsid allowed us to classify the strain into a genotype that was different from the genotype presenting the highest percentage of nucleotide identity with the sample in the RNA polymerase (Table 1).

TABLE 1. Strains presenting low percentages of nucleotide sequence identity with those of reference strains for the RNA polymerase gene[a]

Name of sample	RNA polymerase				Capsid			
	No. of samples	Reference virus	% Nucleotide sequence identity with reference virus	Designation[e]	No. of samples	Reference virus[c]	% Nucleotide sequence identity with reference virus	% Amino acid sequence identity with reference virus
L169[b]					25	Toronto	91–93	93–99
E78[b]					4	Hawaii	88	97
					3	Bristol	91–93	95–96
E673[b]					1	Snow Mountain	95	98
	41	Hawaii[c]	83–86	GGIIb	8	Amplification failed		
	7	Desert Shield[d]	77–81	NA[f]	7	Desert Shield	88–91	90–99
GEA264	1	Hawaii	84	NA	1	Hawaii	84	95
S63	1	Melksham[c]	82	NA	1	Hillingdon	92	99
L23	1	Hillingdon[c]	86	NA	1	Hawaii	82	94
E3	1	Bristol[c]	88	NA	1	Melksham	87	97
E23	1	Sindlesham[d]	86	NA	1	Toronto	98	99
GEA471	1	Hawaii	85	NA	1	Bristol	83	86
S55	1	Melksham	82	NA	1	Bristol	93	95

[a] The reference virus indicated for each sample is the one among all reference viruses representing all genotypes which presented the highest percentage of identity with the sample.

[b] L169 is one of the 25 samples of GGIIb/Toronto, E78 is one of the 4 samples of GGIIb/Hawaii, and E673 is the unique sample GGIIb/Snow Mountain. Their ORF1–ORF2 junctions were sequenced (GenBank accession numbers AY773210, AY580335, and AY682549, respectively). The junction is not available for GGIIb/Bristol.

[c] GenBank accession numbers: Bristol/1993/UK, X76716; Desert Shield 395/1990/SA, U04469; Hawaii/1971/US, U07611; Hillingdon/1990/UK, AJ277607; Melksham/1994/UK, X81879; Toronto 24/1991/CA, U02030; SMA, AY134748.

[d] Sequences are not available in GenBank; the sequences were kindly provided by Marion Koopmans and Harry Vennema (RIVM, National Institute of Public Health and the Environment, Bilthoven, The Netherlands).

[e] Designation at the date of characterization.

[f] NA, not assigned.

Phylogenetic analysis of partial RNA polymerase and capsid sequences. Figure 1 represents the phylogenetic trees for six strains, strains S63, L23, E3, E23, GEA471, and S55. The RNA polymerases of strains S63, L23, and E3 are related to Melksham, Hillingdon, and Bristol viruses, respectively, although they could not be clearly assigned to these genotypes; and their capsid genes can be classified into the Hillingdon, Hawaii, and Melksham genotypes, respectively. Also, for E23, GEA471, and S55, the phylogenetic trees do not give the same classification when the trees are based on the RNA polymerase and the capsid. These results confirm that for each of these strains the RNA polymerase gene and the capsid gene do not belong to the same genotype, thus suggesting the possibility of recombination.

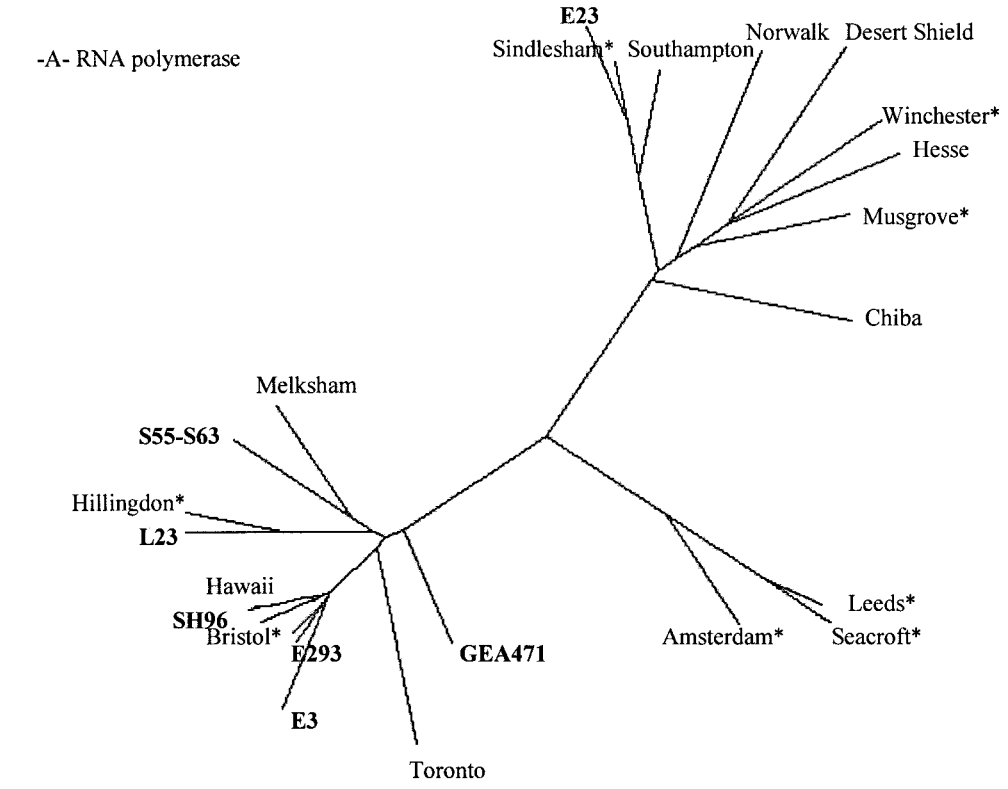
Identification of recombination by nucleotide identity window search. To confirm the existence of recombinant strains and to exclude the possibility of double contamination, long genomic fragments that included part of the RNA polymerase and part of the capsid genes were amplified by RT-PCR and cloned into a plasmid vector before sequencing. We obtained consensus sequences (by overlapping sequences of the polymerase and the cloned genes) of 1,428 bp, 1,371 bp, and 1,416 bp for S63, L23, and E3, respectively, that correspond to the regions located between nucleotides 4233 and 5660, 4290 and 5660, and 4233 and 5648 in the Hawaii virus genome (GenBank accession number U07611), respectively. The other three potentially recombinant strains, GEA471, S55, and E23, were also successfully cloned; but it appeared that they were doubly contaminated samples (data not shown).

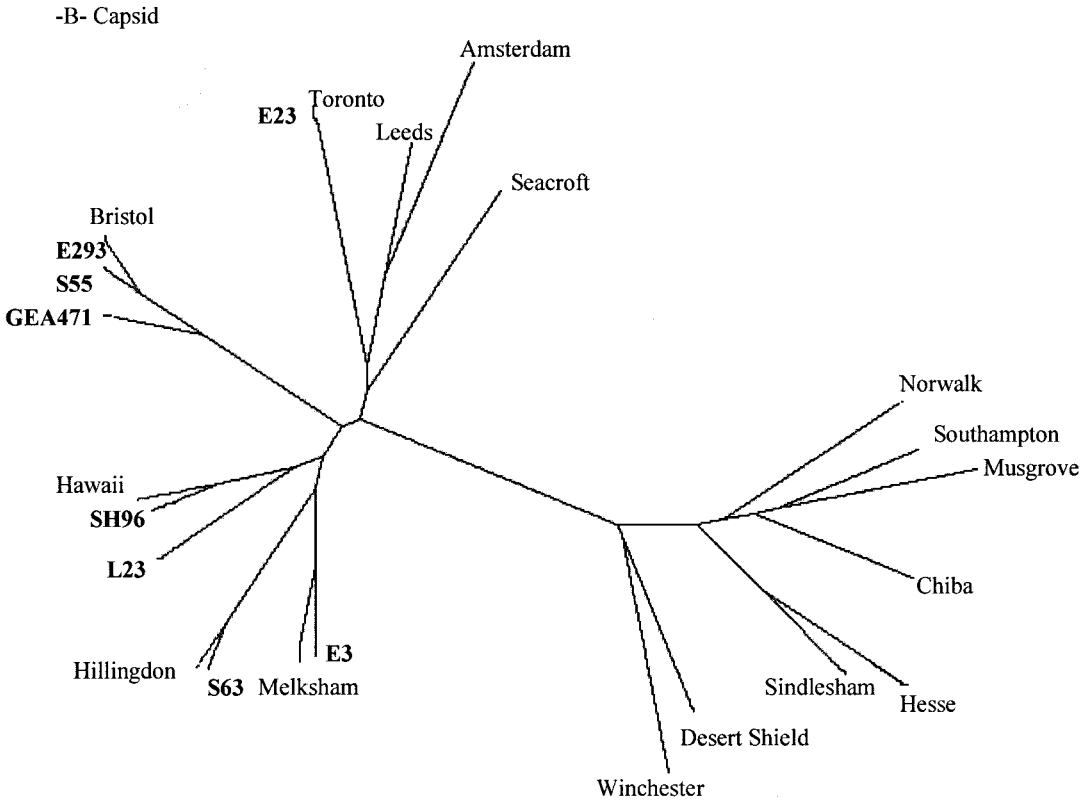
For strains S63, L23, and E3, the long sequences corresponding to ORF1-ORF2 fragments were analyzed with SimPlot software by comparing the sequences of strains representing each genotype of GII. Figure 2 shows the results obtained by comparison with the Melksham, Hawaii, Bristol, and Hillingdon sequences as the reference sequences.

S63 is distant from the four reference strains (and from the other reference strains; data not shown) in the first part of the sequence, which corresponds to ORF1 (residues 1 to 872; residues 4233 to 5104 in the Hawaii sequence): there is less than 85% similarity except at the 3′ end, which corresponds to a more conserved region of the genome, at the ORF1-ORF2 junction (24). This confirms that it could not be classified into one particular group on the basis of the RNA polymerase gene sequence. On the other hand, the S63 sequence shows a close association with Hillingdon virus starting at about position 853 (ORF2 starts at position 853 [position 5085 in Hawaii virus]); this corresponds to the result obtained with the capsid fragment.

For L23, the first part of the sequence, which corresponds to ORF1, between residues 1 and 815 (residues 4290 to 5104 in Hawaii virus) is closer to Hillingdon than to other reference strains but presents only 85% similarity, on average, which is not high enough for the strains to be considered of the same genotype; the second part corresponding to ORF2 and starting at position 796 (position 5085 in Hawaii virus) is closer to Hawaii strain essentially at the 5′ end. It must be noted that the decrease in the percentage of similarity with Hawaii virus starting at about position 900 in Fig. 2B (position 5190 in Hawaii virus) is in accordance with the rather low percentage of nucleotide identity of the capsid fragment with that of Hawaii virus reported in Table 1 (82%). However, the amino acid identity percentage is high (94%).

For E3, the sequence between nucleotides 1 and 872 (ORF1) (nucleotides 4233 to 5104 in Hawaii) is closer to the Bristol sequence, but the sequence between nucleotides 853 and 1416 (ORF2) (nucleotides 5085 to 5648 in Hawaii virus) is


-A- RNA polymerase
E23
Sindlesham*
Southampton
Norwalk
Desert Shield
Winchester*
Hesse
Musgrove*
Chiba
Melksham
S55-S63
Hillingdon*
L23
Hawaii
SH96
Bristol*
E293
GEA471
E3
Toronto
Amsterdam*
Leeds*
Seacroft*


-B- Capsid
Amsterdam
Toronto
E23
Leeds
Seacroft
Bristol
E293
S55
GEA471
Norwalk
Southampton
Musgrove
Hawaii
SH96
L23
Chiba
Hillingdon
S63
Melksham
E3
Sindlesham
Hesse
Desert Shield
Winchester

close to Melksham sequence, which is also the case for the capsid gene analyzed separately. These three examples showing sudden shifts in similarities suggest that a recombination occurred near the junction of ORF1 and ORF2, which is localized between nucleotides 5085 and 5104 in the Hawaii genome.

Identification of recombination by Sawyer's test. The available sequences of strains E3, L23, and S63, including the ORF1-ORF2 junction, were aligned separately with the following reference strains: Hawaii, Bristol, Melksham, and Hillingdon. As controls, three confirmed recombinants were also tested: Arg320 (GenBank accession number AF190817), which has been previously published as a Lordsdale-Mexico recombinant (19), and two strains of the GGIIb group which were detected in our laboratory, strains L169 (GGIIb/Toronto) and E673 (GGIIb/Snow Mountain) (Table 1). The sequences of these three additional strains located in the same region as those of E3, L23, and S63 were aligned with the sequences of the same reference strains, as well as those of Mexico, Toronto, and Snow Mountain strains for Arg320, L169, and E673, respectively. The results allowed detection of different aligned segments with sequences sufficiently similar to be suggestive of gene conversion (Table 2): S63 and Hillingdon presented a 515-bp homologous fragment which corresponds to the totality of the ORF2 sequence that we obtained for S63; L23 shows a short fragment (122 bp) homologous to that of the Hawaii genotype at the 5′ end of ORF2; E3 presents a 185-bp sequence homologous to that of the Bristol genotype at the 3′ end of ORF1 and a 330-bp sequence homologous to that of the Melksham genotype at the 5′ end of ORF2. This suggests the presence of recombination breakpoints upstream and downstream the beginning of ORF2 (Table 2; Fig. 3) located at positions 25, −45, and −6 or −9 (when it is considered that position 0 is the beginning of ORF2) for strains S63, L23, and E3, respectively. For the three confirmed recombinants, breakpoints were identified at positions 26 (strains Arg320 and L169) and −67 (strain E673).

Geographical distribution of recombinant strains. E3 was found during the investigation of a waterborne outbreak, along with other pathogens. This outbreak happened in Crete in June 1997 and concerned approximately 200 individuals. Both S63 and L23 were found in sporadic cases in the center of France in January 1999 and in the west of France in December 1999, respectively (Table 3). No data were available for the patients concerned.

To determine if these strains were found in other places in Europe, we compared the RNA polymerase sequences that we obtained with sequences entered in the database of the European Food-Borne Viruses Network. It appeared that no other sequence corresponds to the S63 RNA polymerase sequence. The L23 RNA polymerase sequence is similar to those of 17 strains in this database, and the E3 RNA polymerase sequence is close to those of 5 strains, but no capsid sequence was available for any of these 22 strains, so it is impossible to certify that these strains are identical to the strains characterized in our study.

By comparison with the sequences in GenBank, we found that the L23 sequence is very close (98 to 99% nucleotide identity) to eight sequences of strains characterized in the United States, but no data concerning their frequency or geographical distribution are available. It must be noted that the complete ORF2 sequences of these strains have been determined, but their ORF1 sequences are at least 550 bp shorter than the L23 ORF1 sequences, thus preventing comparison of the 5′ end of L23 with the sequences of the 5′ ends of these strains. The sequences of S63 and E3 did not correspond to any sequences submitted to GenBank so far.

DISCUSSION

The aim of this study was to better characterize 55 strains previously detected but not classified on the basis of the polymerase gene by sequencing the capsid gene and, if necessary, the ORF1-ORF2 junction. Among the 47 noroviruses whose capsid genes could be further sequenced, 36 were recombinant strains. Thirty-three were of the GGIIb group, which has been described elsewhere (4, 17, 33, 38), and three were identified as three different probably new naturally occurring recombinants. The possibility of a double infection of patients was excluded by cloning the cDNA covering the junction between the RNA polymerase and the capsid genes. This cDNA is generated from a single virus. Indeed, it must be mentioned that three other samples suspected of being recombinants proved to be doubly contaminated when they were cloned. These examples show the importance of working with cloned sequences or long genomic fragments covering the junction between ORF1 and ORF2 but not with separate RNA polymerase and capsid genomic fragments. The E3 strain has a Melksham capsid and a Bristol-related *pol* gene, although it could not be assigned strictly to the Bristol genotype on the basis of the criteria proposed by Vinje et al. (39). This result was confirmed by both the SimPlot test and statistical analysis. As their RNA polymerase genes do not correspond to any cluster, S63 and L23 can be defined as new variants, but they also show the characteristics of recombinant strains. S63 has a capsid gene that belongs to the Hillingdon genotype and a *pol* gene that is distant from the capsid-based genotype. The results of SimPlot and statistical analysis both confirmed recombination in the junction region between ORF1 and ORF2. Strain L23 has a capsid that belongs to the Hawaii genotype (95% amino acid identity). However, the results of SimPlot and Sawyer's test suggest that in this case recombination may not involve a large

FIG. 1. Phylogenetic analysis of partial (A) RNA polymerase (145-bp) and (B) capsid (277-bp) sequences. The strains characterized in our laboratory are represented in boldface. Strains SH96 and E293 were characterized in another study; they present a good correlation between the RNA polymerase and the capsid genes, with SH96 being assigned to the Hawaii genotype and E293 being assigned to the Bristol genotype. The GenBank accession numbers for the calicivirus reference strains are Amsterdam/1998/NL, AF195848; Chiba407/1987/JP, AB042808; Hesse3/1997/GE, AF093797; Leeds/1990/UK, AJ277608; Musgrove/1989/UK, AJ277614; Norwalk/1968/US, M87661; Seacroft/1990/UK, AJ277620; Sindlesham/1995/UK, AJ277615; and Winchester/1994/UK, AJ277609. The others are given in Table 1. *, sequences not available in GenBank (the sequences were kindly provided by Marion Koopmans and Harry Vennema).

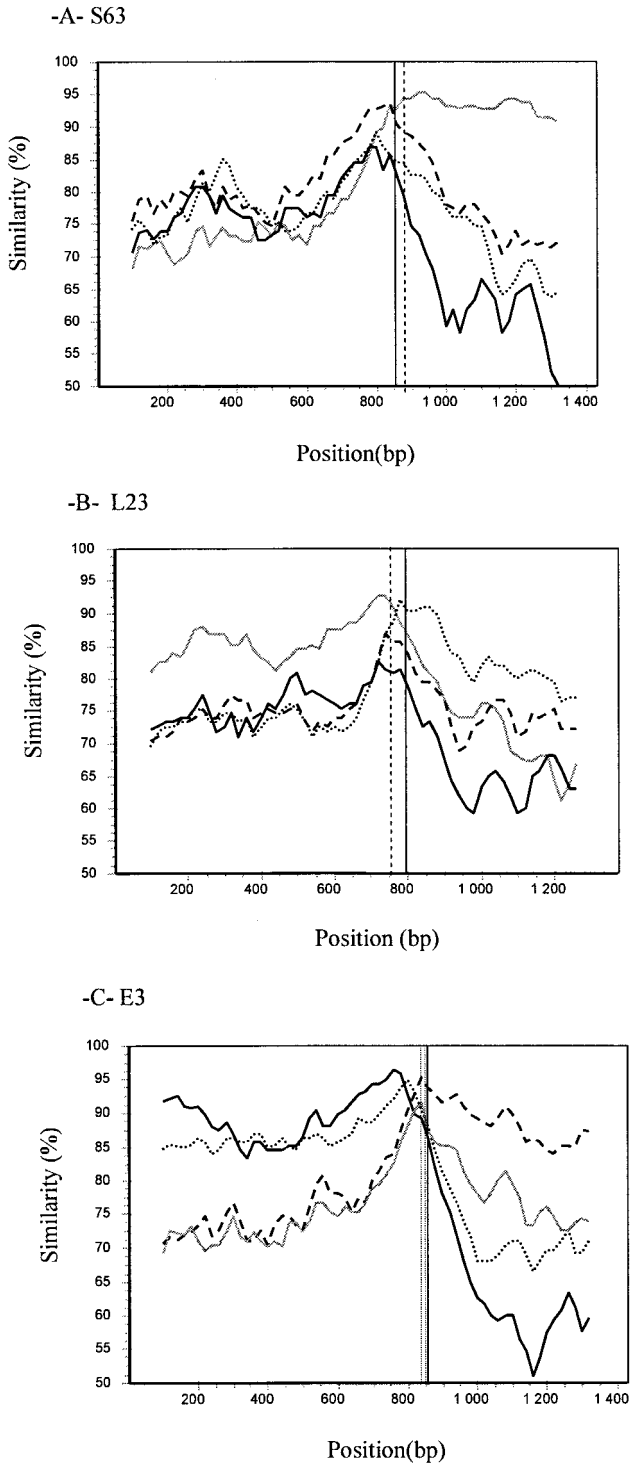


FIG. 2. SimPlot analysis of partial RNA polymerase and capsid gene sequences of recombinant strains. Window size, 200 bp; step, 20 bp. The query sequences for panels A, B, and C are S63, L23, and E3, respectively. The vertical axis indicates the nucleotide identities between the query strain and the four reference strains, expressed as percentages. The horizontal axis indicates the nucleotide positions (in base pairs; corresponding to base pairs 4233 to 5660, 4290 to 5660, and 4233 to 5648 of the Hawaii strain with GenBank accession number U07611 for S63, L23, and E3, respectively). Vertical lines indicate the beginning of ORF2 (continuous vertical line) and putative recombination breakpoints identified by Sawyer's test at the ORF1-ORF2 junction (dotted vertical line). Gray line, Hillingdon; dashed line, Melksham genotype; bold line, Bristol; dotted line, Hawaii.

TABLE 2. Homologous fragments suggestive of recombination according to Sawyer's test

Strain	Parent strain	Fragment position according to 5′ end of ORF2	Length (bp)	*P* value
S63	Hillingdon	**25** to 539	515	0.0000
L23	Hawaii	**−45** to 77	122	0.0007
E3	Bristol	−193 to **−9**	185	0.0022
		−751 to −612	140	0.0115
		−444 to −378	67	0.0237
	Melksham	**−6** to 323	330	0.0000
		453 to 547	95	0.0013
Arg320	Mexico	**26** to 576	551	0.0000
L169	Toronto	**26** to 548	523	0.0000
E673	SMA[b]	**−67** to 282	349	0.0000

[a] Positions in boldface are putative recombination breakpoints in the ORF1–ORF2 junction region. GenBank accession numbers: Hillingdon-related virus, AF397156; Mexico strain, U22498; Snow Mountain strain, AY134748.
[b] SMA, Snow Mountain agent.

fragment of the parental strain. Such recombination may be more similar to the recombination identified recently by Rohayem and Münch (36) in different norovirus capsids which show fragments of different lengths that may result from gene conversion.

Taken together, these results confirm that the junction region between ORF1 and ORF2 may be critical for recombination, as is the case for other recombinant noroviruses described in the literature. Sawyer's test allowed us to identify several recombination breakpoints located in this region, upstream and downstream of the beginning of ORF2, suggesting that there is not a unique breakpoint in this region. Of interest, positions 25 and 26 were identified as breakpoints for three different recombinant strains: S63 (Hillingdon genotype capsid, GII-5), Arg320 (Mexico genotype capsid, GII-3), and L169 (Toronto genotype capsid, GII-3), suggesting that this particular region may be more frequently involved in recombination. Although the exact mechanism of recombination has not been elucidated, it has been suggested that during coinfection, the RNA-dependent RNA polymerase could switch from one RNA molecule to another during viral replication because of the high sequence conservation that increases the interaction between RNA molecules. Our investigations of gastroenteritis outbreaks showed that coinfection of patients by several noroviruses has been observed frequently, especially in water- and oyster-related outbreaks, and has also been described several times in the literature (1, 10, 35).

The capsid of the Melksham strain that is part of the E3 strain has been found to recombine with the polymerase of the Bristol strain to give the recombinant Snow Mountain strain (28). The RNA polymerase of the Bristol strain is also found in recombinant strains Arg320 (19) and Wortley (39) in combination with the capsids of genotypes Mexico and Hawaii, respectively. Of interest, sample E3, considered a recombinant between Bristol and Melksham (*pol* and capsid, respectively), is different from the Snow Mountain agent, which is also a Bristol-Melksham recombinant (86% nucleotide sequence ho-

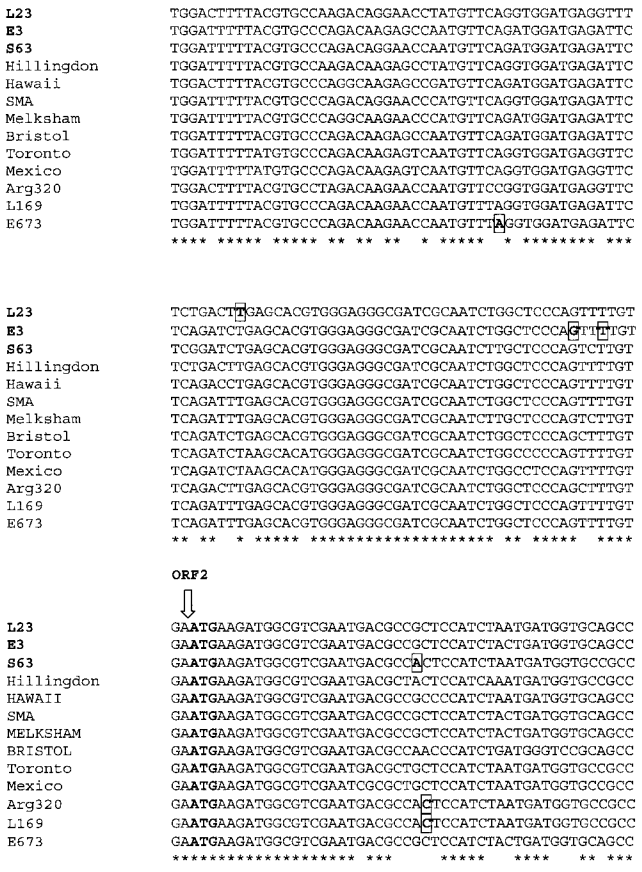


FIG. 3. Recombination breakpoints detected in the junction region between ORF1 and ORF2 of recombinant strains by Sawyer's test. Positions according to the beginning of ORF2 are −45, −6 or −9, and 25 for the three putative recombinants L23, E3, and S63, respectively, and 26, 26, and −67 for the confirmed recombinants Arg320, L169, and E673, respectively.

mology in ORF1 and 92% nucleotide sequence homology in ORF2). Thus, some genotypes may be more prone to recombination and the formation of new strains than others. However, the identification of a greater number of recombinants will be required to confirm this hypothesis.

All the recombinants described in this study, as well as the GGIIb variants, presented a low nucleotide identity percentage with known genotypes in the RNA polymerase, suggesting that the capsid sequences of such "odd" strains should be further sequenced. It is not known if the "odd" strains just mentioned really exist or if recombination between the polymerase and capsid genes of two different strains would induce mutations in the polymerase gene, producing a decrease in nucleotide similarities between the polymerases of the recombinant and parent strains. Some recombinant noroviruses described in the literature present high percentages of homology with two different genotypes in both ORFs, meaning that our observation is not a general characteristic of recombinants.

Unlike GGIIb and other norovirus recombinants, such as the Snow Mountain strain and Arg320, which are widely distributed, the three recombinants characterized in this study do not seem to have diffused widely. One was the cause of an outbreak in Crete in 1997; the other two were detected in two sporadic cases of infection in January and December 1999, respectively. Since then, they have never been identified among caliciviruses detected in France. Similar strains were found in other European countries but only their RNA polymerase genes and not their capsid genes were characterized, so they may be different. However, even if we consider, based on the RNA polymerase, that they are the same strains, it appears that they were not reported many times in other European countries. Also, L23 is similar to strains characterized in the United States, but we have no data about their geographical distribution or frequency. Of interest, it has recently been proposed (Second International Calicivirus Conference, Dijon, France, 6 to 10 November 2004) that these strains define a new genotype represented by strain Tiffin (GenBank accession number AY502010). Whatever the rate of recombination in noroviruses, it seems that some recombinants are able to diffuse in the human population and emerge as new strains, while others cause infection in only a few cases and do not spread. The reasons for this difference are not clear. All the recombinants described so far present known capsids, and since capsids are responsible for the immunogenicity, the mechanism that causes a selective advantage is not clear. Variations in antigen sites and/or receptor binding sites may be critical by causing immune escape and/or increasing the number of susceptible hosts.

In conclusion, the results presented here suggest that recombination is not a rare phenomenon among noroviruses and thus contributes to the genetic diversity of these viruses. Recombinant strains may be underestimated by the fact that characterization of noroviruses is usually based on the RNA polymerase gene sequence only, whereas it is necessary to know both the RNA polymerase and the capsid gene sequences to identify such viruses. Identification of recombination breakpoints in the region located at the junction of ORF1-ORF2 confirms the importance of this region in this phenomenon. Further work is needed to determine more precisely the exact implication of recombinant noroviruses in gastroenteritis as well as the molecular basis for the selective advantage observed for some of these strains.

TABLE 3. Epidemiology of the three recombinant strains

Characteristic	E3	S63	L23
Epidemiology	Outbreak	Sporadic case	Sporadic case
Country	Crete	France	France
Vehicle	Water	Unknown	Unknown
Setting	Hotel		
Date of onset of cases	June 1997	January 1999	December 1999
No. of cases	200	1	1

ACKNOWLEDGMENTS

We thank S. A. Sawyer (Department of Mathematics, Washington University, St. Louis, Mo.) for help with Sawyer's test.

This work was supported by the European Commission, under the Quality of Life and Management of Living Resources, 5th Framework (QLK1-CT-1999-00594).

REFERENCES

1. **Anderson, A. D., A. G. Heryford, J. P. Sarisky, C. Higgins, S. S. Monroe, R. S. Beard, C. M. Newport, J. L. Cashdollar, G. S. Fout, D. E. Robbins, S. A.**

Seys, K. J. Musgrave, C. Medus, J. Vinje, J. S. Bresee, H. M. Mainzer, and R. I. Glass. 2003. A waterborne outbreak of Norwalk-like virus among snowmobilers—Wyoming, 2001. J. Infect. Dis. **187:**303–306.

2. **Ando, T., J. S. Noel, and R. L. Fankhauser.** 2000. Genetic classification of "Norwalk-like viruses." J. Infect. Dis. **181**(Suppl. 2)**:**S336–S338.
3. **Bon, F., H. Giraudon, C. Sancey, C. Barranger, M. Joannes, P. Pothier, and E. Kohli.** 2004. Development and evaluation of a new commercial test allowing the simultaneous detection of noroviruses and sapoviruses by reverse transcription-PCR and microplate hybridization. J. Clin. Microbiol. **42:**2218–2220.
4. **Buesa, J., B. Collado, P. Lopez-Andujar, R. Abu-Mallouh, J. Rodriguez Diaz, A. Garcia Diaz, J. Prat, S. Guix, T. Llovet, G. Prats, and A. Bosch.** 2002. Molecular epidemiology of caliciviruses causing outbreaks and sporadic cases of acute gastroenteritis in Spain. J. Clin. Microbiol. **40:**2854–2859.
5. **Chikhi-Brachet, R., F. Bon, L. Toubiana, P. Pothier, J. C. Nicolas, A. Flahault, and E. Kohli.** 2002. Virus diversity in a winter epidemic of acute diarrhea in France. J. Clin. Microbiol. **40:**4266–4272.
6. **de Wit, M. A., M. P. Koopmans, L. M. Kortbeek, N. J. van Leeuwen, A. I. Bartelds, and Y. T. van Duynhoven.** 2001. Gastroenteritis in sentinel general practices, The Netherlands. Emerg. Infect. Dis. **7:**82–91.
7. **de Wit, M. A., M. P. Koopmans, L. M. Kortbeek, W. J. Wannet, J. Vinje, F. van Leusden, A. I. Bartelds, and Y. T. van Duynhoven.** 2001. Sensor, a population-based cohort study on gastroenteritis in The Netherlands: incidence and etiology. Am. J. Epidemiol. **154:**666–674.
8. **Fankhauser, R. L., S. S. Monroe, J. S. Noel, C. D. Humphrey, J. S. Bresee, U. D. Parashar, T. Ando, and R. I. Glass.** 2002. Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States. J. Infect. Dis. **186:**1–7.
9. **Furione, M., S. Guillot, D. Otelea, J. Balanant, A. Candrea, and R. Crainic.** 1993. Polioviruses with natural recombinant genomes isolated from vaccine-associated paralytic poliomyelitis. Virology **196:**199–208.
10. **Gray, J. J., J. Green, C. Cunliffe, C. Gallimore, J. V. Lee, K. Neal, and D. W. Brown.** 1997. Mixed genogroup SRSV infections among a party of canoeists exposed to contaminated recreational water. J. Med. Virol. **52:**425–429.
11. **Green, J., C. I. Gallimore, J. P. Norcott, D. Lewis, and D. W. Brown.** 1995. Broadly reactive reverse transcriptase polymerase chain reaction for the diagnosis of SRSV-associated gastroenteritis. J. Med. Virol. **47:**392–398.
12. **Green, J., J. Vinje, C. I. Gallimore, M. Koopmans, A. Hale, D. W. Brown, J. C. Clegg, and J. Chamberlain.** 2000. Capsid protein diversity among Norwalk-like viruses. Virus Genes **20:**227–236.
13. **Green, K. Y., R. M. Chanock, A. Z. Kapikian.** 2001. Human caliciviruses, p. 841–873. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, Fourth Edition, vol. 1. Lippincott Williams & Wilkins, Philadelphia, Pa.
14. **Hafliger, D., M. Gilgen, J. Luthy, and P. Hubner.** 1997. Seminested RT-PCR systems for small round structured viruses and detection of enteric viruses in seafood. Int. J. Food Microbiol. **37:**27–36.
15. **Hansman, G. S., K. Katayama, N. Maneekarn, S. Peerakome, P. Khamrin, S. Tonusin, S. Okitsu, O. Nishio, N. Takeda, and H. Ushijima.** 2004. Genetic diversity of norovirus and sapovirus in hospitalized infants with sporadic cases of acute gastroenteritis in Chiang Mai, Thailand. J. Clin. Microbiol. **42:**1305–1307.
16. **Hardy, M. E., S. F. Kramer, J. J. Treanor, and M. K. Estes.** 1997. Human calicivirus genogroup II capsid sequence diversity revealed by analyses of the prototype Snow Mountain agent. Arch. Virol. **142:**1469–1479.
17. **Hohne, M., and E. Schreier.** 2004. Detection and characterization of norovirus outbreaks in Germany: application of a one-tube RT-PCR using a fluorogenic real-time detection system. J. Med. Virol. **72:**312–319.
18. **Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. J. Virol. **74:**1234–1240.
19. **Jiang, X., C. Espul, W. M. Zhong, H. Cuello, and D. O. Matson.** 1999. Characterization of a novel human calicivirus that may be a naturally occurring recombinant. Arch. Virol. **144:**2377–2387.
20. **Jiang, X., M. Wang, K. Wang, and M. K. Estes.** 1993. Sequence and genomic organization of Norwalk virus. Virology **195:**51–61.
21. **Johansson, P. J., M. Torven, A. C. Hammarlund, U. Bjorne, K. O. Hedlund, and L. Svensson.** 2002. Food-borne outbreak of gastroenteritis associated with genogroup I calicivirus. J. Clin. Microbiol. **40:**794–798.
22. **Kageyama, T., M. Shinohara, K. Uchida, S. Fukushi, F. B. Hoshino, S. Kojima, R. Takai, T. Oka, N. Takeda, and K. Katayama.** 2004. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to norovirus in Japan. J. Clin. Microbiol. **42:**2988–2995.
23. **Karst, S. M., C. E. Wobus, M. Lay, J. Davidson, and H. W. T. Virgin.** 2003. STAT1-dependent innate immunity to a Norwalk-like virus. Science **299:** 1575–1578.
24. **Katayama, K., H. Shirato-Horikoshi, S. Kojima, T. Kageyama, T. Oka, F. Hoshino, S. Fukushi, M. Shinohara, K. Uchida, Y. Suzuki, T. Gojobori, and N. Takeda.** 2002. Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. Virology **299:**225–239.
25. **Koopmans, M., H. Vennema, H. Heersma, E. van Strien, Y. van Duynhoven, D. Brown, M. Reacher, and B. Lopman.** 2003. Early identification of common-source foodborne virus outbreaks in Europe. Emerg. Infect. Dis. **9:**1136–1142.
26. **Lai, M. M.** 1992. RNA recombination in animal and plant viruses. Microbiol. Rev. **56:**61–79.
27. **Lambden, P. R., E. O. Caul, C. R. Ashley, and I. N. Clarke.** 1993. Sequence and genome organization of a human small round-structured (Norwalk-like) virus. Science **259:**516–519.
28. **Lochridge, V. P., and M. E. Hardy.** 2003. Snow Mountain virus genome sequence and virus-like particle assembly. Virus Genes **26:**71–82.
29. **Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray.** 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. **73:**152–160.
30. **Lopman, B., Y. van Duynhoven, F. X. Hanon, M. Reacher, M. Koopmans, and D. Brown.** 2002. Laboratory capability in Europe for foodborne viruses. Euro. Surveill. **7:**61–65.
31. **Lopman, B. A., D. W. Brown, and M. Koopmans.** 2002. Human caliciviruses in Europe. J. Clin. Virol. **24:**137–160.
32. **Noel, J. S., T. Ando, J. P. Leite, K. Y. Green, K. E. Dingle, M. K. Estes, Y. Seto, S. S. Monroe, and R. I. Glass.** 1997. Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995. J. Med. Virol. **53:**372–383.
33. **Nygard, K., M. T., C. Ancker, S. B. Knauth, K. O. Hedlund, J. Gieseke, Y. Andersson, and L. Svensson.** 2003. Emerging genotype (GGIIb) of norovirus in drinking water, Sweden. Emerg. Infect. Dis. **9:**1548–1552.
34. **Oberste, M. S., S. Penaranda, and M. A. Pallansch.** 2004. RNA recombination plays a major role in genomic change during circulation of coxsackie B viruses. J. Virol. **78:**2948–2955.
35. **Parshionikar, S. U., S. Willian-True, G. S. Fout, D. E. Robbins, S. A. Seys, J. D. Cassady, and R. Harris.** 2003. Waterborne outbreak of gastroenteritis associated with a norovirus. Appl. Environ. Microbiol. **69:**5263–5268.
36. **Rohayem, J., and R. A. Münch.** 2005. Evidence of Recombination in the norovirus capsid gene. J. Virol. **79:**4977–4990.
37. **Suzuki, Y., T. Gojobori, and O. Nakagomi.** 1998. Intragenic recombinations in rotaviruses. FEBS Lett. **427:**183–187.
38. **Vennema, H., E. Kohli, D. Brown, E. Schreier, C. H. von Bonsdorff, L. Svensson, F. Hanon, M. Koopmans for the Consortium of Foodborne Viruses in Europe.** 2002. International spread in 8 countries in Europe of new pol-type calicivirus associated with 3 different capsids facilitated by food- and waterborne transmission. XIIth International Congress of Virology, Paris, France.
39. **Vinje, J., J. Green, D. C. Lewis, C. I. Gallimore, D. W. Brown, and M. P. Koopmans.** 2000. Genetic polymorphism across regions of the three open reading frames of "Norwalk-like viruses." Arch. Virol. **145:**223–241.
40. **Vinje, J., R. A. Hamidjaja, and M. D. Sobsey.** 2004. Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses. J. Virol. Methods **116:**109–117.
41. **Vinje, J., H. Vennema, L. Maunula, C. H. von Bonsdorff, M. Hoehne, E. Schreier, A. Richards, J. Green, D. Brown, S. S. Beard, S. S. Monroe, E. de Bruin, L. Svensson, and M. P. Koopmans.** 2003. International collaborative study to compare reverse transcriptase PCR assays for detection and genotyping of noroviruses. J. Clin. Microbiol. **41:**1423–1433.
42. **Walter, J. E., J. Briggs, M. L. Guerrero, D. O. Matson, L. K. Pickering, G. Ruiz-Palacios, T. Berke, and D. K. Mitchell.** 2001. Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. Arch. Virol. **146:**2357–2367.
43. **Worobey, M., and E. C. Holmes.** 1999. Evolutionary aspects of recombination in RNA viruses. J. Gen. Virol. **80:**2535–2543.
44. **Zheng, D. P., T. Ando, R. I. Glass, and S. Monroe.** 2004. Norovirus classification and proposed strain nomenclature. Second International Calicivirus Conference.