The Isolation and Characterization of a Norwalk Virus-specific cDNA

Suzanne M. Matsui,† Jungsu P. Kim,‡ Harry B. Greenberg,* Wanchuang Su,* Qiming Sun,‡
Philip C. Johnson,‡ Herbert L. DuPont,‡ Lyndon S. Oshiro,† and Gregory R. Reyes*†

*Departments of Medicine (Gastroenterology) and Microbiology and Immunology, Stanford University, Stanford, California 94305; and 
Veterans Administration Medical Center, Palo Alto, California 94304; ‡Molecular Virology Department, Genelabs, Incorporated, 
Redwood City, California 94063; †Department of Medicine, The University of Texas Health Science Center at Houston, Houston, Texas 77030; and ‡California Department of Health Services, Viral and Rickettsial Disease Laboratory, Berkeley, California 94704

Abstract

Norwalk virus, an important cause of epidemic, acute, nonbacterial gastroenteritis in adults and children, has eluded adaptation to tissue culture, the development of an animal model, and molecular cloning. In this study, a portion of the Norwalk viral genome encoding an immunoreactive region was cloned from very small quantities of infected stool using sequence-independent single primer amplification. Six overlapping complementary DNA (cDNA) clones were isolated by immunologic screening. The expressed recombinant protein from a representative clone reacted with sera from a large group of recovering individuals. Nucleic acid sequence analysis of the clones revealed a single open reading frame of 981 nucleotides encoding a 15K protein. Only oligonucleotide probes specific for the positive strand (defined by the open reading frame) hybridized to an RNaseA-resistant, DNaseI-resistant nucleic acid sequence extracted from Norwalk-infected stool. Furthermore, RNA extracted from stool after infection, but not preinfection, stools from three of five volunteers hybridized to a Norwalk virus cDNA probe. Clone-specific oligonucleotide probes hybridized with cesium chloride gradient fractions containing purified Norwalk virion. In conclusion, an antigenic, protein-coding region of the Norwalk virus genome has been identified. This epitope has potential utility in future sero- and molecular epidemiologic studies of Norwalk viral gastroenteritis. (J. Clin. Invest. 1991. 87:1456-1461.) Key words: sequence-independent single primer amplification • gastroenteritis virus • viral diarrhea • cloning

Introduction

Norwalk and Norwalk-like viruses comprise a group of serologically diverse, small round-structured viruses implicated in common source outbreaks of gastroenteritis (1, 2). It is estimated that up to 65% of acute, nonbacterial gastroenteritis in the United States is attributable to these agents (3). Norwalk virus, named after a 1968 outbreak of gastroenteritis at a school in Norwalk, Ohio, (4) was the first of this heterogeneous group of viruses to be identified by immune electron microscopy (IEM) (5). It alone appears to cause at least one-third of all cases of gastroenteritis in U.S. epidemics (6). While this illness typically lasts only 1–2 days, its impact on days lost from work or school may be substantial (7).

The study of these viruses has been restricted by the inability to cultivate these agents in vitro and lack of an animal model. Much of what is known about the immune response to this virus has come from human volunteer studies (8–10) or epidemiologic surveys (6). Positive identification of these viruses relies on IEM, a technique not well-suited to large-scale screening. Seroconversion and detection of viral antigens in fecal specimens are also used to establish the association of these viruses with episodes of acute gastroenteritis (11, 12). Unequivocal diagnosis, however, requires the use of paired human sera (pre- and postinfection). The molecular characterization of this agent has likewise been difficult, because the virus is shed in very small amounts in stool.

We report the cloning of an immunoreactive region of the Norwalk virus from infected fecal material. Sequence-independent single primer amplification (SISPA) (13; Reyes, G. R., and J. P. Kim, manuscript submitted for publication), a recently described and generally applicable technique of nonselective cDNA amplification, was used to overcome the restrictions imposed by the very limited amount of volunteer clinical material containing viral particles and the lack of existing nucleotide sequence information. The specificity of these clones for Norwalk virus was shown by (a) the immunoreactivity of the expressed peptide for only postinfection sera from several volunteers; (b) the hybridization of a clone-specific probe exclusively to RNA extracted from Norwalk antigen-positive volunteer stools; and (c) the convergence of the clone-specific probe hybridization peak and the Norwalk viral antigen peak in a cesium chloride gradient fraction of buoyant density 1.37 g/ml.

Methods

Virus and antisera. LT1-8, the eighth diarrheal stool specimen collected from volunteer LT (14) after oral inoculation with infectious
fetal filtrate 8Fla (8), was used to construct a cDNA library (see below). Occasional Norwalk virus particles were demonstrated by IEM in this specimen and in LT1-3, the third postinfection stool from the same volunteer. Particle counts were estimated between 10^2 and 10^4 viral particles per g of feces. LT0, the preinfection stool from this volunteer, was IEM negative.

Serial stool samples from other volunteers (L-V) inoculated with 8Fla were used for dot blot hybridization (see below). In each case, viral antigen was not detected by RIA (11) and/or ELISA (12) in preinfection specimens, but was present in one or more postinfection specimens from symptomatic volunteers. Viral particles were detected by IEM only in pooled specimens from volunteer IV.

Seven pairs of volunteer sera demonstrated a significant rise in titer between the pre- and the postinfection specimens by ELISA. A pair of sera that showed no change in titer (1:100) between pre- and postinfection samples was used as the negative control.

**Virus purification and nucleic acid extraction.** 7.5 g of specimen LT1-8 was partially purified using fluorocarbon extraction and pelleting through a sucrose suspension onto a cesium chloride cushion, as described previously (15). Viral antigen was monitored at each step by ELISA. LT0 control stool and LT1-3 were processed in tandem with LT1-8.

Serial stool samples from volunteers L-V, including uninfected controls, were partially purified by a modified procedure. 500 µl of a 10% fecal suspension (in PBS) was combined with 125 µl of 40% polyethylene glycol (PEG, mol wt 8,000) and precipitated overnight at 4°C. The suspension was pelleted at 10,000 g for 20 min.

Nucleic acid was extracted from half of the partially purified stool by a one-step guanidinium/phenol extraction procedure (16). Aliquots of nucleic acid were tested for sensitivity to treatment with DNase-free, RNaseA or RNase-free, DNaseI. The remainder of the partially purified stool was layered onto a 1.2–1.6 g/ml discontinuous cesium chloride gradient (15). The density and Norwalk antigenicity (by ELISA; (12) of each gradient fraction was determined.

**cDNA cloning and analysis.** ~ 10 µg of nucleic acid obtained from 1.5 g of original fecal material was reverse transcribed into cDNA (17), using an oligo (dT) primer. Double-stranded linker-primer oligonucleotides were directionally ligated onto the cDNA amplified by SISPA (13; Reyes, G. R., and J. P. Kim, manuscript submitted for publication) and then digested with EcoRI. Excess linkers were removed by passage through Sephadex 300, before inserting the amplified cDNA into bacteriophage λgt11.

Libraries were screened immunologically (18) with postinfection serum from volunteer DD (Norwalk ELISA titer ≥ 1:25,600). The serum was preadsorbed with λgt11 and used at a 1:200 dilution for immunoscreening. After three immunoscreenings, a purified plaque for each clone was obtained and the specificity assessed using DD preinfection sera (ELISA titer = 1:100).

cDNA from clones of interest was tested by Southern blot (19) or dot blot hybridization for similarity to human lymphocyte and Escherichia coli strain 1088 DNA and to other enteric viruses investigated in our laboratories (NCDV rotavirus and astrovirus type 1). Direct sequence comparisons with the hepatitis E genome (Tam, A. W., et al., manuscript submitted for publication) were also made.

**Subcloning and sequencing.** cDNA was subcloned into Bluescript SK+ (Stratagene Inc., La Jolla, CA) for sequencing (20). 7-deaza-dGTP was used in confirmatory sequencing reactions. Nucleic and amino acid sequences were compared with the Genbank database.

**Dot blot hybridization.** Nucleic acid was prepared for dot blot as recommended by the manufacturer of the apparatus (Schleicher and Schuell, Keene, NH). Hybridization was carried out in 50% formamide and 1 x hybridization buffer (5 x Denhardt's reagent, 5 x SSC, 50 mM NaH2PO4, 1 mM sodium pyrophosphate/Na2HPO4, 100 µg/ml salmon sperm DNA, 100 µg/ml AT) at room temperature, then 0.1 x SSC and 0.1% SDS at 65°C for 1 h.

**Radiolabeled probes.** Probes were made using direct polymerase chain reaction (PCR) (21) and primers of known λgt11 sequence (5' GGCAGACATGCGTGC CCGG 3' and 5' TCGACGTTTCTCA- TATGGGG 3'). The typical PCR cycle (denaturation, 94°C, 30 s; annealing, 50°C, 1 min; extension, 72°C, 2 min) was repeated for 30 cycles. The Norwalk-specific amplified fragment was separated from λgt11 sequences by EcoRI digestion, purified by preparative 1.5% agarose gel electrophoresis, and radiolabeled by random priming.

An N49-specific fragment was amplified by PCR, using oligonucleotide primers shown in Fig. 1 A and modified PCR conditions (each segment 30 s). Radiolabeled, random-primed probes were synthesized. Single-stranded oligonucleotide probes were made by 3-end-labeling N49-specific, synthetic oligonucleotide primers with γ[32P]ATP. Hybridization was carried out in 30% formamide and 1 x hybridization buffer at 42°C. Filters were washed in 2 x SSC at room temperature for 30 min.

**Results**

Our cloning strategy was dictated by the small amount of Norwalk-infected, IEM-positive, fecal specimen available. The original stool contained 10^3–10^4 virion particles per g as estimated by IEM and 1.5 g of original fecal material was used in cDNA synthesis. cDNA was synthesized from partially purified virion and the heterogeneous population nonselectively amplified by SISPA (13; Reyes, G. R., and J. P. Kim, manuscript submitted for publication). The cDNA library was constructed in bacteriophage λgt11 for clone identification.

The first three rounds of immunoscreening were accomplished using DD postinfection serum. In the first screening, 41 reactive plaques were identified from a total of 1.2 x 10^5 recombinant phage. After two more rounds of immunoscreening and plaque purification, a final screening yielded six clones (N28, N35, N40, N48, N49, and N51) that were chosen by virtue of their specific reactivity with DD postinfection, but not DD preinfection serum.

The initial characterization of these clones demonstrated the absence of hybridization to Southern blots (19) of either human or E. coli DNA (data not shown). In addition, the representative N49 probe (see below) did not hybridize to NCDV rotavirus RNA or astrovirus type 1 cDNA (data not shown). N49 nucleotide and peptide sequences (see below) were distinct from the full length hepatitis E virus sequences (Tam, A. W., et al., manuscript submitted for publication).

Two approaches established these immunoreactive clones as derived from only Norwalk-infected but not preinfection fecal specimens. First, we demonstrated that probes from all clones hybridized only to the SISPA-expanded cDNA population from LT1-8 and not to a similarly prepared LT0 preinfection cDNA specimen (data not shown). In a second test, a [32P]labeled random-primed probe derived from clone N49 hybridized only to RNA extracted from two Norwalk IEM-positive postinfection specimens, LT1-8 and LT1-3. This probe did not hybridize with nucleic acid extracted from negative controls, LT0 or NCDV rotavirus (data not shown).

The relationship of the six immunoreactive clones to one another was next examined by hybridization and sequence analysis. The clones ranged in size from 172 bp (N40) to 380 bp (N35) (Fig. 1 A) and all six cross-hybridized (data not shown). Each of the clones had a single open reading frame, contiguous with the β-galactosidase reading frame in λgt11. Regions of overlap were verified by sequencing and are shown in Fig. 1. Five different start sites were represented among six isolated clones. Two regions of disparity between the N35 reference sequence and the other clone sequences were found. The heterogeneity in the 3' end suggested that no authentic, 3'-end,
poly (A) tail was present in these clones derived from the oligo (dT)-primed library. These sequences were unique at both the nucleic and amino acid levels when compared to the GenBank\textsuperscript{[a]} database. Two of these clones, N49 and N51, had identical nucleotide sequences and N49 was selected for closer study.

Immunological screening of N49 with an expanded panel of seven paired Norwalk volunteer sera provided further sup-
port that these clones expressed an immunogenic epitope of Norwalk virus. As shown in Table I, the expressed product from clone N49 reacted with six of the seven high titer, postinfection sera. Only preinfection serum AT, which had high preexisting levels of reactivity to Norwalk antigen by ELISA, also reacted with the N49 clone product. No other preinfection sera or the negative control paired sera (EM) reacted with the N49 clone product.

As a second proof that clone N49 represented a portion of the Norwalk virus genome, RNA was extracted from preinfection and serial postinfection stools from five additional volunteers who became ill after receiving 8FIIa. As shown in Fig. 2, the N49 probe hybridized to nucleic acid from test specimens I-4 and -5, III-4, -5, -6, and -7, and IV-8 and -9. These represented the third through eighth stools collected from three symptomatic volunteers. In each case, N49 hybridization identified a peak of viral shedding in these serial postinfection stools that contained viral antigen. Hybridization of this probe could not be demonstrated in the specimens of volunteers II and V who had only one antigen-positive stool specimen each. No hybridization could be demonstrated with nucleic acid isolated from preinfection stools.

Furthermore, the clone N49-specific probe hybridized most prominently to sample LT1-8 cesium chloride density gradient fractions 7 through 10 (Fig. 3). Peak hybridization signal and peak Norwalk antigen signal by ELISA were both found in fraction 8. Buoyant density of this fraction was 1.37 g/ml. Simultaneously prepared LTO cesium chloride gradient fractions showed neither hybridization with the N49-specific probe nor Norwalk antigenic activity.

Table I. Summary of Immunologic Reactivity of Norwalk Clone N49 with Paired Sera from Infected Volunteers

<table>
<thead>
<tr>
<th>Serum</th>
<th>Reactivity of sera to Norwalk by ELISA*</th>
<th>Immuneactivity of clone N49†</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD preinfection</td>
<td>1:100</td>
<td>-</td>
</tr>
<tr>
<td>DD postinfection</td>
<td>1:25,600</td>
<td>+</td>
</tr>
<tr>
<td>AT preinfection</td>
<td>1:1,600</td>
<td>+</td>
</tr>
<tr>
<td>AT postinfection</td>
<td>1:102,400</td>
<td>+</td>
</tr>
<tr>
<td>SL preinfection</td>
<td>&lt;1:100</td>
<td>-</td>
</tr>
<tr>
<td>SL postinfection</td>
<td>1:6,400</td>
<td>+</td>
</tr>
<tr>
<td>DF preinfection</td>
<td>1:400</td>
<td>-</td>
</tr>
<tr>
<td>DF postinfection</td>
<td>1:6,400</td>
<td>+</td>
</tr>
<tr>
<td>JY preinfection</td>
<td>&lt;1:100</td>
<td>-</td>
</tr>
<tr>
<td>JY postinfection</td>
<td>1:25,600</td>
<td>+</td>
</tr>
<tr>
<td>CS preinfection</td>
<td>&lt;1:100</td>
<td>-</td>
</tr>
<tr>
<td>CS postinfection</td>
<td>1:6,400</td>
<td>-</td>
</tr>
<tr>
<td>519 preinfection</td>
<td>1:100</td>
<td>-</td>
</tr>
<tr>
<td>519 postinfection</td>
<td>1:25,600</td>
<td>+</td>
</tr>
<tr>
<td>EM preinfection</td>
<td>&lt;1:100</td>
<td>-</td>
</tr>
<tr>
<td>EM postinfection</td>
<td>&lt;1:100</td>
<td>-</td>
</tr>
</tbody>
</table>

* The Norwalk ELISA was performed as described previously (12).
† Assay performed by mixing plaque purified N49 with nonrecombinant lambda gt11 (internal negative control) and plating at a 1:1 ratio. Reactive plaques were detected after first antibody incubation ("serum") using an alkaline phosphatase-conjugated anti-IgG second antibody. The absence (−) or presence (+) of immunoreactivity was determined by two independent readers. Selected reactive plaques were confirmed to contain the N49-specific insert by PCR (21).

Figure 2. Hybridization of an N49-specific, radiolabeled probe to RNA from serial, Norwalk volunteer, fecal specimens. Specimens are numbered in order of collection, with specimen 1 representing the antigen-negative, preinfection stool for each. All postinfection fecal samples tested were antigen positive. Viral particles were detected by IEM only in pooled, postinfection stools from volunteer IV (*). N49-specific hybridization was found in nucleic acid extracts of stools from volunteers I (specimens 4 and 5), III (specimens 4, 5, 6, and 7), and IV (specimens 8 and 9).

Figure 3. Correlation of Norwalk antigenic activity with N49-specific, radiolabeled probe hybridization in cesium chloride gradient fractions of buoyant density 1.31–1.40 g/ml. Peak antigenic activity and peak hybridization signal were detected in fraction 8, buoyant density 1.37 g/ml.
Finally, we examined the nature of the extracted nucleic acid from which the cDNA was synthesized. As shown in Fig. 4 A, RNaseA treatment of LT1-8 and LT1-3 nucleic acid extracts eliminated hybridization with the N49 probe, while DNaseI treatment had no effect. Only single-stranded probes of negative polarity (as defined by open reading frame analysis) hybridized to LT1-8 nucleic acid (Fig. 4 B).

Discussion

We conclude that we have cloned an immunoreactive portion of the Norwalk virus genome. The isolated clones were derived from nucleic acid extracted from partially purified fecal material that contained a very small number of intact viral particles. The SISPA technique (13; Reyes, G. R., and J. P. Kim, manuscript submitted for publication) was used to successfully amplify the LT1-8 cDNA, before the identification of a set of overlapping immunoreactive clones specific to postinfection sera specimens. SISPA coupled to conventional cloning and immunoscreening techniques appears to be a generally applicable strategy for identifying cDNA clones of fastidious infectious pathogens derived from limited amounts of starting material.

The expressed recombinant protein of representative clone N49 was recognized specifically by a panel of high titer postinfection sera from several volunteers, but not by antibody-negative preinfection sera or by paired sera without antibody to Norwalk virus (Table I). The lack of reactivity of the N49 product with CS postinfection serum (ELISA titer = 1:6,400) is possibly due to antibodies in the CS specimen being directed to other Norwalk virus epitopes.

The N49-specific cDNA probe specifically hybridized to nucleic acid from only the postinfection stools of three different volunteers (Fig. 2). The positive hybridization signals correlated well with the period of symptomatic illness and antigen detection in the stool by RIA (11) and/or ELISA (12). This cDNA-derived probe was more sensitive than IEM for detecting small quantities of virus in stool. Generally, the limit of detection by IEM in feces is in the range of 10^3–10^6 particles per ml (5). RIA and/or ELISA, however, may be able to detect viral antigens in the stool before and after the period of shedding of intact viral particles.

Analysis of LT1-8 cesium chloride density gradient fractions by ELISA and dot blot hybridization provided further support that the isolated clones represent portions of the Norwalk viral genome (Fig. 3). Peak Norwalk antigenic activity and peak clone N49-specific hybridization signal were both observed in fraction 8. The buoyant density of fraction 8, 1.37 g/ml, is in the range of that which has been reported for Norwalk (15) and Norwalk-like viruses (2).

The viral protein structure of Norwalk virus (15) and Norwalk-like viruses, such as Snow Mountain agent (22), most closely resembles caliciviruses that have a single structural protein of molecular mass 62 kD (23). Furthermore, a one-way serologic relationship between caliciviruses and Norwalk virus has been demonstrated by RIA (24). These observations have led to the hypothesis that Norwalk virus may be a plus-sense, single-stranded RNA virus with a 3'-end poly (A) tail, like caliciviruses. We demonstrate here that the Norwalk virus has a single-stranded RNA genome of positive polarity (Fig. 3). These findings lend support to the speculation that Norwalk virus may be related to caliciviruses (23), but further characterization will be required to confirm this classification. An authentic poly (A) tail was not identified in these clones. The most likely explanation for the 11 to 12-bp stretches of (A) residues at the 3' end of clones N49/N51, N48, N28, and N40, is that the (A)-rich regions of the sequence (Fig. 1 A) served as the initiation sites of cDNA synthesis by oligo (dT).

The identified immunoreactive clones can now be used to obtain a complete set of overlapping genomic cDNA clones and nucleotide sequence of the Norwalk virus. Both will be helpful in studying Norwalk virus and establishing its relatedness to other viruses. With the PCR and sequence information now available, the development of useful and highly sensitive diagnostic reagents should be relatively straightforward.

The Norwalk sequence identified in these studies appears to encode a common immunoreactive viral epitope. It seems likely that this clone can be employed to produce a recombinant antigen for use in large scale, antibody prevalence, seroepidemiologic studies. Such analyses should help further clarify the distribution and importance of Norwalk infection.

Acknowledgments

We would like to thank Drs. K. Fry and E. Mackow for helpful discussions. We also gratefully acknowledge the assistance of R. Lo, K. Yun, J. Fernandez, and the Genelabs Visual Arts Department.

---

**Figure 4.** The Norwalk genome is composed of single-stranded RNA of positive polarity. (A) The N49-specific, radiolabeled probe failed to hybridize with RNaseA-treated nucleic acid (column 1), but hybridized with both DNaseI-treated (column 2) and untreated (column 3) nucleic acid from Norwalk-infected stools (LT1-3 and LT1-8). (B) A negative sense, N49 oligonucleotide probe (N49 residues 208 to 227, Fig. 1 A) hybridized to RNA extracted from infected stool LT1-8, indicating that the RNA is of positive polarity (spot 1). A positive sense probe (N49 residues 71 to 91, Fig. 1 A) did not hybridize with LT1-8 RNA (spot 2).
This work was supported in part by Public Health Service grants DK-01811 (to S. Matsui), DK-38707 (to the Stanford Digestive Disease Center), and RR-02558 (to the University of Texas Clinical Research Center), and a Veterans Administration Merit Review grant (to H. Greenberg).

Note added in proof: During the revision of this manuscript, a paper describing the cloning of a portion of the Norwalk virus was published elsewhere (Jiang, X., D. Y. Graham, K. Wang, and M. K. Estes. 1990. Science (Wash. DC). 250:1580–1583). We have identified an immunoreactive region of the Norwalk virus that is different from the polymerase region reported by Jiang et al.

References


Norwalk Virus-specific cDNA 1481