Diarrhea, characterized by frequent liquid or loose stools, commonly results from gastroenteritis caused by infection with bacteria, parasites, or viruses. Patients with mild diarrhea do not require medical attention; the illness is typically self-limited, and disease symptoms usually resolve quickly. However, diarrheal diseases can result in severe illness and death worldwide and are the second leading cause of death around the world in children <5 years of age, particularly in low- and middle-income countries (1). For many cases of diarrhea in humans, no causative agent is identified.

In recent years, many novel viruses have been identified in human and animal blood, respiratory secretions, and fecal material through viral metagenomic studies consisting of random amplification in combination with next-generation sequencing methods (2–5). To identify unknown human viruses in the enteric tracts of persons with diarrhea, we performed sequence-independent amplification on purified viral nucleic acid from fecal samples obtained from patients with diarrhea in Bangladesh (6,7). We identified a novel calicivirus and classified it in the proposed genus Recovirus. Caliciviruses, which are nonenveloped, positive-stranded RNA viruses with a polyadenylated genome of ~6.4–8.4 kb, cause illness in animals and humans (8,9), including gastroenteritis in humans. The family Caliciviridae consists of 5 genera, Norovirus, Sapovirus, Lagovirus, Vesivirus, and Nebovirus, and 3 proposed genera, Recovirus, Valovirus, and chicken calicivirus (8–10).

The Study

Each year, >100,000 diarrhea patients are admitted to the Dhaka hospital of the International Centre for Diarrheal Disease Research, Bangladesh (ICDDR,B). Fecal samples from 2% of these patients are collected and examined as part of systematic routine surveillance system for the presence of enteric pathogens (11). All procedures were performed in compliance with relevant laws and institutional guidelines and in accordance with the Declaration of Helsinki.

Figure 1. Schematic outline of the strategies used for PCR amplification of calicivirus Bangladesh/289/2007. A) Schematic representation of the calicivirus Bangladesh/289/2007 genome. Boxes represent the open reading frames encoding the calicivirus proteins. Indicated are the poly(A)-tail (An); putative cleavage sites indicated by XX↓XX. The 5′ end of the genome was not obtained. The bottom of the panel shows a schematic outline of the reverse transcription PCRs employed to amplify calicivirus Bangladesh/289/2007 sequences by using random amplification, degenerate PCR, and 3′ rapid amplification of cDNA ends (RACE) PCR. The orientations and positions of the oligonucleotides on the calicivirus genome are depicted and sequences shown in online Technical Appendix Table 2 (wwwnc.cdc.gov/EID/pdfs/12-0344-Techapp.pdf). B) Genome organization of caliciviruses in the genera Vesivirus, Nebovirus, Norovirus, Sapovirus, and Lagovirus and proposed genera Valovirus and Recovirus, for comparison with the new calicivirus Bangladesh/289/2007. The 5′ end of the genome is shown with a Vpg protein (black dots). Numbers indicate the nucleotide positions according to the virus genome for which the GenBank accession number is indicated.

DISPATCHES

Calicivirus from Novel Recovirus Genogroup in Human Diarrhea, Bangladesh

Saskia L. Smits, Mustafizur Rahman, Claudia M.E. Schapendonk, Marije van Leeuwen, Abu S.G. Faruque, Bart L. Haagmans, Hubert P. Endtz, and Albert D.M.E. Osterhaus

To identify unknown human viruses in the enteric tract, we examined 105 stool specimens from patients with diarrhea in Bangladesh. A novel calicivirus was identified in a sample from 1 patient and subsequently found in samples from 5 other patients. Phylogenetic analyses classified this virus within the proposed genus Recovirus.
Table 1. Percentage amino acid identity of calicivirus Bangladesh/289/2007 with representative caliciviruses of other genera*

| Calicivirus genus | Strain          | GenBank accession no. | NS polyprotein† | VP1   | VP2   |
|------------------|-----------------|-----------------------|-----------------|-------|-------|
| Norovirus        | Norwalk         | M87661                | 27.7            | 30.4  | 24.7  |
|                  | Southampton     | L07418                | 29.0            | 31.2  | 23.3  |
| Sapovirus        | Manchester      | X86560                | 20.1            | 18.6  | 4.1   |
|                  | PEC Cowden      | AF182760              | 18.3            | 18.1  | 6.8   |
| Vesivirus        | FCV CF68        | U13992                | 18.4            | 19.3  | 15.1  |
|                  | SMSV 1          | U15301                | 19.0            | 19.3  | 5.5   |
| Lagovirus        | RHDV FRG        | K67473                | 18.0            | 19.6  | 9.6   |
|                  | EBHSV GD        | Z69620                | 17.3            | 17.8  | 5.5   |
| Nebovirus        | BECT Newbury 1  | AY082991              | 18.7            | 16.8  | 11.0  |
|                  |                 | DQ013304              | 19.0            | 16.8  | 11.0  |
| Recovirus‡       | Tulane          | EU391643              | 57.7            | 49.5  | 53.1  |
| Valovirus‡       | AB104           | FJ355930              | 35.8            | 43.5  | 26.0  |

*NS, nonstructural; VP, viral protein.
†All polyprotein sequences were aligned without the capsid protein encoding sequences.
‡Tentative genus not yet accepted by the International Committee on Taxonomy of Viruses.

Calicivirus, Bangladesh

Stool specimens from a subset of patients from routine surveillance during 2007–2009 (1,614 samples total) were available for further studies. These specimens were prescreened for adenovirus and rotavirus A by using TaqMan EZ RT-PCR Core Reagents (Applied Biosystems, Foster City, CA, USA), rotavirus primers RVNSP3R and RVNSP3F and probe 5'-FAM-AGTTAAAAGCTAA-CACTGTCAAA-TAMRA-3' (12), and TaqMan Universal Mastermix (Applied Biosystems) (13). Sequence-independent nucleic acid amplification and next-generation sequencing were performed on 105 stool specimens from diarrhea patients enrolled during 2007 by using a 454 GS Junior Instrument (Roche, Indianapolis, IN, USA) as described (6,7). More than 725,000 trimmed reads were assembled by using de novo assembly and analyzed according to BLAST searches (online Technical Appendix Figure 1 and Table 1, wwwnc.cdc.gov/EID/pdfs/12-0344-Techapp.pdf) (6,7). Sequences were classified on the basis of the taxonomic origin of the best-hit sequence (6,7). An E (expect) value of 0.001 was used as cutoff value of significant virus hits. The largest proportion of virus-related sequences in human diarrhea samples from Bangladesh in 2007 was related to known bacteriophages and mammalian viruses (online Technical Appendix Figure 1).

One sample, no. 289, yielded a novel mammalian virus from the family Caliciviridae that we further characterized by near full-length genome sequencing using random amplification with next-generation sequencing, specific reverse transcription PCRs, and 3' rapid amplification of cDNA ends PCR (Figure 1, panel A) as described (6,7). We named the virus isolate calicivirus Bangladesh/289/2007 (GenBank accession no. JQ745645).

The Caliciviridae genome encodes a polyprotein precursor for nonstructural proteins, and 2 structural capsid proteins, viral protein (VP) 1 and VP2, from 2 or 3 open reading frames (ORFs) (Figure 1, panel B) (8,9). The genome organization of Bangladesh/289/2007 is most closely related to that of viruses in the genus Norovirus, with ORF2 encoding VP1 overlapping with ORF1 (Figure 1). The partial genome encodes a polyprotein precursor for nonstructural proteins, and 2 structural capsid proteins, viral protein (VP) 1 and VP2, from 2 or 3 open reading frames (ORFs) (Figure 1, panel B) (8,9). The genome organization of Bangladesh/289/2007 is most closely related to that of viruses in the genus Norovirus, with ORF2 encoding VP1 overlapping with ORF1 (Figure 1). The partial genome encodes a polyprotein precursor for nonstructural proteins, and 2 structural capsid proteins, viral protein (VP) 1 and VP2, from 2 or 3 open reading frames (ORFs) (Figure 1, panel B) (8,9). The genome organization of Bangladesh/289/2007 is most closely related to that of viruses in the genus Norovirus, with ORF2 encoding VP1 overlapping with ORF1 (Figure 1). The partial genome encodes a polyprotein precursor for nonstructural proteins, and 2 structural capsid proteins, viral protein (VP) 1 and VP2, from 2 or 3 open reading frames (ORFs) (Figure 1, panel B) (8,9). The genome organization of Bangladesh/289/2007 is most closely related to that of viruses in the genus Norovirus, with ORF2 encoding VP1 overlapping with ORF1 (Figure 1).
tial polyprotein precursor and complete VP1 and VP2 proteins were aligned with corresponding sequences of representative caliciviruses. Divergence analysis demonstrated that the calicivirus Bangladesh/289/2007 showed 57.7%, 49.5%, and 53.1% amino acid identity in the polyprotein, VP1, and VP2, respectively, to its closest relative, Recovirus Tulane (GenBank accession no. EU391643), which was identified from rhesus macaques (Table 1; Figure 2).

Additionally, on the basis of the identity profiles of noroviruses (14)—strains, genotypes, and genogroups showed 85.9%–100%, 56.2%–85.7%, and 38.6%–55.1% identity, respectively—Bangladesh/289/2007 may represent a new genogroup in the proposed genus Recovirus. This conclusion was corroborated by comparison of a partial RNA-dependent RNA polymerase (RdRp) sequence of calicivirus Bangladesh/289/2007 to known recovirus RdRp sequences (online Technical Appendix Figure 2).

We performed a diagnostic real-time recovirus PCR targeting the RdRp of all 1,614 available samples from patients with diarrhea (11). Reverse transcription PCR–grade viral nucleic acid was extracted by using the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche) and amplified by using reverse transcription PCR with primers VS665 and VS666 and probe VS664 (online Technical Appendix Table 2) and TaqMan EZ RT-PCR Core Reagents (Applied Biosystems). The cycling program consisted of 50°C for 2 min, 60°C for 30 min, 95°C for 5 min, and 50 cycles of 95°C for 20 s and 59°C for 1 min, resulting in a 164-bp amplicon.

In addition to sample 289, 5 other human diarrhea samples (Table 2) were sequence-confirmed to be positive for a recovirus, with high homology (>98%) to calicivirus Bangladesh/289/2007; this finding indicates that recovirus Bangladesh infects humans. Clinical data indicate that all patients with a recovirus present in their feces had 6 to 21 watery stools in the first 24 hours after illness onset (Table 2); 4 patients experienced vomiting and 2 patients had fever. Patient ages ranged from 3 months to 50 years. Three recovirus-positive patients showed evidence of co-infection with other pathogens that are known to cause diarrhea in humans, such as rotavirus A, adenovirus, *Vibrio cholerae*, or *Salmonella* spp.; the other 3 patients did not. Although viruses such as norovirus, sapovirus, and astrovirus were not detected in the recovirus-positive samples by sequence-independent amplification assays, all samples were not analyzed for all known enteric pathogens. Of 514 diarrhea samples gathered by the Diagnostic Unit, Department of Virology, Erasmus Medical Center, Rotterdam, the Netherlands, for diagnosis of gastrointestinal infections during 2007 and 2009, none was positive for Recovirus Bangladesh (data not shown).

Conclusions

For a large proportion of human diarrhea cases, no etiologic agent can be identified, despite multiple metagenomic studies of viruses in human stool aimed at identifying new etiologic agents (3–5). In addition, it cannot be known when and where emerging and reemerging viruses will appear in the human population. To identify potential etiologic agents of diarrhea in humans, we performed a metagenomic viral inventory in diarrhea samples from Bangladesh, which led to the identification of a novel calicivirus.

Although no species demarcation criteria have been defined for the family *Caliciviridae* by the International Committee on the Taxonomy of Viruses, we classified calicivirus Bangladesh/289/2007 in the proposed genus Recovirus, primarily on the basis of phylogenetic analyses (8). Bangladesh/289/2007 may also be considered a new genogroup of the genus Recovirus for the following reasons: 1) the genetic distance between Recovirus Tulane and Bangladesh/289/2007 VP1 is similar to that of VP1 capsids of noroviruses belonging to different genogroups (14); 2) the genetic distances between macaque recoviruses and Recovirus Bangladesh/289/2007 RdRP are similar to that of RdRP sequences of recoviruses belonging to different genogroups (15); 3) the genome organization of Recovirus Tulane and Bangladesh/289/2007 differs; and 4) Recovirus Tulane and Bangladesh/289/2007 were identified in different host species.
In conclusion, this identification of a novel calicivirus, classified as Recovirus Bangladesh/289/2007, from human diarrhea samples provides PCR-based evidence that recoviruses can infect humans. A previous study found high prevalence of virus-neutralizing antibodies against a closely related calicivirus, Recovirus Tulane, in serum samples from animal caretakers (15). Larger epidemiologic studies using genetic and serologic screening will be necessary to provide more insight into the distribution and pathogenic potential of recoviruses in humans.

This work was partially funded by the European Community’s Seventh Framework Program (FP7/2007–2013) under the project “European Management Platform for Emerging and Reemerging Infectious disease Entities” European Commission agreement no. 223498 and the Virgo Consortium. This research study was funded in part by ICDDR,B and its donors which provide unrestricted support to ICDDR,B for its operations and research. Current donors providing unrestricted support include the following: Australian Agency for International Development (AusAID), Government of the People’s Republic of Bangladesh; Canadian International Development Agency (CIDA), Swedish International Development Cooperation Agency (SIDA), and the Department for International Development, UK (DFID).

Dr Smits works at the Virology Department, Erasmus Medical Center, and Viroclinics Biosciences BV, Rotterdam, the Netherlands. Her research interests are hepatitis C virus and severe acute respiratory syndrome coronavirus.

References

1. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375:1969–87. http://dx.doi.org/10.1016/S0140-6736(10)60549-1
2. Tang P, Chiu C. Metagenomics for the discovery of novel human viruses. Future Microbiol. 2010;5:177–89. http://dx.doi.org/10.2217/fmb.09.120
3. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, et al. Metagenomic analyses of an uncultured viral community from human feces. J Bacteriol. 2003;185:6220–3. http://dx.doi.org/10.1128/JB.185.20.6220-6223.2003
4. Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, et al. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol. 2009;83:4642–51. http://dx.doi.org/10.1128/JVI.02301-08
5. Zhang T, Breitbart M, Lee WH, Run QJ, Wei CL, Soh SW, et al. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol. 2006;4:e3. http://dx.doi.org/10.1371/journal.pbio.0040003
6. van den Brand JM, van Leeuwen M, Schapendonk CM, Simon JH, Haagmans BL, Osterhaus AD, et al. Metagenomic analysis of the viral flora of pine marten and European badger feces. J Virol. 2012;86:2360–5. http://dx.doi.org/10.1128/JVI.01637-11
7. van Leeuwen M, Williams MM, Koraka P, Simon JH, Smits SL, Osterhaus AD. Human picobirnaviruses identified by molecular screening of diarrhea samples. J Clin Microbiol. 2010;48:1787–94. http://dx.doi.org/10.1128/JCM.02452-09
8. Farkas T, Sestak K, Wei C, Jiang X. Characterization of a rhesus monkey calicivirus representing a new genus of Caliciviridae. J Virol. 2008;82:5408–16. http://dx.doi.org/10.1128/JVI.00070-08
9. L’Homme Y, Sansregret R, Plante-Fortier E, Lamontagne AM, Ouardani M, Lacroix G, et al. Genomic characterization of swine caliciviruses representing a new genus of Caliciviridae. Virus Genes. 2009;39:66–75. http://dx.doi.org/10.1007/s11262-009-0360-3
10. Wolf S, Reetz J, Otto P. Genetic characterization of a novel calicivirus from a chicken. Arch Virol. 2011;156:1143–50. http://dx.doi.org/10.1007/s00705-011-0964-5
11. Stoll BJ, Glass RI, Huq MI, Khan MU, Banu H, Holt J. Epidemiologic and clinical features of patients infected with Shigella who attended a diarrheal disease hospital in Bangladesh. J Infect Dis. 1982;146:177–83. http://dx.doi.org/10.1093/infdis/146.2.177
12. Svraka S, Duizer E, Vennema H, de Bruin E, van der Veer B, Dorrresteijn B, et al. Etiological role of viruses in outbreaks of acute gastroenteritis in The Netherlands from 1994 through 2005. J Clin Microbiol. 2007;45:1389–94. http://dx.doi.org/10.1128/JCM.02305-06
13. Heim A, Ebnet C, Harste G, Pring-Akerblom P. Rapid and quantitative detection of human adenovirus DNA by real-time PCR. J Med Virol. 2003;70:228–39. http://dx.doi.org/10.1002/jmv.10382
14. Zheng DP, Ando T, Fankhauser RL, Beard RS, Glass RI, Monroe SS. Norovirus classification and proposed strain nomenclature. Virology. 2006;346:312–23. http://dx.doi.org/10.1016/j.virol.2005.11.015
15. Farkas T, Cross RW, Hargitt E III, Lerche NW, Murray AL, Sestak K. Genetic diversity and histo-blood group antigen interactions of rhesus enteric caliciviruses. J Virol. 2010;84:8617–25. http://dx.doi.org/10.1128/JVI.00630-10

Address for correspondence: Saskia Smits, Department of Virology, Erasmus Medical Center/Viroclinics Biosciences BV, PO Box 2040, 3000 CA Rotterdam, The Netherlands; email: smits@viroclinics.com
Detection of Human Calicivirus in Novel Recovirus Genogroup, Bangladesh

Technical Appendix

Figure 1. Distribution of the percentage of sequence read matches. (A) The 725305 sequence reads were assembled in 78676 contigs and 156008 single reads, which were analyzed using BLASTn searches. The contigs and single reads were classified based on the taxonomic origin in the annotation of best hit sequence (E-score <10^{-3}) as being of bacterial, eukaryotic, other or viral origin and

Technical Appendix Figure 1. Distribution of the percentage of sequence read matches. (A) The 725305 sequence reads were assembled in 78676 contigs and 156008 single reads, which were analyzed using BLASTn searches. The contigs and single reads were classified based on the taxonomic origin in the annotation of best hit sequence (E-score <10^{-3}) as being of bacterial, eukaryotic, other or viral origin and
depicted in a pie chart colored according to the key. The percentage of each category is based on the number of reads per category. A large percentage of reads showed no hits with any sequence in the Genbank database. (B) Taxonomic classification of sequences with similarity to known viruses. Circles located next to taxa are proportional to the total number of sequence reads identified in the indicated category of viruses with BLASTn (E<10⁻³). The number of sequence reads is included in brackets and the different colors indicate whether the virus infects mammals, plants, insects or bacteria. The table indicates which virus species were identified.

Technical Appendix Figure 2. Phylogenetic tree of the nucleotide sequences of the partial RdRP sequence of selected representative recoviruses and the newly identified recovirus Bangladesh/289/2007. Phylograms were generated using MEGA5, with the Neighbor-joining method with p-distance and 1,000 bootstrap replicates. Significant bootstrap values and Genbank accession numbers are shown, as well as genogroups and genotypes.
### Technical Appendix Table 1. Overview of sequence reads and contigs

| No. samples | Total no. reads | Average length, nt | No. assembled | No. unassembled | % Assembled | No. contigs | Average no. reads/sample |
|-------------|-----------------|--------------------|---------------|-----------------|-------------|-------------|--------------------------|
| 105         | 725305          | 222                | 569297        | 156008          | 78.5        | 78676       | 6908                     |

### Technical Appendix Table 2. Oligonucleotides primers

| Primer | Nucleotide sequence, 5'→3' | Orientation | Position* |
|--------|----------------------------|-------------|-----------|
| VS431  | GCGAGCACAGAATATAACGACT     | Reverse     | Not applicable |
| VS432  | CGCGATCCGAAATATACGACT      | Reverse     | Not applicable |
| VS627  | GGAGGAGAGAGATGCTG          | Forward     | 527–543   |
| VS828  | GACAGGTTCTGCTTGTAGG         | Reverse     | 1642–1659 |
| VS629  | GCCGCTTTCCGCAGCTTGG        | Forward     | 4396–4414 |
| VS630  | GACTGGCTCTAGGCAAGCTTC      | Forward     | 4407–4427 |
| VS631  | CAATTATACCAAGAAAGATC       | Forward     | 1530–1549 |
| VS634  | CCAGAACCAGGAAGATC          | Reverse     | 4285–4302 |
| VS635  | CTATAGTATGTATGTGG          | Forward     | 4195–4212 |
| VS836  | GTGAGCTAGAGAAAGAC          | Reverse     | 4905–4921 |
| VS837  | CAGGACCTGAGAGATGCTG        | Forward     | –2 to 16  |
| VS640  | CAGGCTTGATGTATGG           | Reverse     | 1129–1145 |
| VS664  | FAM-CACTGGAGTGCTCAAC-TAMRA | Forward     | 2077–2093 |
| VS665  | CAAGGTGATTGTAATGACGAACACT | Forward     | 1989–2012 |
| VS866  | CGATTGCGACTTACAGC          | Reverse     | 2134–2152 |

*Position on genome of calicivirus Bangladesh/289/2007 (GenBank accession no. JQ745645).