

- allele in this tumor actually carried a functionally important mutation outside of the region sequenced (exons 5 to 9); or (iii) some mutated p53 alleles behave differently than others, or differently depending on cell type.
22. That mutated p53 has a function is also suggested by a study [D. W. Wolf, N. Harris, V. Rotter, *Cell* 38, 119 (1984)] in which a presumably mutated p53 gene was transfected into Ab-MuLV-transformed murine leukemia cells that lacked endogenous p53 expression. This resulted in cells with more malignant behavior in tumorigenicity assays.
 23. J. R. Jenkins, P. Chumakov, C. Addison, H. W. Sturzbecher, A. Wade-Evans, *J. Virol.* 62, 3902 (1988).
 24. E. Harlow, L. V. Crawford, D. C. Pim, N. M. Williamson, *ibid.* 39, 861 (1981).
 25. J.-Y. Shew *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 6 (1990); P. L. Chen *et al.*, *Cell* 58, 1193 (1989).
 26. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
 27. We thank W. Hammerschmidt and B. Sugden for retroviral vector 477, and E. Harlow for monoclonal antibody pAb421 and p53 cDNA. This study was supported by grants from the National Eye Institute (EY05758 and EY00278) and National Cancer Institute (CA51495), and a fellowship to P.-L.C. from the Ministry of Education, Republic of China.

2 August 1990; accepted 11 October 1990

Norwalk Virus Genome Cloning and Characterization

XI JIANG, DAVID Y. GRAHAM, KENING WANG,* MARY K. ESTES†

Major epidemic outbreaks of acute gastroenteritis result from infections with Norwalk or Norwalk-like viruses. Virus purified from stool specimens of volunteers experimentally infected with Norwalk virus was used to construct recombinant complementary DNA (cDNA) and derive clones representing most of the viral genome. The specificity of the clones was shown by their hybridization with post- (but not pre-) infection stool samples from volunteers infected with Norwalk virus and with purified Norwalk virus. A correlation was observed between the appearance of hybridization signals in stool samples and clinical symptoms of acute gastroenteritis in volunteers. Hybridization assays between overlapping clones, restriction enzyme analyses, and partial nucleotide sequence information of the clones indicated that Norwalk virus contains a single-stranded RNA genome of positive sense, with a polyadenylated tail at the 3' end and a size of at least 7.5 kilobases. A consensus amino acid sequence motif typical of viral RNA-dependent RNA polymerases was identified in one of the Norwalk virus clones. The availability of Norwalk-specific cDNA and the new sequence information of the viral genome should permit the development of sensitive diagnostic assays and studies of the molecular biology of the virus.

ACUTE GASTROENTERITIS IS ONE OF the most common illnesses in the United States (1, 2), with a large number (up to 42% of outbreaks) of cases estimated to be caused by Norwalk or Norwalk-like viruses (3). Both water- and food-borne transmissions of Norwalk virus have been documented, and particularly large epidemic outbreaks of illness have occurred after consumption of contaminated shellfish including clams, oysters, and cockles (4-10).

Norwalk virus was discovered in 1973, but knowledge about the virus has remained limited because it has not been possible to propagate the virus in cell culture and suitable animal models have not been found (2). Human stool samples obtained from out-

breaks and from human volunteer studies have been the only source of virus. Moreover, the concentration of virus in stool is usually so low that virus detection by routine electron microscopy is not possible (11-13). Current methods of Norwalk virus detection include immune electron microscopy and other immunologic methods such as radioimmunoassays (RIAs) or biotin-avidin enzyme-linked immunosorbent assays (ELISAs), which use sera from humans in acute and convalescent phases. To date, because of either insufficient quantities or unusual properties of the viral antigen, no hyperimmune serum from animals has been successfully prepared. Preliminary biophysical characterization of virions has indicated that particles contain one polypeptide (14), but efforts to characterize the viral genome have failed. Therefore, these viruses have remained unclassified and difficult to study. Molecular cloning was seen as an approach to overcome these problems.

To permit better diagnosis and molecular characterization of Norwalk virus, we constructed a cDNA library derived from nu-

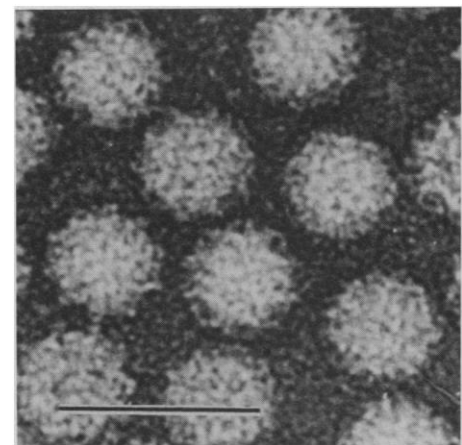


Fig. 1. Electron micrograph of Norwalk virus particles after CsCl gradient purification. Virus was visualized by staining with 1% ammonium molybdate. Scale bar, 50 nm.

cleic acid extracted from virions purified from stool samples obtained from volunteers. Norwalk virus (8FIIa) was administered orally to adult volunteers (15, 16). Stool samples collected after infection were examined for Norwalk virus by either RIAs or direct or immune electron microscopy. Stools from two patients containing the highest amount of Norwalk virus were used as the source of virus for cloning. Norwalk virus was purified by methods used previously for preparation of hepatitis A and rotaviruses from stool, with some modifications (17, 18). Basically, stools containing Norwalk virus were treated with Genetron (1,1,2-trichloro-1,2,2-trifluoroethane) to remove lipid and water-insoluble materials. Virus in the aqueous phase was then centrifuged through a 40% sucrose cushion, and the resultant pellets were suspended, sonicated, and loaded in a CsCl gradient for isopycnic centrifugation. Approximately 10^9 physical particles of virus of relatively high purity (Fig. 1) were obtained from 500 g of stool.

Nucleic acids were extracted from these purified viruses by proteinase K treatment of the samples followed by phenol-chloroform extraction and ethanol precipitation (19). Because the nature of the viral genome was unknown, the extracted nucleic acids were denatured with methylmercuric hydroxide before cDNA synthesis. Random primed cDNA was synthesized and a small amount of cDNA was obtained (19). Direct cloning of this small amount of cDNA was unsuccessful so a step of amplification of the DNA was performed by synthesizing more copies of the DNA with random primers and the Klenow fragment of DNA polymerase I before cloning. This cloning method was developed based on control experiments in which we amplified a known cDNA fragment of hepatitis A virus (20). The proce-

X. Jiang, K. Wang, M. K. Estes, Division of Molecular Virology, Baylor College of Medicine, Houston, TX 77030.

D. Y. Graham, Gastroenterology Section, Veterans Administration Medical Center, Houston, TX 77030.

*Present address: Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030.

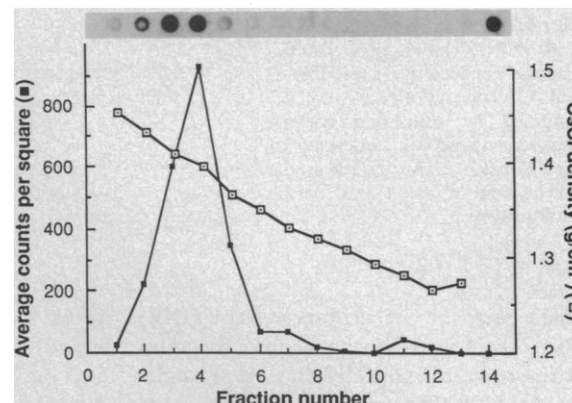
†To whom correspondence should be addressed.

cedure involved manual cycles of denaturation, addition of random primers and the Klenow fragment of DNA polymerase I, reannealing, and elongation (19). With this procedure, incorporation of labeled nucleotides into product was observed as the number of cycles of synthesis was increased. The number of cycles performed was limited to avoid the synthesis of an excess of smaller fragments. In the case of Norwalk cDNA, eight cycles of amplification were performed and approximately 2.5 µg of DNA were obtained. This amplified cDNA was cloned into the plasmid vector pUC13 by blunt-end ligation and a positive clone (pUCNV-953) was isolated.

To obtain the positive Norwalk virus clone, minipreparations of plasmid DNAs from white colonies of transformed bacteria containing potential inserts were screened by agarose gel electrophoresis (19). The larger clones were digested with restriction enzymes to release the insert DNA. The insert DNA was labeled with the Prime-a-Gene labeling system (Promega) and used to probe individually paired stool samples (before and after Norwalk infection) from two volunteers (Fig. 2A). One clone (pUCNV-953) reacted with early post- but not with preinfection stool samples from both volunteers.

The results were confirmed with six more paired stool samples (20). In additional tests, no signal was observed by dot blot hybridization of the clone with stools collected before acute illness, strong signals were observed only with stools from acute phase, and a weak signal was seen in a stool collected 4 days after illness from one of three volunteers (Fig. 2B). This result was

Fig. 3. Dot blot hybridization for detection of Norwalk viruses in a CsCl gradient. Partially purified viruses from stool samples were mixed with a CsCl solution (refractive index 1.368) and centrifuged at 35,000 rpm for 24 hours in an SW50.1 rotor (18). Aliquots of 50 µl from each fraction of the CsCl gradient were dotted onto a Zeta-bind filter. Duplicates of filters were made and hybridized with two ss-RNA probes made from pGEMNV-953 (top). The two probes were subsequently designated cRNA (positive hybridization with the viral nucleic acid) and vRNA [no hybridization with the viral nucleic acid (20)]. The number of virus particles (■) in each gradient fraction was determined by electron microscopy (bottom). Five squares from each electron microscopic grid were counted and the average number of viral particles per square was calculated.



consistent with previous RIA tests for viral antigen detection in which convalescent sera from volunteers with Norwalk diarrhea were used and with immune electron microscopy studies of the samples for viral particle examination. This result also agrees with the patterns of virus shedding in stool in the course of the disease (13). When RNA transcripts from the pUCNV-953 clone were hybridized with fractions of a CsCl gradient from the Norwalk virus purification scheme, an excellent correlation was observed between hybridization and viral particle counts as seen with the electron microscope (Fig. 3). The peaks of the hybridization signals and viral particle counts both were observed in fractions with a density of 1.39 to 1.40 g/cm³, which agrees with previous reports of the biophysical properties of Norwalk virus (21). Finally, the clone was tested by direct hybridization

with highly purified Norwalk virus particles after electrophoresis on an agarose gel. A single hybridization band was observed with Norwalk virus but not with hepatitis A virus and rotavirus (20). Sequence analysis of the pUCNV-953 cDNA showed that this fragment contains 512 bp and a potential open reading frame. No significant nucleotide or deduced amino acid sequence homology was found by comparison with other sequences in GenBank (22).

To obtain clones representing the entire genome of Norwalk virus, we constructed two additional cDNA libraries using pUC13 and Lambda-ZAPII vectors (Stratagene) from a second cDNA synthesis performed without the amplification step. More than 400 positive clones were identified by colony hybridization with probes made from the original clone (pUCNV-953) and from subsequent positive clones. A schematic of the first positive clone and four large overlapping clones identified from cross-hybridization and restriction mapping analyses is shown (Fig. 4). Overlapping regions of the clones were confirmed with clones isolated in the Lambda-ZAPII library. A polyadenylated [poly(A)] tail (~80 bp) was found at the 3' end of clone pUCNV-4145. At the 5' end, clone pUCNV-5101 was the longest of all the clones identified. More than 30 clones were identified from the libraries by means of a 600-bp fragment from the 5' end of pUCNV-5101 as probe. Sequence analyses have shown these clones contain either the same 5' end or they are one or two bases shorter than clone pUCNV-5101. When the viral RNA was hybridized with the probes, a band of approximately 7.5 kb was identified on a native agarose gel as previously described (20, 23). This result agreed with our estimated size of the viral genome based on restriction mapping of the clones.

The Norwalk virus cDNAs were used to

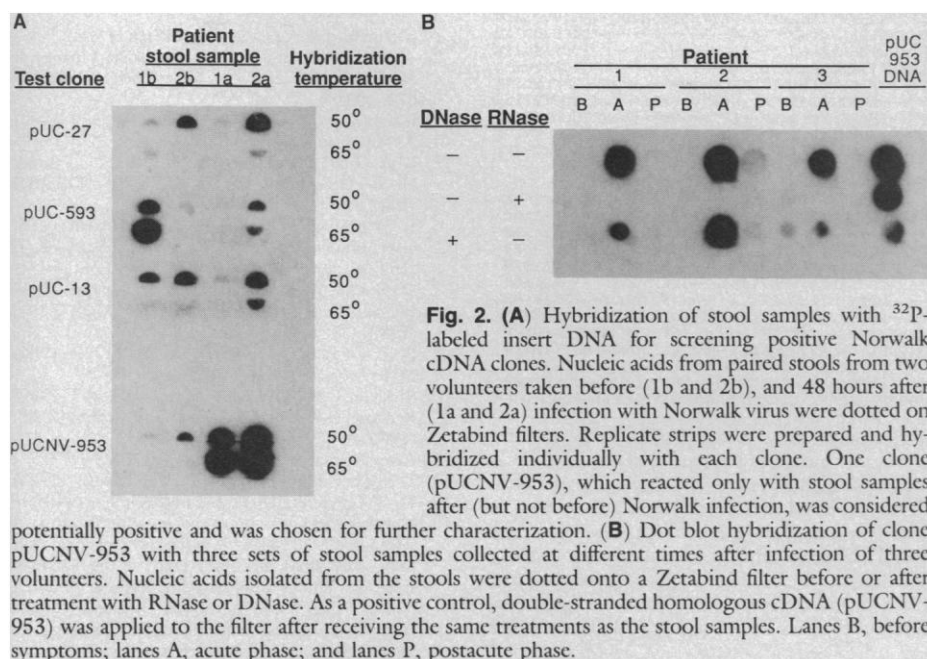
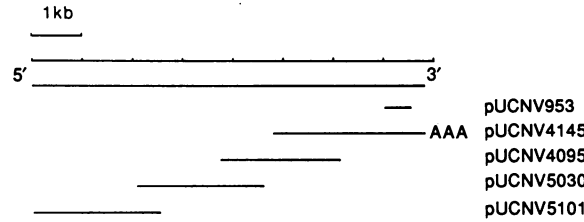


Fig. 4. Schematic diagram of Norwalk cDNA clones. pUCNV-953 was the first positive clone identified. Overlapping clones were determined by restriction enzyme analyses and partial sequencing of the clones. AAA indicates the poly(A) tail at the 3' end of the viral genome.



characterize the viral genome. The pUCNV-953 cDNA was subcloned into the transcription vector pGEM-3Zf(+) and single-stranded RNA (ssRNA) probes were generated by in vitro transcription with SP6 and T7 polymerases (Promega) (24). When two opposite-sense ssRNA probes were hybridized with the viral nucleic acid separately, only one strand reacted with the virus, indicating that the viral genome is single-stranded (20). A long open reading frame was found in one of the two strands of the DNA insert in pUCNV-953. The ssRNA

probe with the same sequence as this coding strand did not react with the viral nucleic acid but the complementary ssRNA probe did react in the hybridization tests. These results, together with the finding that ribonuclease (RNase) [but not deoxyribonuclease (DNase)] treatment of Norwalk viral nucleic acid removed hybridization signals (Fig. 2B), indicate that Norwalk virus contains a positive sense, ssRNA genome.

Partial DNA sequencing of the clones has shown that a putative open reading frame is present in the DNA sequence of each clone.

A long open reading frame with an amino acid sequence motif that is present in all positive-strand RNA viruses was identified in clone pUCNV-4095. This amino acid motif is believed to encode the RNA-dependent RNA polymerase. When the motifs were aligned, the Norwalk virus protein showed the highest similarity with the polymerases of viruses in the picornavirus family compared to other groups of viruses (Fig. 5). No other homology has been found between the available sequence and other sequences currently in the nucleic acid database.

Further studies with this cDNA should now allow specific diagnostic assays to be developed. Such assays could be based on detection of Norwalk virus nucleic acid (hybridization or polymerase chain reaction), or on the production of antibodies to proteins expressed from the cDNAs or to synthetic peptides made based on knowledge of the genome sequence. Sequencing of the cDNA clones will further our understanding of the genome organization of Norwalk virus and the relatedness of this virus with other Norwalk-like viruses, caliciviruses, and hepatitis E virus. The application of similar methods to other antigenic types among the Norwalk-like agents and possibly other non-characterized agents of gastroenteritis should also permit their molecular characterization (25). The availability of Norwalk-specific cDNA will facilitate studies of the molecular biology, immunology, and pathogenesis of the virus. The development of vaccines by recombinant DNA technology against acute epidemic gastroenteritis also may now be feasible.

A	
cat ttt gat gca gat tat aca gca tgg gac tca aca caa aat aga caa att atg aca gaa tcc ttc tcc att atg	
H F D A D Y T A W D S T Q N R Q I M T E S F S I M	
tcg cgc ctt acg gcc tca cca gaa ttg gcc gat gtt gtg gcc caa gat ttg cta gca cca tct gag att gat gta	
S R L T A S P E L A G V A Q D L L A C P S E M D V	
ggg gat tat gtc atc agg gtc aaa gag ggg ctg cca tct gga ttc cca tgt act tcc cag gtg aac agc ata aat	
G D Y V I R V K E G L P S G F P C T S Q V N S I N	
cac tgg ata att act ctc tgt gca ctg tct gag gcc act ggt tta tca cct gat gtg gtg caa tcc atg tca tat	
H W I I T L C A L S E A T G L S P D V V Q S M S Y	
ttc tca ttt tat ggt gat gat gag att gtg tca act gac ata gat ttt gac cca gcc cgc ctc act caa att ctc	
F S F Y G D D E I V S T D I D F D P A R L T Q I L	
aag gaa	
K E	
B	
NV	HFDAHYTAWDSTQNRQIMT-ESFSIMSR--LTASPEL-AEVVAQDLDLAPSEMDV-----GDYIVRVK--EG-LPS@FPCTSQV@
HEV	VFENDFSEFSTQNNFSLG-LECAIMEEC @MPQWLI-RLYHLIRSAWILQAPKESLR-GFWKKH-----S-KHS@EPGTL@W@
HCV	GFSY@TRCFSTVTESDIR-TEEAIVQCCD@DQ@RV-AIKSLTERIVYVGGPLTNSR--GENCGYRRCRAS-RAS@VLT@SCG@
HAV	GLDL@SFA@BASLSPFMIREAGRIMSELS-@TPSHFGTALINTIISKHLLYNCCHVCGS-----MPS@SPCTAL@
JE	MYAD@TAGW@TRITRTDLE-NEAKVLELLD@EHRMILARAI@ELTYRHKVVKVMRPA@E-GKTVMDVISREDQ@RGS@QVVTYAL@
POLIO	FA-F@YTGY@BASLS-PAWFEAL-KMVLEKIGFGDRVDY-IDYLNHSHHLYKNTKTCVKGG-----MPS@CSCT@SIF@
FMD	VWDV@YSAF@ANHCSDAMNIMFEEV@RTD@F@H@PNAEWILKTLVNT@HAYENK@RTV-EGG-----MPS@CSAT@SII@
EMC	VYDV@YSN@F@STH@SVAMFRLLAE@F@TPENG@F-DPL@TREY@LES@LAISTHAF@E@K@R@FLIT@G-----LPS@CAAT@SML@
SNBV	VLE@T@IAS@F@BKS-QDDAM@L@GLM@LEDL-@VD@P@LLD@LIE@CA@F@E@IS@THL@PT@G@TR@F@K@GA-M-----MKS@M@F@L@LFV@
TMV	VLE@L@ISK@Y@BKS-QNEF@CAVEY@E@WRRL-@FED@F@L@GEV@K@G--HRK@TLK-DITA-GYKTC-IWY-QRKS@DV@T@F@IG@
AMV	FKE@B@SK@F@BKS-QNELH@L@Q@R@F@L@K@YL-@IPNE@F@L@L@W@FNA--HRKS@RIS-DSKN-GVFFN-VDF-QRRT@DAL@Y@L@G@
BMV	FLE@B@L@SK@F@BKS-QGELH@L@Q@R@E@I@LLAL-@FPA@L@TN@W@S@D@F--HRDS@YLS-DPHAK@V@GMS--VSF-QRRT@DA@F@Y@F@G@
CpMV	VLCC@Y@S@F@B@L@L@SK@Q@VMDIASMINELCG@E-DQLKN@RRNLLMACCS@L@K@NTV@R@VECG-----IFS@F@P@T@VIV@
NV	SINH@IIT@L@C@A@S@E@AT@GLSPD-----VVQ@S@Y@F@S@FY@D@E@IVS---TDID@F@P-ARLT@QILK--E
HEV	TVNNMAVITHC-----YDFRD@F@Q@A@F@K@D@S@IVL-CSEY@R@Q@S@P@-A-A--VLI@AC
HCV	TLTCYIKARAACRAAGLQDCTMLVC-----@D@LVVI-CESAG@Q@E@-A-A--SL-RAF
HAV	SIINN@N@LYV@FSKI-----FGKSPVFF@C@Q@LK@Y@C@D@VLI@V--FSRDV@Q@B@L@D@LIG@Q@K@VDE@F
JE	TFTNIA@V@L@V@R@L@ME@AGVIG@P@H@L@EQ@L@P@R@K@T@IA@V@RT@W@LF@ENG@E@V@R@T@MAIS@D@C@V@K-----PLD@R@F@A@L@H@F@L-NAM
POLIO	SMINN@L@IRT@LL@K@Y@K@ID-----LDHLK@M@I@Y@D@D@V@AS---Y@P@H@D@S-----L@AQ@S
FMD	TILNNI@V@Y@L@R@R@H@Y@E@V@E-----LDY@T@M@I@S@D@D@IV@A---SDY@L@D@F-----E@AK@P-H
EMC	TMNNI@I@R@AG@L@Y@L@Y@K@N@F-----PDDV@K@V@S@Y@D@D@L@V@A--TNY@Q@L@D@F@K@V--RASL@K@T@G
SNBV	TVLN@V@I@AS@R@V@L@E-----ERL@K@S@R@C@A@F@I@D@D@N@I@H--GVVS@B@K@E@A@R@C@A@T@W@L-N
TMV	TVIIA@C@L@S@M@L-----PMEKI@K@G@A@F@C@D@D@S@L@Y-FPK@G@C@F@P@D@V@Q@H@S@N-L@W@N@F@E
AMV	TIVT@L@A@C@L@H@V@D@L@M-----DPNV@K@F@V@A@S@D@D@S@L@IG--TVE@L@P@R@Q@E@F-LFT@L@F@N@L@E
BMV	TLV@T@M@I@A@Y@A@D@L@S-----DCDCA@I@F@S@D@D@S@L@I@I--SKVK@P@V@L@D@T@M--FTS@L@F@N@M@E
CpMV	SIFNE@IL@R@Y@H@Y@K@L@M@R@E@Q@Q@A@P@E-----LMV@S@P@D@K@L@IG@L@V@T@Y@D@D@N@L@IS@V@N@V@V@T@Y@F@D@G@K@L-Q@S@I@A@Q@G@

Fig. 5. Norwalk virus encodes an RNA-dependent RNA polymerase sequence. The sequence of a portion of clone pUCNV-4095 (A) was derived by the dideoxy method (Sequenase, version 2.0, U.S. Biochemical Corporation). The deduced amino acid sequence of this portion (NV) is compared with consensus amino acid residues (B) thought to encode putative RNA-dependent RNA polymerases of hepatitis E virus (HEV), hepatitis C virus (HCV), hepatitis A virus (HAV), Japanese encephalitis virus (JE), poliovirus (POLIO), foot-and-mouth disease virus (FMD), encephalomyocarditis virus (EMC), Sindbis virus (SNBV), tobacco mosaic virus (TMV), alfalfa mosaic virus (AMV), bromo mosaic virus (BMV), and cowpea mosaic virus (CpMV). Sequences for viruses other than NV are from figure 3 of (26). No comparison is shown for caliciviruses because the deduced sequence information for caliciviruses is not available.

REFERENCES AND NOTES

1. J. Dingle *et al.*, *Am. J. Hyg.* **58**, 16 (1953).
2. A. Z. Kapikian and R. M. Chanock, in *Virology*, B. N. Fields, Ed. (Raven, New York, 1990), pp. 671-693.
3. J. Kaplan *et al.*, *Am. J. Pub. Health* **72**, 1329 (1982).
4. A. M. Murphy *et al.*, *Med. J. Aust.* **2**, 329 (1979).
5. R. A. Gunn, H. T. Janowski, S. Lieb, E. C. Prather, H. B. Greenberg, *Am. J. Epidemiol.* **115**, 348 (1982).
6. R. Wilson, L. J. Anderson, R. C. Holman, G. W. Gary, H. B. Greenberg, *Am. J. Pub. Health* **72**, 72 (1982).
7. O. N. Gill, W. D. Cubitt, D. A. McSwiggan, B. M. Wanney, C. L. R. Bartlett, *Br. Med. J.* **287**, 1532 (1983).
8. H. L. DuPont, *N. Engl. J. Med.* **314**, 707 (1986).
9. D. L. Morse *et al.*, *ibid.*, p. 678.
10. S. Sekine *et al.*, *Microbiol. Immunol.* **33**, 207 (1989).
11. R. Dolin *et al.*, *Proc. Soc. Exp. Biol. Med.* **140**, 578 (1972).
12. A. Z. Kapikian *et al.*, *J. Virol.* **10**, 1075 (1972).
13. T. S. Thornhill *et al.*, *J. Infect. Dis.* **132**, 28 (1975).
14. H. B. Greenberg *et al.*, *J. Virol.* **37**, 944 (1981).
15. Norwalk virus was produced by administration of safety-tested Norwalk virus (8FIIa) to adult volunteers. The virus inoculum used in the volunteer study was kindly supplied by A. Kapikian (Laboratory of Allergy and Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, MD). This virus originated from an out-

- break of acute gastroenteritis in Norwalk, Ohio. Two milliliters of a 1 to 100 dilution of 8FIIa in tris-buffered saline (TBS) was administered orally to each individual with 80 ml of Milli-Q water (Millipore). Sodium bicarbonate solution was taken by each person 2 min before and 5 min after virus administration. The volunteer studies were approved by the Institutional Review Board for Human Research at Baylor College of Medicine and at the Clinical Research Unit. The virus was administered to the volunteers in the General Clinical Research Center where the volunteers were hospitalized and under medical care for 4 days. All stools were collected and kept at -70°C for later use.
16. R. Dolin *et al.*, *J. Infect. Dis.* **123**, 307 (1971).
 17. X. Jiang, M. K. Estes, T. G. Metcalf, J. L. Melnick, *Appl. Environ. Microbiol.* **52**, 711 (1986).
 18. Norwalk virus from stool samples was purified by the following procedure. A 10% solution of stool samples in TBS was first clarified by low-speed centrifugation at 3000 rpm for 15 min. The resultant supernate was then extracted two or three times with Genetron in the presence of 0.5% Zwittergent 3-14 detergent (Calbiochem). Virus in the aqueous phase was concentrated by pelleting at 36,000 rpm for 90 min through a 40% sucrose cushion in a 50.2 Ti rotor (Beckman Instruments). The pellets were suspended in TBS, sonicated, and mixed with CsCl solution (refractive index 1.368) in TBS and centrifuged at 35,000 rpm for 24 hours in a SW50.1 rotor (Beckman). The CsCl gradient was fractionated by bottom puncture and each fraction was monitored for virus by examination with an electron microscope. The peak fractions containing Norwalk virus were pooled, and CsCl in the samples was diluted with TBS and removed by centrifugation at 35,000 rpm for 1 hour in a SW50.1 rotor. The purified virus was stored at -70°C .
 19. The cDNA libraries were prepared from nucleic acid extracted from purified Norwalk virus. In order to extract nucleic acids, purified Norwalk virus from CsCl gradients was first treated with proteinase K (400 $\mu\text{g}/\text{ml}$) in $1\times$ proteinase K buffer [0.1 M tris-HCl (pH 7.5), 12.5 mM EDTA, 0.15 M NaCl, 1% (w/v) SDS] at 37°C for 30 min. The samples were then extracted once with phenol-chloroform and once with chloroform. Nucleic acid in the aqueous phase was concentrated by precipitation with 2.5 volumes of ethanol in the presence of 0.2 M sodium acetate followed by centrifugation for 15 min in a microcentrifuge. The nucleic acid was then denatured with 10 mM CH_3HgOH , and cDNA was synthesized (from both DNA and RNA) by means of the cDNA synthesis kit with the supplied random hexanucleotide primer (cDNA Synthesis System Plus, Amersham). After synthesis of the second strand the reaction mixture was extracted once with phenol-chloroform and once with chloroform followed by ethanol precipitation. Amplification of DNA was performed with buffers as described in the Prime-a-Gene Labeling System (Promega). Eight cycles of denaturation (100°C for 2 min), reannealing (2 min of cooling to room temperature), and elongation (room temperature for 30 min) were performed after addition of the Klenow fragment of DNA polymerase I (Promega). A cDNA library was constructed in pUC13 with blunt-end ligation into the Sma I site. White colonies from transformed DH5 α (Bethesda Research Laboratories) cells were picked, and both a master plate and minipreps of plasmid DNA were prepared for each clone. Clones containing inserts were identified after electrophoresis of the plasmid DNA in an agarose gel. The insert DNA was excised by digestion with restriction enzymes, and insert DNA that was resolved in low-melting temperature agarose gels was cut out and labeled with ^{32}P by random priming. Nucleic acids extracted from paired stool samples (before and after Norwalk infection) from two volunteers (543 and 544) were dotted onto Zetabind filters (AMF, Cuno, Meriden, CT). Replicate filter strips were prepared and hybridized with each labeled insert probe individually at 65°C without formamide. Potential positive clones were judged by their different reactions with the pre- and postinfection stools. Clones that reacted with post- (but not pre-) infection stools of volunteers were considered positive, and these clones on the master plates were characterized further.
 20. X. Jiang and M. K. Estes, unpublished data.
 21. A. Z. Kapikian, J. L. Gerin, R. G. Wyatt, T. S. Thornhill, R. M. Chanock, *Proc. Soc. Exp. Biol. Med.* **142**, 874 (1974).
 22. GenBank sequences were checked with the Molecular Biology Information Resource, Eugene Software, Baylor College of Medicine.
 23. X. Jiang, M. K. Estes, T. G. Metcalf, *J. Clin. Microbiol.* **27**, 874 (1989).
 24. ———, *Appl. Environ. Microbiol.* **53**, 2487 (1987).
 25. The knowledge that the Norwalk virus genome is single-stranded RNA has permitted the synthesis of additional cDNA clones from nucleic acid that was extracted from small amounts (~ 1 g) of stool samples. These preliminary results indicate that it will be possible to obtain cloned DNAs from many clinical specimens.
 26. G. R. Reyes *et al.*, *Science* **247**, 1335 (1990).
 27. We thank A. Opekun for help with the volunteer studies; T. Tanaka for establishing and performing the RIA assays; E. Calomeni for the electron microscopy; and J. Butel, L. Donehower, and J. L. Melnick for helpful discussions of the manuscript. Supported by Food and Drug Administration grant 223-88-2182. The General Clinical Research Center is supported by the Division of Research Resources of the NIH under grant RR 00350.

10 July 1990; accepted 22 October 1990

Homologous Recombination and Stable Transfection in the Parasitic Protozoan *Trypanosoma brucei*

MARY GWO-SHU LEE AND LEX H. T. VAN DER PLOEG*

Development of methods for the manipulation of the genomes of parasitic protozoa will lead to enhanced understanding of parasite biology and host-parasite relationships. Efficient gene transfer and targeted integration by homologous recombination were achieved in the parasitic protozoan *Trypanosoma brucei*, the causative agent of sleeping sickness. An expression vector with the neomycin phosphotransferase gene (*neo*), under the control of a procyclic acidic repetitive protein (PARP) gene promoter, was targeted into an intergenic region in $\beta\alpha$ -tubulin-gene tandem array. Sixteen copies of *neo* were found in a tandem array in one of the transfectants where the PARP promoter controlled α -amanitin-resistant transcription of *neo*, whereas transcription of tubulin genes remained α -amanitin-sensitive.

TRYPANOSOMES ARE MEMBERS OF A large group of unicellular eukaryotic flagellates, causative agents of a wide spectrum of diseases in humans and livestock. The study of this group of protozoa has revealed many aspects of host-parasite interaction and novel biochemical control mechanisms, including precursor mRNA maturation by transsplicing, polycistronic transcription of protein coding genes, and RNA editing (1).

Transient expression of foreign genes has been achieved in several trypanosome species (2–6). Stable transfection of the protozoan *Leishmania* was also reported on the basis of the expression of *neo* from extrachromosomal elements. We had established transient DNA transfection of the protozoan *Trypanosoma brucei* by means of a plasmid construct in which the chloramphenicol acetyl transferase (CAT) gene was expressed from the promoter of one of the PARP genes (plasmid BNsp-CAT) (6). Positive selection was needed to identify trypanosomes in which exogenous DNA had been incorporated into chromosomes by targeted

homologous recombination; therefore, the PARP promoter was used to direct expression of *neo* in the plasmid BNsp-Neo-T in *T. brucei* (Fig. 1A). To avoid disruption of essential trypanosome genes, we targeted the BNsp-Neo-T plasmid into the intergenic region of one of the tandem arrays of $\beta\alpha$ -tubulin genes. To mediate the homologous recombination at the tubulin intergenic regions, we inserted a 639-bp fragment derived from the intergenic region of the $\beta\alpha$ -tubulin genes (a blunt-ended Eco RI–Bgl II intergenic region fragment) (Fig. 1B) into the region immediately downstream of *neo*.

In *Saccharomyces cerevisiae*, the introduction of a double-stranded break in the region of homology between the transfected DNA and the target sequences drastically increases the efficiency of homologous recombination (7). Procyclic *T. brucei*, the form of the parasite found in its insect vector, the tsetse fly, was therefore transfected by electroporation with a BNsp-Neo-T plasmid, that had been linearized at a unique Mlu I site; this restriction site is located centrally in the 639-bp $\beta\alpha$ -tubulin intergenic region (see Fig. 1 legend). When the linearized plasmid was used, G418-resistant trypanosomes could be observed 2 weeks after transfection (in four individual experiments), whereas cells transfected with circular plasmids never developed G418 re-

M. G.-S. Lee, Division of Tropical Medicine, School of Public Health, Columbia University, New York, NY 10032.

L. H. T. Van der Ploeg, Department of Genetics and Development, Columbia University, New York, NY 10032.

*To whom correspondence should be addressed.