

Prediction of Human Rotavirus Serotype by Nucleotide Sequence Analysis of the VP7 Protein Gene

KIM Y. GREEN,^{1*} JOHNNA F. SEARS,¹ KOKI TANIGUCHI,² KAREN MIDTHUN,¹ YASUTAKA HOSHINO,¹
MARIO GORZIGLIA,¹ KAZUO NISHIKAWA,¹ SHOZO URASAWA,² ALBERT Z. KAPIKIAN,¹
ROBERT M. CHANOCK,¹ AND JORGE FLORES¹

Laboratory of Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892,¹ and Department of Hygiene and Epidemiology, Sapporo Medical College, Sapporo, Japan²

Received 12 November 1987/Accepted 8 February 1988

Human rotavirus field isolates were characterized by direct sequence analysis of the gene encoding the serotype-specific major neutralization protein (VP7). Single-stranded RNA transcripts were prepared from virus particles obtained directly from stool specimens or after two or three passages in MA-104 cells. Two regions of the gene (nucleotides 307 through 351 and 670 through 711) which had previously been shown to contain regions of sequence divergence among rotavirus serotypes were sequenced by the dideoxynucleotide method with two different synthetic oligonucleotide primers. The resulting nucleotide sequences were compared with the corresponding sequences from rotaviruses of known serotype (serotype 1, 2, 3, or 4). A total of 25 field isolates and 10 laboratory strains examined by this method exhibited marked sequence identity in both areas of the gene with the corresponding regions of 1 of the 4 reference strains. In addition, the predicted serotype from the sequence analysis correlated in each case with the serotype determined when the rotaviruses were examined by plaque reduction neutralization or reactivity with serotype-specific monoclonal antibodies. These data suggest that as a result of the high degree of sequence conservation observed among rotaviruses of the same serotype, it is possible to predict the serotype of a rotavirus isolate by direct sequence analysis of its VP7 gene.

Four serotypes were originally identified among the many rotaviruses recovered from humans (25), and two new human rotavirus serotype candidates have been described recently (1, 5, 18). Rotavirus serotypes have been defined using plaque reduction neutralization (PRN) in vitro (16, 24), tube neutralization (25), fluorescent focus neutralization (3) and, recently, enzyme-linked immunosorbent assay (ELISA) using serotype-specific monoclonal antibodies (7, 15, 21, 22). The major serotype-specific neutralization antigens have been demonstrated (13, 17) to be located on the VP7 outer capsid protein which is encoded by gene 8 or 9. In a recent report from this laboratory, the close relatedness of the gene sequences encoding the VP7 of rotavirus strains belonging to the same serotype was demonstrated by RNA-RNA hybridization experiments (19). In these experiments, denatured double-stranded RNAs from field strains were hybridized to ³²P-labeled single-stranded RNAs (ssRNAs) transcribed in vitro from single-gene substitution rotavirus reassortants (in which the VP7 gene was derived from human strains belonging to each of the four epidemiologically important rotavirus serotypes), which allowed the deduction of the rotavirus serotype specificity. More recently, we reported that at least six specific regions of the VP7 protein that are divergent in amino acid sequence among different serotypes are highly conserved among rotaviruses of the same serotype (12). On the basis of these observations, we proposed that it may be possible to predict the serotype of a rotavirus field isolate by comparison of its VP7 gene nucleotide sequence with that of a reference virus from each serotype (serotypes 1 through 4).

In the current study, a method was developed for obtain-

ing and sequencing ssRNA transcripts from rotaviruses partially purified either directly from stool material or after passage in tissue culture. The places and dates of collection of 25 human rotavirus field isolates and laboratory strains used in this study are shown in Table 1. Strains B, G, and C were obtained from I. Sarov (Ben Gurion University of the Negev, Beer-Sheva, Israel), and strains 3819, 5083, 2743, and 2899 were obtained from T. Naguib (Egyptian Organization for Biological Products and Vaccines, Agouza-Guiza, Egypt). Strain P118 was obtained from R. Dolin (University of Rochester, Rochester, New York), strain McN13 was obtained from R. Bishop (Royal Children's Hospital, Melbourne, Australia), and strains Hoshi and Hosokawa were obtained from Y. Inaba (National Institute of Animal Health, Ibaraki, Japan). Stool specimens positive for rotavirus by the ELISA were made into a 10% suspension in phosphate-buffered saline. To grow virus in tissue culture, portions of some stool suspensions were treated with trypsin (5 µg/ml) for 30 to 60 min at 37°C before absorption onto MA-104 cell monolayers. Viruses were partially purified either from stool suspensions (2 ml) or from tissue culture harvests by fluorocarbon extraction, followed by centrifugation through a 30% sucrose cushion as described previously (9). ssRNAs were produced by transcription from partially purified virus, extracted once with phenol-chloroform, and precipitated with 3 volumes of ethanol. Additional purification of virus was not required to obtain sufficient ssRNAs for sequence analysis. Dideoxynucleotide sequencing of rotavirus RNA was performed as described previously (12) with two synthetic oligonucleotides, 5'CCATTGGATTACACAACCAT TC3' and 5'GCTACGTTTCTCTTGGTCC3', targeted to nucleotides 531 through 552 and 801 through 819 of the transcript RNA encoding VP7, respectively. These primers were deduced from published sequences of the VP7 gene of

* Corresponding author.

TABLE 1. Comparison of the serotype designation of 35 human rotavirus strains as determined by three methods

Strain	Country of origin	Year collected	Serotype determined by following method:		
			PRN	MAB ^a	Sequence ^b
D	United States	1974	1	1	1
HN289	Venezuela	1981		1	1
HN256	Venezuela	1981		1	1
HN307	Venezuela	1981		1	1
M37	Venezuela	1982	1	1	1
B	Israel	1983		1	1
G	Israel	1983		1	1
P118	United States	1985		1	1
DS-1	United States	1976	2	2	2
HN126	Venezuela	1981	2	2	2
2743	Egypt	1982		2	2
2899	Egypt	1982		2	2
C	Israel	1983		2	2
V197	Venezuela	1985		2	2
V206	Venezuela	1986		2	2
P	United States	1974	3	3	3
McN13	Australia	1980	3	3	3
HN330	Venezuela	1981		3	3
HN257	Venezuela	1981		3	3
V251	Venezuela	1985		3	3
V460	Venezuela	1986		3	3
V101	Venezuela	1986		3	3
V52	Venezuela	1986		3	3
V96	Venezuela	1986		3	3
ST3	England	1975	4	4	4
HN11	Venezuela	1980		4	4
HN14	Venezuela	1980		4	4
HN5	Venezuela	1980		4	4
HN19	Venezuela	1980		4	4
VA70	Italy	1981	4	4	4
HN98	Venezuela	1981		4	4
3819	Egypt	1983		4	4
5083	Egypt	1983		4	4
Hosokawa	Japan	1983	4	4	4
Hochi	Japan	1983	4	4	4

^a Monoclonal antibody ELISA performed as described by Taniguchi et al. (22).

^b Nucleotide sequence analysis.

bovine rotavirus strain Nebraska calf diarrhea virus (11) and are complementary to two regions in the VP7 gene which appear to be conserved in all rotaviruses examined thus far. These two priming sites are each downstream from a region of the gene that codes for a discrete cluster of amino acids in VP7 which are conserved within a serotype but are divergent among serotypes. The two regions sequenced, nucleotides 307 through 351 and 670 through 711, correspond to amino acids 87 through 101 and 208 through 221, respectively.

The nucleotide and deduced amino acid sequences obtained for each field strain in the two divergent regions were compared with the sequences of reference strains D (serotype 1), DS-1 (serotype 2), P (serotype 3), and ST3 (serotype 4). In Fig. 1, the sequence of each field strain is shown in comparison with the reference strain to which it showed the greatest identity. When there was marked conservation of amino acid sequence ($\geq 85\%$) between a field strain and the reference strain in these two regions, the field strain was assigned the serotype of the reference strain. In contrast, amino acid similarity among different serotypes in these two regions ranged from 33 to 62%. No more than two amino acid substitutions occurred per region among strains of the same predicted serotype.

All 35 strains in this study were tested in a monoclonal antibody serotyping ELISA, and 10 were serotyped by PRN. A comparison of these results with the deduced serotype of each virus determined by sequence analysis showed complete agreement (Table 1).

Early after the discovery of rotaviruses in 1973, the adaptation of a series of animal and human strains to tissue culture and their comparison by neutralization assays permitted the designation of several serotypes (23). However, serotype identification has been difficult until recently. Serotype analysis of rotaviruses will undoubtedly be an essential component in the evaluation of rotavirus vaccine efficacy, because recent studies indicate the importance of serotype-specific immunity in humans. For example, phase II vaccine trials with a monovalent (serotype 3) vaccine, containing the rhesus rotavirus strain, have been conducted in several areas of the world, and the preliminary data indicate that the vaccine was successful in inducing significant resistance to severe diarrhea in an area where the prevalent serotype was the same as that of the vaccine strain (10) but was not successful in places where other serotypes were prevalent (17a).

Two regions of the VP7 gene shown to be conserved within a serotype but divergent among serotypes were sequenced from ssRNA transcripts of 25 rotavirus strains. These strains had been collected in widely scattered localities over a 12-year period. In every case, it was possible to correctly predict the serotype by comparing the sequence in two selected regions of the VP7 gene to those of strains of known serotype. These two regions of the gene were analyzed because they contained the sequences for amino acids (residues 94, 96, and 211) which had been previously assigned to two VP7 serotype-specific neutralization epitopes on the basis of sequence analysis of neutralization-resistant mutants selected with neutralizing monoclonal antibodies (8). Comparison of the nucleotide sequence of the isolates with the reference sequence to which it matched most closely showed that nucleotide substitutions did occur but that most substitutions were in position 3 of the codon and did not result in an amino acid substitution. Some nucleotide substitutions which resulted in an amino acid substitution

FIG. 1. Sequence analyses of 25 rotavirus field isolates of unknown serotype and 10 laboratory strains of known serotype. Nucleotide sequences from two areas of the VP7 gene transcript RNA (nucleotides 307 through 351 encoding amino acids 87 through 101 and nucleotides 670 through 711 encoding amino acids 208 through 221) which were divergent in sequence among serotypes were determined and compared with reference strains D (serotype 1), DS-1 (DS1) (serotype 2), P (serotype 3), or ST3 (serotype 4). The deduced amino acid sequences for each reference strain in both regions are enclosed in open boxes above the nucleotide sequence. The nucleotide sequences determined from each strain analyzed are shown beneath the reference strain to which the highest degree of similarity was detected. Nucleotide substitutions within strains corresponding to the same serotype are underlined, and amino acid substitutions are enclosed in open boxes. Each strain analyzed matched closely with only one of the reference strains in both divergent regions. The serotype of the reference strain to which the field isolate correlated was assigned as the predicted serotype of the field isolate. nt, Nucleotide; aa, amino acid.

were located at positions to which amino acid substitutions in monoclonal antibody-selected variants had been mapped. In addition, sequence analysis was performed in these two divergent regions of the gene approximately 400 bases apart to determine whether field isolates maintained serotype-specific sequences throughout the gene or whether recombination of divergent regions occurred. Evidence for chimeric VP7 sequences was not obtained. Each of the 25 isolates possessed both of the variable region sequences of the reference rotavirus it most closely resembled. In addition, the isolates matched closely with the reference virus in other regions of the VP7 gene (including nucleotides 163 through 198 and 745 through 774) shown to be variable among different serotypes (data not shown). However, since the neutralization specificity of VP7 appears to be dependent on protein conformation (2, 14), it is possible that strains may exist which contain mixed divergent regions which yield a new serotype specificity.

Taniguchi et al. (22), using monoclonal antibodies specific for the VP7 of rotaviruses from serotypes 1 through 4, identified the serotype of rotaviruses in 38 out of 57 rotavirus-positive stool specimens examined. Of the 19 specimens studied that were not typed, 10 of these appeared to contain two viruses, 4 were nonreactive with the VP3 common antigen control antibody, and 5 reacted with the VP3 common antigen control antibody but did not react with any of the serotype-specific monoclonal antibodies. These latter viruses could represent either new serotypes or naturally occurring variants. Coulson (6) has described the existence of serotype 1 rotaviruses with various reactivity patterns with three different serotype 1-specific monoclonal antibodies and has proposed that these variants be designated monotypes. It is apparent from sequence data presented here and previously that although there is a high degree of sequence conservation within rotaviruses of the same serotype, there are sites in the VP7 gene where genetic variation occurs. The emergence of variants from the genetic pool of rotaviruses could explain the different reactivities of some rotaviruses with certain monoclonal antibodies, i.e., monotypes. It is not known whether existing host antibody pressures play a role in the emergence of these variants. However, it is evident that care must be used in the selection of monoclonal antibodies for a serotyping assay. For example, serotype 1-specific monoclonal antibody 2C9, described by Shaw et al. (21), neutralizes serotype 1 strain D but does not neutralize serotype 1 strain M37 by PRN. Strain M37 contains an amino acid substitution (Asn → Ser) at residue 94 when compared with strain D and thus could represent a naturally occurring variant which would not be recognized by monoclonal antibody 2C9, which maps to residue 94.

The serotype of each test strain predicted by the sequence analysis in this study correlated in each case with the serotype of the strain determined by a monoclonal antibody serotyping ELISA. Furthermore, PRN assays performed with a number of these specimens also yielded results indicating concordance with sequence analyses. Although sequence analysis may not be a practical method for routine typing in most laboratories, sequencing of VP7 genes has several features of practical interest. (i) If sufficient virus is present in the stool, rotavirus particles can be partially purified and transcribed from the stools without adaptation of the virus to growth in tissue culture and, if not, two passages in tissue culture will provide sufficient virus. (ii) Comparison with the reference strains will reveal at the molecular level whether the strain is a variant or a new serotype. Sequence analysis has been used recently in

another virus system for a similar purpose. Subgrouping of foot-and-mouth disease virus by sequence analysis revealed that several outbreaks of the disease in Europe were likely caused by a subtype of the virus identical to the vaccine strain (4). Genotyping of poliovirus is now being used for epidemiologic studies (20). In the latter case, direct sequence analysis was performed on an area of the genome relevant to serotype. Similar studies with rotaviruses are in progress in our laboratory using sequence analysis as an epidemiologic tool.

LITERATURE CITED

1. Albert, M. J., L. E. Unicomb, and R. F. Bishop. 1987. Cultivation and characterization of human rotaviruses with "super short" RNA patterns. *J. Clin. Microbiol.* **25**:183-185.
2. Bastardo, J. W., J. L. McKimm-Breschkin, S. Sonza, L. D. Mercer, and I. H. Holmes. 1981. Preparation and characterization of antisera to electrophoretically purified SA11 virus polypeptides. *Infect. Immun.* **34**:641-647.
3. Beards, G. M., H. N. Pilford, M. E. Thouless, and T. H. Flewett. 1980. Rotavirus serotypes by serum neutralization. *J. Med. Virol.* **5**:231-237.
4. Beck, E., and K. Strohmaier. 1987. Subtyping of European foot-and-mouth disease virus strains by nucleotide sequence determination. *J. Virol.* **61**:1621-1629.
5. Clark, H. F., Y. Hoshino, L. M. Bell, J. Groff, G. Hess, P. Bachman, and P. A. Offit. 1987. Rotavirus isolate WI61 representing a presumptive new human serotype. *J. Clin. Microbiol.* **25**:1757-1762.
6. Coulson, B. S. 1987. Variation in neutralization epitopes of human rotaviruses in relation to genomic RNA polymorphism. *Virology* **159**:209-216.
7. Coulson, B. S., L. E. Unicomb, G. A. Pitson, and R. F. Bishop. 1987. Simple and specific enzyme immunoassay using monoclonal antibodies for serotyping human rotaviruses. *J. Clin. Microbiol.* **25**:509-515.
8. Dyall-Smith, M. L., I. Lazdins, G. W. Tregear, and I. H. Holmes. 1986. Location of the major antigenic sites involved in rotavirus serotype-specific neutralization. *Proc. Natl. Acad. Sci. USA* **83**:3465-3468.
9. Flores, J., J. Myslinski, A. R. Kalica, H. B. Greenberg, R. G. Wyatt, A. Z. Kapikian, and R. M. Chanock. 1982. In vitro transcription of two human rotaviruses. *J. Virol.* **43**:1032-1037.
10. Flores, J., I. Perez-Schael, M. Gonzalez, D. Garcia, M. Perez, N. Daoud, W. Cunto, R. M. Chanock, and A. Z. Kapikian. 1987. Protection against severe rotavirus diarrhea by rhesus rotavirus vaccine in Venezuelan infants. *Lancet* **i**:882-884.
11. Glass, R. I., J. Keith, O. Nakagomi, T. Nakagomi, J. Askaa, A. Z. Kapikian, R. M. Chanock, and J. Flores. 1985. Nucleotide sequence of the structural glycoprotein VP7 gene of Nebraska calf diarrhea virus rotavirus: comparison with homologous genes from four strains of human and animal rotaviruses. *Virology* **141**:292-298.
12. Green, K. Y., K. Midthun, M. Gorziglia, Y. Hoshino, A. Z. Kapikian, R. M. Chanock, and J. Flores. 1987. Comparison of the amino acid sequences of the major neutralization protein of four human rotavirus serotypes. *Virology* **161**:153-159.
13. Greenberg, H. B., J. Flores, A. R. Kalica, R. G. Wyatt, and R. Jones. 1983. Gene coding assignments for growth restriction, neutralization and subgroup specificities of the W and DS-1 strains of human rotavirus. *J. Gen. Virol.* **64**:313-320.
14. Gunn, P. R., F. Sato, K. F. H. Powell, A. R. Bellamy, J. R. Napier, D. R. K. Harding, W. S. Hancock, L. J. Siegman, and G. W. Both. 1985. Rotavirus neutralizing protein VP7: antigenic determinants investigated by sequence analysis and peptide synthesis. *J. Virol.* **54**:791-797.
15. Heath, R., C. Birch, and I. Gust. 1986. Antigenic analysis of rotavirus isolates using monoclonal antibodies specific for human serotypes 1, 2, 3, and 4, and SA11. *J. Gen. Virol.* **67**:2455-2466.
16. Hoshino, Y., R. G. Wyatt, H. B. Greenberg, J. Flores, and A. Z. Kapikian. 1984. Serotype similarity and diversity of rotaviruses

- of mammalian and avian origin as studied by plaque-reduction neutralization. *J. Infect. Dis.* **149**:694-702.
17. Kalica, A. R., H. B. Greenberg, R. G. Wyatt, J. Flores, M. M. Sereno, A. Z. Kapikian, and R. M. Chanock. 1981. Genes of human (strain Wa) and bovine (strain UK) rotavirus that code for neutralization and subgroup antigens. *Virology* **112**:385-390.
 - 17a. Kapikian, A. Z., J. Flores, K. Midthun, Y. Hoshino, K. Green, M. Gorziglia, K. Taniguchi, K. Nishikawa, R. Chanock, L. Potash, I. Perez-Schael, R. Dolin, C. Christy, M. Santosham, N. Halsey, M. Clements, J. Sears, R. Black, M. Levine, G. Lonsky, M. Rennels, L. Gothefors, G. Wadell, R. Glass, T. Vesikari, E. Anderson, R. Belshe, P. Wright, and S. Urasawa. 1988. Development of a rotavirus vaccine by a "Jennerian" and a modified "Jennerian" approach, p. 151-159. *In* R. Chanock, F. Brown, R. Lerner, and H. Ginsberg (ed.), *Vaccines '88*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 18. Matsuno, S., A. Hasegawa, A. Mukoyama, and S. Inouye. 1985. A candidate for a new serotype of human rotavirus. *J. Virol.* **54**:623-624.
 19. Midthun, K., J. Flores, K. Taniguchi, S. Urasawa, A. Z. Kapikian, and R. M. Chanock. 1987. Genetic relatedness among human rotavirus genes coding for VP7, a major neutralization protein, and its application to serotype identification. *J. Clin. Microbiol.* **25**:1269-1274.
 20. Rico-Hesse, R., M. A. Pallansch, B. K. Nottay, and O. M. Kew. 1987. Geographic distribution of wild poliovirus type 1 genotypes. *Virology* **160**:311-322.
 21. Shaw, R. D., D. L. Stoner-Ma, M. K. Estes, and H. B. Greenberg. 1985. Specific enzyme-linked immunoassay for rotavirus serotypes 1 and 3. *J. Clin. Microbiol.* **22**:286-291.
 22. Taniguchi, K., T. Urasawa, Y. Morita, H. B. Greenberg, and S. Urasawa. 1987. Direct serotyping of human rotavirus in stools by an enzyme-linked immunosorbent assay using serotypes 1-, 2-, 3-, and 4- specific monoclonal antibodies to VP7. *J. Infect. Dis.* **155**:1159-1166.
 23. Thouless, M. E., A. S. Bryden, and T. H. Flewett. 1978. Serotypes of human rotavirus. *Lancet* **i**:39.
 24. Wyatt, R. G., H. B. Greenberg, W. D. James, A. L. Pittman, A. R. Kalica, J. Flores, R. M. Chanock, and A. Z. Kapikian. 1982. Definition of human rotavirus serotypes by plaque reduction assay. *Infect. Immun.* **37**:110-115.
 25. Wyatt, R. G., H. D. James, Jr., A. L. Pittman, Y. Hoshino, H. B. Greenberg, A. R. Kalica, J. Flores, and A. Z. Kapikian. 1983. Direct isolation in cell culture of human rotaviruses and their characterization into four serotypes. *J. Clin. Microbiol.* **18**: 310-317.