

## Comparison of group B rotavirus genes 9 and 11

Martin Petric, Kumudini Mayur, Steven Vonderfecht and Joseph J. Eiden\*

*Division of Infectious Disease, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, U.S.A.*

Group B rotaviruses (GBRs) were recognized recently as causative agents of gastroenteritis. Investigations into the relatedness of various heterologous GBR strains have been hindered by the difficulty of growing these viruses in cell culture. Viral RNA extracted from experimentally infected rats was used to prepare cDNA clones. From these, the nucleotide sequences of genes 9 and 11 of the IDIR strain of GBR were determined and compared with the corresponding sequences of the human ADRV strain of GBR. IDIR gene 11 is 643 bp

in length with a single open reading frame (ORF) encoding 174 amino acids; IDIR gene 9 is 804 bp in length with a single ORF encoding 246 amino acids. Comparison of the IDIR sequences with those of ADRV showed that nucleotide sequence similarity was 60.6% and 71.9% for genes 9 and 11, respectively. The deduced amino acid sequence similarity was 51.2% for the gene 9 and 66.5% for the gene 11 product. This sequence diversity indicates that GBRs are more distantly related than strains of group A rotavirus.

Rotaviruses are currently classified into antigenically distinct groups, A to F (Bridger, 1987). Group B rotaviruses (GBRs) have been detected in a number of mammalian species including humans, rats, lambs, calves and pigs (Hung *et al.*, 1984; Vonderfecht *et al.*, 1984, 1986; Chasey & Banks, 1984; Bridger, 1987; Theil *et al.*, 1985; Snodgrass *et al.*, 1984). These viruses are antigenically and genetically distinct from the group A rotaviruses (GARs) commonly associated with epidemics of infantile gastroenteritis in winter. Investigations of their molecular biology and epidemiology have been hindered by the inability to grow these viruses in cell culture. To date, cDNA clones have been synthesized from faecal isolates of the IDIR and ADRV strains of GBR (Sato *et al.*, 1989; Chen *et al.*, 1990a). Using these clones, GBR nucleotide sequences have been determined for gene 3 of the IDIR strain, and genes 9 and 11 of the ADRV strain. However, the nucleotide and amino acid sequences of equivalent genomic segments of heterologous GBR strains have not been compared. We have determined the nucleotide sequence of genomic segments 9 and 11 of IDIR and compared them to those of the equivalent ADRV gene segments, as well as GAR sequences.

The IDIR strain of GBR was propagated in infant CD-1 rats (Charles River Laboratories) as previously described (Vonderfecht *et al.*, 1984). Viral genomic RNA was extracted from infected intestinal contents and purified by CF-11 cellulose chromatography (Theil *et al.*, 1981). Synthesis of IDIR cDNA clones and their insertion into the plasmid pUC9 were accomplished by

oligo(T) priming (Both *et al.*, 1982; Gething *et al.*, 1979). Northern blot analysis was used to identify cDNA clones of each IDIR genome segment (Eiden *et al.*, 1989), the sequences of which were determined by dideoxynucleotide sequencing (Sanger *et al.*, 1977) using synthetic oligodeoxynucleotide primers and direct primer extension from recombinant pUC clones. Additional sequences were analysed after subcloning into M13 vectors to enable complete sequence analysis from both cDNA strands. The 5'- and 3'-terminal sequences were confirmed by direct RNA sequencing (Bassel-Duby *et al.*, 1986). Analyses of the IDIR nucleotide sequences and deduced primary protein sequences were performed with the GCG sequence analysis software package (Genetics Computer Group, Inc., Madison, Wis., U.S.A.), and comparison with the EMBL and GenBank databases.

The sequence of IDIR gene 11 (available under DDJ/EMBL/GenBank DNA database accession number D00912:IDIR11) was determined by sequence analysis from two nearly full-length cDNA clones. The length of gene 11 was 643 bp with one major open reading frame (ORF) encoding a translation product of 174 amino acids, or 20.1K if unmodified (Fig. 1a). There was a 58 bp untranslated sequence at the 5' end, with the first AUG codon at base 59, in the optimal sequence context for translation initiation [(A/G)NNAUGG; Kozak, 1981]. A second in-frame AUG at base 158 was also within an optimal initiation context. No significant translation products were deduced to be encoded by other reading frames, as has been shown for GAR gene



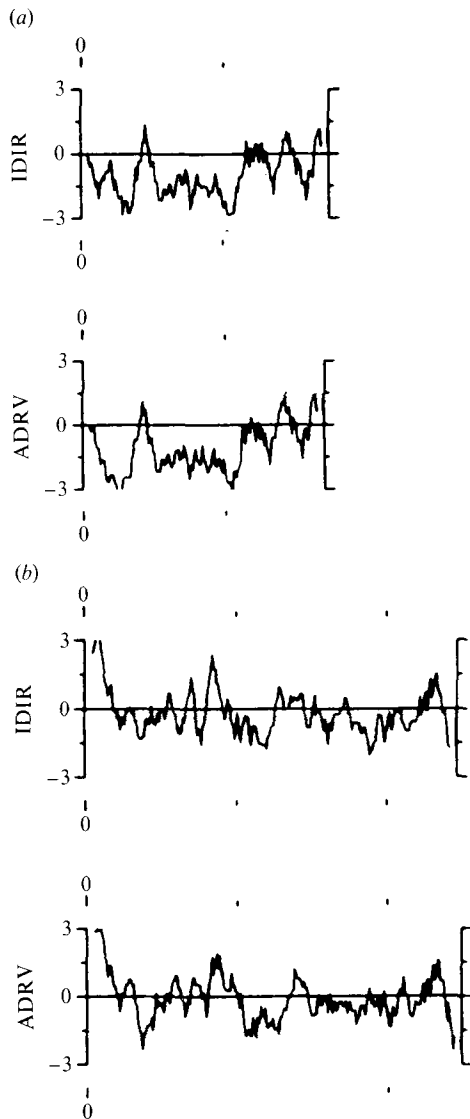


Fig. 2. Comparative hydrophilicity/hydrophobicity plots of the deduced amino acid sequences of GBR genes 11 (a) and 9 (b). The plots were generated by the method of Kyte & Doolittle (1982) using a computer analysis program (Devereux *et al.*, 1984).

conserved, being at identical equivalent positions, which implies that the polypeptide product of gene 9 may be a molecule with a high level of cross-linkage. The ADRV gene 9 product has three potential glycosylation sites, at amino acids 45, 91 and 105, whereas that of IDIR has only two sites, at amino acids 43 and 129.

It has been reported recently that the ADRV gene 9 protein sequence is 28% identical to the VP7 consensus sequence of the major structural glycoprotein of GARs (Chen *et al.*, 1990b). If substitutions of similar amino acids are included, this similarity increases to 78%, despite the disparity in size (ADRV, 249 amino acids;

GAR VP7, 326 amino acids). When the IDIR gene 9 protein sequence was aligned with the VP7 sequence in the same manner as described for ADRV, only 18% of the amino acids were identical. However, four of the six cysteines of the IDIR gene 9 product have the same positions in the VP7 sequence. If GBR gene 9 encodes a protein which is functionally and structurally equivalent to GAR VP7, IDIR may be more distantly related than ADRV to GARs. Future efforts should be aimed at identifying the gene encoding the GBR major outer capsid protein.

Although both the nucleotide and amino acid sequences of genes 9 and 11 of ADRV and IDIR are similar, neither pair shares the same degree of sequence conservation that is observed among heterologous strains of GARs. For example, a review of the EMBL and GenBank databases showed that the nucleotide and amino acid sequences of gene 11 of the Wa, OSU and UK strains of GAR are  $\geq 85.9\%$  and  $\geq 82.4\%$  conserved. In contrast, only 66.5% of the amino acids and 71.9% of the nucleotide sequence of ADRV and IDIR gene 11 are identical. Gene 9 of the two GBR strains is 60.6% and 51.2% identical at the nucleotide and amino acid sequence levels. This degree of similarity is substantially less than the 85% reported for heterologous VP7 genes obtained from different serotypes of GAR (Kapikian & Chanock, 1990). However, the differences noted between the ADRV and IDIR genes are largely due to substitutions of amino acids which conserve charge and/or polarity, as illustrated in Fig. 1(b) and 2(b).

The differences between the genes of IDIR and ADRV indicate that strains of GBR are more distantly related than different strains of GAR. This diversity should be confirmed by sequence analysis of additional GBR genomic segments from IDIR and ADRV, as well as other GBR strains. Since each of the rotavirus groups possesses very similar morphology, as determined by electron microscopic examination, the differences within the individual rotavirus groups as well as characteristics shared among the rotavirus groups should aid in our understanding of the mechanisms of pathogenesis and replication, and of structure.

This work was supported in part by funds provided by the National Institutes of Health grant no. 1 R29 AI24922-01A1. M.P. is on sabbatical leave from The Hospital for Sick Children, Toronto, Canada.

## References

- BASSEL-DUBY, R., SPRIGGS, D. R., TYLER, K. L. & FIELDS, B. N. (1986). Identification of attenuated mutations on the reovirus  $\sigma$  type 3, S1 double-stranded RNA segment with a rapid sequencing technique. *Journal of Virology* **60**, 64-67.

- BOTH, G. W., BELLAMY, A. R., STREET, J. E. & SIEGMAN, L. J. (1982). A general strategy for cloning double stranded RNA: nucleotide sequence of the simian-11 rotavirus gene 8. *Nucleic Acids Research* **10**, 7075-7088.
- BRIDGER, J. C. (1987). Novel rotaviruses in animals and man. *Ciba Foundation Symposium* **128**, 5-23.
- CHASEY, D. & BANKS, J. (1984). The commonest rotaviruses from neonatal lamb diarrhoea in England and Wales have atypical electropherotypes. *Veterinary Record* **115**, 326-327.
- CHEN, G. M., HUNG, T. & MACKOW, E. R. (1990a). cDNA cloning of each genomic segment of group B rotavirus ADRV: molecular characterization of the 11th RNA segment. *Virology* **175**, 605-609.
- CHEN, G. M., HUNG, T. & MACKOW, E. R. (1990b). Identification of the gene encoding the group B rotavirus VP7 equivalent: primary characterization of the ADRV segment 9 RNA. *Virology* **178**, 311-315.
- DEVEREUX, J., HAEBERLI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387-395.
- EIDEN, J. J., FIROOZMAND, F., SATO, S., VONDERFECHT, S. L., FANG, Z. Y. & YOLKEN, R. H. (1989). Detection of group B rotavirus in fecal specimens by dot hybridization with a cloned cDNA probe. *Journal of Clinical Microbiology* **27**, 422-426.
- EIDEN, J. J., WILDE, J., FIROOZMAND, F. & YOLKEN, R. (1991). Detection of animal and human group B rotaviruses by polymerase chain reaction. *Journal of Clinical Microbiology* **29**, 539-543.
- ESTES, M. K. & COHEN, J. (1989). Rotavirus gene structure and function. *Microbiological Reviews* **53**, 410-449.
- GETHING, M. J., BYE, J., SKEHEL, J. J. & WATERFIELD, M. (1979). Cloning and DNA sequence of double stranded copies of hemagglutinin genes from H<sub>2</sub> and H<sub>3</sub> strains of influenza virus. In *Structure and Variation in Influenza Viruses*, pp. 1-10. Edited by G. Laver & G. Air. New York: Elsevier/North-Holland.
- HUNG, T., CHEN, G., WANG, C., YAO, H., FANG, Z., CHAO, T., CHOU, Z., YE, W., CHANG, X., DEN, S., LIONG, X. & CHANG, W. (1984). Waterborne outbreak of rotavirus diarrhoea in adults in China caused by a novel rotavirus. *Lancet* **i**, 1139-1142.
- KAPIKIAN, A. Z. & CHANOCK, R. M. (1990). Rotaviruses. In *Virology*, pp. 1353-1404. Edited by B. N. Fields. New York: Raven Press.
- KOZAK, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Research* **9**, 5233-5250.
- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* **157**, 105-132.
- MITCHELL, D. B. & BOTH, G. A. (1988). Simian rotavirus SA11 segment 11 contains overlapping reading frames. *Nucleic Acids Research* **16**, 6244.
- PEDLEY, S., BRIDGER, J. C., BROWN, J. F. & MCCRAE, M. A. (1983). Molecular characterization of rotaviruses with distinct group antigens. *Journal of General Virology* **64**, 2093-2101.
- PROUDFOOT, N. J. & BROWNLEE, G. G. (1976). 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature, London* **263**, 211-214.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463-5467.
- SATO, S., YOLKEN, R. H. & EIDEN, J. J. (1989). The complete nucleic acid sequence of gene segment 3 of the IDIR strain of group B rotavirus. *Nucleic Acids Research* **17**, 10113.
- SNODGRASS, D. R., HERRING, A. J., CAMPBELL, I., INGLIS, J. M. & HARGREAVES, F. D. (1984). Comparison of atypical rotaviruses from calves, piglets, lambs, and man. *Journal of Virology* **65**, 909-914.
- SPIEGELBERG, H. L., ABEL, C. A., FISHKIN, B. G. & GREY, H. M. (1970). Localization of the carbohydrate within the variable region of light and heavy chains of human gamma G myeloma proteins. *Biochemistry* **9**, 4217-4223.
- THEIL, K. W., MCCLOSKEY, C. M., SAIF, L. J., REDMAN, D. R., BOHL, E. H., HANCOCK, D. D., KOHLER, E. M. & MOORHEAD, P. D. (1981). Rapid, simple method of preparing rotaviral double-stranded ribonucleic acid for analysis by polyacrylamide gel electrophoresis. *Journal of Clinical Microbiology* **14**, 273-280.
- THEIL, K. W., SAIF, L. J., MOORHEAD, P. D. & WHITMOYER, R. E. (1985). Porcine rotavirus-like (group B rotavirus): characterization and pathogenicity for gnotobiotic pigs. *Journal of Clinical Microbiology* **21**, 340-345.
- VONDERFECHT, S. L., HUBER, A. C., EIDEN, J. J., MADER, L. C. & YOLKEN, R. H. (1984). Infectious diarrhea of infant rats produced by a rotavirus-like agent. *Journal of Virology* **52**, 94-98.
- VONDERFECHT, S. L., EIDEN, J. J., TORRES, A., MISKUFF, R., MEBUYS, C. A. & YOLKEN, R. H. (1986). Identification of a bovine enteric syncytial virus as a non-group A rotavirus. *American Journal of Veterinary Research* **47**, 1913-1918.
- WELCH, S.-K. W., CRAWFORD, S. E. & ESTES, M. K. (1989). Rotavirus SA11 genome segment 11 protein is a nonstructural phosphoprotein. *Journal of Virology* **63**, 3974-3982.

---

(Received 28 March 1991; Accepted 30 July 1991)