Molecular epidemiology of rotaviruses in India

Shobha Broor, Dhrubaa Ghosh & Purva Mathur

Department of Microbiology, All India Institute of Medical Sciences, New Delhi, India

Received July 16, 2003

Rotaviruses cause an estimated 140 million cases of gastroenteritis and 800,000 deaths in children between the ages of 6 months to 2 yr in developing countries. In India, one of every 250 children or about 100-150,000 children die of rotavirus diarrhoea each year. The prevalence of rotavirus diarrhoea in India has been found to vary from 5-71 per cent in hospitalized children <5 yr of age with acute gastroenteritis. The seasonal variation of rotavirus diarrhoea in India varies in different geographical regions with high incidence in winter months at low relative humidity in north India. The distinctive features of rotavirus infection in India include the occurrence of severe disease at an early age and common neonatal rotavirus infections which are often asymptomatic. Rotavirus shows genetic and antigenic diversity in terms of subgroup, electropherotypes and G and P serotypes/genotypes. There are a few studies in terms of prevalence of different antigenic and genetic variants from various regions of India. In most studies on subgroup distribution from India a higher prevalence of subgroup II was reported compared to subgroup I. Electropherotyping has also demonstrated that a number of multiple electropherotypes co-circulate at one time in a particular community leading to extensive genomic variation and the appearance of new strains which may become the predominant electropherotype during the peak season.

The most common G types reported from India are G1 and G2 and P types are P[4] and P[8]. A significant number of children also have mixed rotavirus infections. G9 strains are also quite commonly seen in Indian children. In addition P6 strains of probable bovine origin have been reported from India. A novel neonatal strain P type 11 human rotavirus (116 E) was isolated from neonates in Delhi, the VP4 of which was closely related to the bovine serotype G10P[11] strain B223 and VP7 was closely related to the human serotype G9 strain. Another neonatal strain G10P[11] was reported from Bangalore. G10P[11] strains also have a high prevalence in calves with diarrhoea, in India. The occurrence of these unusual rotavirus strains which are natural reassortants of human and bovine rotaviruses, suggests that reassortment may be an important mechanism for generation of rotavirus strains of newborns. This is catalyzed by the age old traditions of calves and humans living in the same household and socio-economic conditions in India.

The diversity of rotavirus strains and the high prevalence of mixed infections in India are unique features of rotavirus epidemiology in India and emphasizes that vaccines should be formulated against a broad range of strains. Another important aspect is that vaccines in India should also target G9 strains. Since neonates acquiring rotavirus infection are protected against severe diarrhoea, newborn rotavirus strains can be effective potential vaccine candidates and vaccines based on these neonatal strains are being indigenously developed in India.

Key words Diarrhoea - electropherotypes - genetic diversity - molecular epidemiology - rotavirus

Acute diarrhoeal diseases are a major cause of childhood morbidity and mortality all over the world¹. Rotavirus has been recognized as the most common cause of severe diarrhoea in children below 5 yr of age. It is estimated that each year rotavirus accounts for 140 million cases of gastroenteritis leading to 800,000 deaths in developing countries accounting for 20-70 per cent of hospitalized cases of diarrhoea in
children between the ages of 6 months to 2 yr. In India, 1 of every 250 children or about 100,000 children die of rotavirus diarrhoea each year, accounting for 17 per cent of the world’s estimated rotavirus deaths. The incidence of rotavirus infection is similar among children in industrialized and developing countries. This signifies that improvement in hygiene and access to safe water may not significantly reduce the prevalence of rotavirus. Thus immunization offers one of the most promising methods to reduce the prevalence of rotavirus disease. However, as the immunity to rotavirus is type specific and there is a lot of heterogeneity in rotaviruses, knowledge of the geographical and temporal distribution of the common circulating genotypes is essential. This would help in developing an appropriate vaccine strategy for this highly morbid infection. This review examines the epidemiology of rotavirus infection in India with special emphasis on the molecular epidemiology of rotavirus circulating in different parts of India.

**Rotavirus**

Rotaviruses have a distinct wheel like appearance by negative-stain electron microscopy (EM) and thus have been named *rota* which in Latin means wheel. A member of the family Reoviridae, the virus has a genome of 11 segments of double-stranded RNA, of molecular weight 2 x 10⁵ to 2.2 x 10⁶ daltons, which is encased within a triple, layered capsid, the innermost layer of which, the core, contains the viral genome. Each RNA segment encodes a single viral polypeptide for 6 structural (VP1, VP2, VP3, VP4, VP6, VP7) and 5 non structural proteins (NSP1-NSP5). The structural proteins are located in the core (VP1-VP3), inner shell (VP6) and outer shell (VP4, VP7) (Fig. 1, Table).

![Schematic representation of rotavirus particle showing the distribution of structural proteins.](image)

Fig. 1. Schematic representation of rotavirus particle showing the distribution of structural proteins.

Rotaviruses display a wide host range and include agents of human infantile diarrhoea, Nebraska calf diarrhoea, epizootic diarrhoea of infant mice, SA11 virus of monkeys and swine rotaviruses.

The classification of rotaviruses is based on antigenic and genomic composition. The VP6 inner-shell polypeptide induces cross reactive non-neutralizing antibodies used to delineate rotavirus groups and subgroups. Thus group A-G can be differentiated with VP6 polyclonal and monoclonal antibodies by immunofluorescence, ELISA and immunoelectron microscopy. Groups A,B and C have been associated with human and animal infections whereas groups D,E, F and G have been found only in animals. Within group A rotavirus, subgroups I, II, I+II and non I and II are defined according to exclusive reactivities of two VP6- specific monoclonal antibodies. Types within group A are determined by cross-neutralization studies as serotypes or by sequence comparison as genotype. The VP4 and VP7 outer capsid proteins independently induce antibodies associated with type specific neutralization and protection from infection. Because two surface proteins carry neutralization specific antigens, a double classification has been established, similar to that developed for influenza virus, *i.e.*, G types (for VP7) and P types (for VP4). Till date, 14 G types and more than 21 P types have been described. Roman letters are used to designate groups, Roman numerals for subgroups and Arabic numbers for both G and P types.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene segment</th>
<th>Size of the protein (kD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP1</td>
<td>1</td>
<td>125,000</td>
</tr>
<tr>
<td>VP2</td>
<td>2</td>
<td>94,000</td>
</tr>
<tr>
<td>VP3</td>
<td>3</td>
<td>88,000</td>
</tr>
<tr>
<td>VP4</td>
<td>4</td>
<td>88,000</td>
</tr>
<tr>
<td>NSP1</td>
<td>5</td>
<td>53,000</td>
</tr>
<tr>
<td>VP6</td>
<td>6</td>
<td>41,000</td>
</tr>
<tr>
<td>NSP3</td>
<td>7</td>
<td>34,000</td>
</tr>
<tr>
<td>NSP2</td>
<td>8</td>
<td>35,000</td>
</tr>
<tr>
<td>VP7</td>
<td>9</td>
<td>38,000</td>
</tr>
<tr>
<td>NSP4</td>
<td>10</td>
<td>28,000</td>
</tr>
<tr>
<td>NSP5</td>
<td>11</td>
<td>26,000</td>
</tr>
<tr>
<td>NSP6</td>
<td>11</td>
<td>12,000</td>
</tr>
</tbody>
</table>
Thus, rotaviruses demonstrate tremendous heterogeneity. The observed diversity is due to the fact that these proteins are encoded by different RNA segments and thus independent reassortment of different viral genome segment can occur. The typing in other rotavirus groups is rudimentary. As an alternative to classification into serotype, which requires cultivation of virus, rotavirus can be classified into electropherotypes on the basis of migration pattern of their RNA segments on polyacrylamide gel electrophoresis (PAGE). The most prominent variation is observed between ‘short’, ‘long’ and ‘super-short’ RNA patterns, based on the difference in the mobility of gene segments 10 and 11. Moreover, a strong association of short and long RNA electropherotypes with subgroups I and II respectively has been observed by some workers.

Clinical features

The incubation period of rotavirus diarrhoea varies from 1-7 days. In newborns, the infection is usually asymptomatic, but 8-24 per cent of neonates may have minimal diarrhoea, and vomiting associated with fever. In infants and young children, there is an abrupt onset of severe vomiting and diarrhoea. Vomiting usually precedes the onset of diarrhoea. Stools are usually loose and watery, mucus may be present in 25 per cent of cases but blood is very rare. Mild to moderate dehydration is seen in 80 per cent of cases and severe loss of fluids and electrolytes may be fatal if untreated. Mild fever is seen in a large majority of cases. The severity of vomiting and diarrhoea may be more in subgroup II rotavirus infection. The illness usually lasts 3-8 days, but virus shedding continues for about 10 days to 1 month. In immunodeficient children, rotavirus can persist for months. Older children and adults are infected but they generally suffer from subclinical infections and virus is infrequently detected in their stool samples.

Laboratory diagnosis

Rotavirus is excreted in large numbers in the faeces (>10⁶ particles/g faeces) and thus can be easily identified on electron microscopy of stool samples which is one of the most specific tests for diagnosis. Direct EM examination of stools for rotavirus has a sensitivity of 80-90 per cent. However EM requires expensive equipment and trained personnel and thus cannot be used in field studies. Other methods like immuno-electrophoresis and modified complement fixation test were developed, but they lacked sensitivity. Following this, many rapid and economical assays like latex agglutination (LA), reverse passive haemagglutination assay (RPHA), solid phase agglutination of coated erythrocytes (SPACE), enzyme immuno assay (EIA) and polyacrylamide gel electrophoresis (PAGE) were used for diagnosis. Of these, the most widely used methods currently are LA, ELISA and PAGE. ELISA using polyclonal sera may give false positive results and requires validation by confirmatory or blocking ELISA. However, specific ELISA assays based on monoclonal antibodies have been developed. The detection of rotavirus by PAGE followed by silver staining though a highly specific technique lacks sensitivity as a minimum of 3-4ng of viral RNA is needed for detection. New methods like dot blot hybridization using radio labeled cDNA probes and reverse transcriptase - polymerase chain reaction (RT-PCR) are now being used as confirmatory methods for detecting rotavirus in stool samples. RT-PCR has been found to be a highly sensitive and specific method for diagnosis of rotavirus in stool sample from patients with acute diarrhoea.
Epidemiology

World

Rotavirus infection has a worldwide distribution and is the single most important cause of gastroenteritis in young children. Typically 50-60 per cent of instances of acute gastroenteritis in hospitalized children throughout the world are caused by rotaviruses. The virus mainly spreads via the feco-oral route, through respiratory route, person-to-person contact, or contaminated environmental surfaces and fomites. Animal to human transmission is not common but recombinants can arise during co-infections. Rarely, rotavirus can cause epidemics of diarrhoea in nurseries, nursing homes and isolated islands. Symptomatic infections are most common in children between the ages of 6 months to 2 yr with a peak incidence at 9-12 months.

Males are more frequently affected than females. In temperate climates rotavirus diarrhoea has a higher prevalence during the winter month. In the tropical countries however, the seasonal variation is not clear. In certain studies higher incidences have been reported during the rainy season while in others no seasonal variation has been found.

Group A rotavirus are by and large the most widely distributed worldwide. Group B have caused widespread outbreaks of diarrhoeal illness in adults and children in China and had not since then been reported from outside China. Almost 16 yr after this Chinese outbreak, sporadic cases of adult group B rotavirus diarrhoea were reported from Kolkata. Group C rotavirus cause sporadic cases and occasional outbreaks of diarrhoeal illness in infants and young children worldwide.

At any one time, there is co-circulation of rotaviruses of different G types, types G1-G4 representing 95 per cent of strains worldwide of which type G1 accounts for 50 per cent. Although theoretically reassortment of independent viral RNA segments could result in more than 100-P-G serotype combinations, the majority of human isolates characterized so far from children with diarrhoea fall into four groups G1 P[8], G3 P[8], G4 P[8] and G2 P[4]. Human isolates with other G and P serotype specificities (eg. G8 P[10], G1 P[9]) have so far been detected rarely. The rotavirus commonly designated as newborn strains have a unique P serotype P[6] and G type 1-4.

India

It is estimated that close to 100,000 annual deaths are caused by rotavirus in India, that is, 1 in every 250 children born in India will die from rotavirus by the age of 5 yr. India accounts for 17 per cent of the world’s estimated rotavirus associated deaths. A number of studies have been conducted on the prevalence of childhood rotavirus diarrhoea in various parts of the country in which rotavirus was detected in 5 - 71 per cent of the hospitalized children less than 5 yr of age with acute gastroenteritis. This variation may be due to the duration of the study, number of children studied and the seasonal variation of rotavirus diarrhoea in different regions of the country. Most of the studies have used ELISA for detection of rotavirus. The maximum number of studies have been conducted in the northern states of Delhi and Punjab and southern states of Karnataka and Tamil Nadu. In contrast, only limited data are available on neonatal rotavirus diarrhoea as well as the distribution of various serotypes/genotypes in India. The prevalence of childhood rotavirus in the north Indian cities of Delhi, Chandigarh and Aligarh has been reported to vary from 6-45 per cent. Studies from Delhi alone have shown highly variable prevalences with one study showing a low 6 per cent prevalence of rotavirus in acute diarrhoea. Other studies, revealed rotavirus infection rates from 15-18 per cent, 24 per cent, 32 per cent and 45 per cent in cases with diarrhoea. In Chandigarh, rotavirus was detected in 16-19 per cent of instances of acute gastroenteritis in children ≤ 5 yr of age while in Aligarh it was detected in 19 per cent of cases with diarrhoea. In the western states of India, in Pune, rotavirus was detected in 28-30 per cent of children ≤5 yr of age with acute diarrhoea. In eastern India, in Kolkata the incidence of rotavirus associated diarrhoea varied from 5-22 per cent. On the other hand, in Manipur the incidence was as high as 41 per cent.

Although group A rotaviruses are the most important pathogens, human group B rotavirus (HuGBR) was identified in five adult patients during a hospital surveillance programme on diarrhoeal pathogens at the Infectious Diseases Hospital in Kolkata. These strains were distinct from the earlier reported Chinese strain.
of group B rotavirus. These strains have been characterized as Cal strain of group B rotavirus. In southern India, rotavirus was detected in 18 per cent of children with acute gastroenteritis in Vellore, 20.8 per cent in Chennai and 16-22 per cent in Bangalore. A high incidence of 71 per cent was observed in Calicut.

Very few studies have been done on rotavirus infection in rural populations. In a study from north India, the incidence of diarrhoea caused by rotavirus was 34 per cent. In rural coastal Karnataka rotavirus infection rate of 11 per cent was reported while rotavirus was detected in 23.9 per cent of children with diarrhoea in a rural community in Tirupati.

Rotavirus diarrhoea shows a seasonal variation with a high incidence of the disease in winter months at low relative humidity in north India. In a study from Punjab, rotavirus infection has been observed throughout the year with maximum occurrence in November and another peak in the hot and dry months of May. The maximum incidence in Pune occurred in winter and the minimum in the rainy season.

There are several distinctive features of rotavirus infection in India. The age distribution of hospitalized children with rotavirus indicates that severe rotavirus disease occurs at an early age in Indian children and neonatal rotavirus infections although often asymptomatic, are common. To date six clinical studies have been conducted on neonatal rotavirus diarrhoea (three in Delhi and one each in Vellore, Kolkata and Pondicherry) in which rotavirus was detected in stools of a mean of 20.5 per cent neonates. In 3 of the 4 cities (7/10 hospitals) surveyed, a significant rate (>15%) of rotavirus infection was found among neonates. A large number of infections in these studies were asymptomatic. It was also found that the rate of infection was directly related to the duration of hospital stay.

While infants, up to 1 yr of age, account for only 50 per cent of all rotavirus hospitalization in the United States, approximately 80 per cent of all rotavirus hospitalization in India occur in infants. Studies on the relationship of age to susceptibility showed the percentage of rotavirus positive children from 6-24 months of age to be 86.09 per cent but the most susceptible age was 6-12 months as in other developing countries. Rotavirus shows genetic and antigenic diversity in terms of subgroup, electropherotypes and G and P serotypes/genotypes. There are a few studies in terms of prevalence of different antigenic and genetic variants from various regions of the country. Studies on subgroup distributions have been carried out in north India in Chandigarh and in Delhi and in all these studies a higher prevalence of subgroup II varying from 50-70 per cent was reported whereas subgroup I was reported in about 30 per cent. In Delhi non I and non II as well as I + II subgroup were also detected in a few cases. The clinical severity of the disease was found to be higher in subgroup II infections as compared to subgroup I infection.

In south India also subgroup II strains of rotavirus were found to predominate, in Vellore 69 per cent of the strain had a subgroup II specificity, 29 per cent had subgroup I specificity and 6 per cent were non subgroup I - non subgroup II strains. In Chennai and Hyderabad also subgroup II was predominant constituting 60-63 per cent, subgroup I 20-37 per cent and dual specificity of subgroup I + II was seen in a minority of rotavirus strains. Usually subgroup II strains have a long RNA pattern and subgroup I have a short RNA pattern but variation in their correlation can exist. In a study from Chandigarh all subgroup II strains were found to have long electropherotype whereas subgroup I had a short RNA pattern. Electropherotyping has also demonstrated that a number of variant strains can co-circulate at one particular time but usually a particular electropherotype predominates. In 2 different studies from Delhi multiple RNA patterns of ‘long’ and ‘short’ electropherotypes were detected with a sequential appearance of electropherotypes with every seasonal peak of infection. Since multiple electropherotypes co-circulate simultaneously in the community this can lead to extensive genomic variation in rotavirus strains with the appearance of new strains which may become the predominant electropherotype during the peak season. In Manipur variant strains of rotavirus were detected, 6 isolates had a subgroup I specificity with ‘long’ electropherotype whereas 1 isolate had a subgroup II specificity with ‘short’ electropherotype. In Mysore subgroup I strains having a ‘long’ electropherotype were detected from children with acute gastroenteritis which were thought to be of animal origin.
India being a vast country it is essential to study the distribution of various serotypes/genotypes in different parts of the country as this will have a bearing on the development of a vaccine for India. Fifteen G serotypes of rotaviruses are recognized, depending on the molecular characterization of VP7 (glycosylated outer capsid protein). However, G1 to G4 are the most pre-dominant genotypes in humans. Moreover, a number of unusual genotypes, G5, G8, G6 and G9, have also been reported recently from various countries. The most common P serotypes infecting humans are P[8] and P[4] genotypes. Although the role of VP4 protein in protective immunity is not very well established, information on G and P typing is important for identifying unusual or new virus strains circulating in different populations.

In most of the studies from India the most common G types reported were G1 and G2 and P types were P[4] and P[8]. A significant number of children also had mixed rotavirus infections. In a study from Delhi, 60 rotavirus positive stool specimens were characterized into G and P genotypes by oligo-hybridization and nested PCR. Of the 60, 44 could be G typed, of which 17 were G1, 13 G2, 5 G3, 4 G4 and 5 had multiple G types. Forty three samples were P typed, of these 23 were P[8], 14 were P[4], 4 were P[6] and 2 had multiple P types. All G2 strains had P[4] genotype. The G1, G3 and G4 strains were characterized into P[6] and P[8] genotypes in a more extensive study also from Delhi. 287 strains were G and P genotyped by RT PCR. Of the four strains common globally, three (G1 P[8] 15%, G2 P[4] 22%; G4 P[8] 6%), were found in 43 per cent of samples whereas G9 strains made up 17 per cent of the total. Three different G9 strains were seen, one P[8] G9 strain, which displayed the long electropherotype and subgroup II VP6 specificity, and two P[6] G9 strains, one with the long electropherotype and subgroup II specificity and the other with the short electropherotype and subgroup I specificity. Serotype G2 strains were detected more often in infections caused by single strains than in mixed infections, whereas serotype G1 strains were found more often in mixed infections than in infections caused by single strains. In a study on 63 rotavirus strains from 5 Indian cities, 10 different genotypes were identified with five different G genotypes and found distinct P types by using RT-PCR. The common worldwide strains G1 P[8], G2 P[4], G3 P[8], and G4 P[8] were underrepresented among Indian children (33%), whereas strains of P type [6] (G1 P[6], G2 P[6], G3 P[6], G4 P[6], and G9 P[6], which primarily infect asymptomatic newborns but are rare in children with diarrhoea were common in India (43%). Of these, G9 P[6], a strain not previously reported to be found in children with diarrhoea, was the most prevalent (22%). Eleven per cent of the strains were non-typeable, and another 11 per cent of the specimens had mixed infections. The epidemiological significance of G9 rotavirus strains, if confirmed in other settings, may have important implications for vaccine development.

From north east India serotype 2 was found to be the predominant type circulating. In a study from Kolkata, the predominant genotype was G1 P[8] (20%), followed by G2P[4] (15%) and G4P[8] (6%). A number of uncommon genotypes, G1 P[4] (4%), G2 P[8] (2.5%), G2 P[6] (0.6%), G4 P[4] (2.5%), and G4 P[6] (1.25%), were also observed. Twenty two per cent of specimens showed mixed infections, and 24 per cent of the total samples remained untypeable. A novel rotavirus strain G4 P[8] was reported from children with acute diarrhoea in this study. The studies on serotyping and genotyping are rather limited from south India. In a study from Chennai serotyping was done on 48 strains and serotype 2 was most predominant (68.8%) followed by serotype 1 (14.6%), serotype 3 in 1 sample and mixed infections were seen in 14.6 per cent of samples. In another study from Vellore in Tamil Nadu on 100 rotavirus strains the commonest G types seen were G1, G4, G2, G9, G3 and G8 in the order of frequency, and the P types were P[4], P[8], and P[6] and the most common G:P combinations were G1 P[8], G1 P[4], G2 P[4] and G4 P[8]. In a study between 1988-1994 in Mysore and Banglore serotype G3 was found to predominate in Bangalore and serotype G1 in Mysore. Among non typable strains from both cities several exhibited dual subgroup (SGI+II) or subgroup I specificity and ‘long’ RNA pattern indicating their probable animal origin. Similar strains were also reported in Mysore from 1993-94 by Jagannath and colleagues again suggesting their animal origin. These strain were related to P6[1], G8 bovine rotavirus. From Pune in west India serotyping was done on 107 rotavirus strains, serotype 2 were found to be commonest (45%), followed by serotype 1(13.8%), serotype 4 (8.3%) and serotype 3 (0.09%). Mixed infections were seen in 22.9 per cent of strains.
A novel neonatal strain P type 11 human rotavirus (116 E) was isolated from a newborn in a nursery of Delhi. The VP4 of this strain was closely related to that of the bovine serotype G10 P[11] of strain B223. However the VP7 protein of 116E was closely related to human serotype G9 strain, F45 and W16135. Another neonatal strain G10 P[11] was reported from Bangalore67. These unusual rotavirus strains which are natural reassortants of human and bovine rotaviruses suggest that reassortment may be an important mechanism for generation of rotavirus strains of newborns. The year round circulation of high titres of rotavirus and consequent natural reassortment may also lead to the diversity of strains. A notable finding is the predominance of G10 P[11] strain in south, west and central India among calves with diarrhoea, in contrast to a high prevalence of G6 in other countries. This may be the source of G10 P[11] in neonates of India catalyzed by the age old traditions and socio-economic conditions in India35,68.

The diversity of rotavirus strains and the high prevalence of mixed infections in India are unique features of rotavirus epidemiology in India and emphasize that vaccines should be formulated against a broad range of strains. Another important aspect is that the vaccines in India should also target G9 strains.

It is observed that neonates acquiring rotavirus infection are protected against severe diarrhoea. Thus, neonatal rotavirus strains are effective potential vaccine candidates69. Some of these vaccines are being indigenously developed in India.

References


67. Das M, Dunn SJ,Woode GN, Greenberg HB, Rao CD. Both surface proteins (VP4 and VP7) of an asymptomatic neonatal rotavirus strain (I321) have high levels of sequence identity with the homologous proteins of a serotype 10 bovine rotavirus. *Virology* 1993; 194: 374-9.


Reprint requests: Dr Shobha Broor, Professor, Department of Microbiology, All India Institute of Medical Sciences Ansari Nagar, New Delhi 110029, India e-mail:broor@hotmail.com