Genomic characterization of coxsackievirus type B3 strains associated with acute flaccid paralysis in south-western India

Rongala Laxmivandana,1 Sarah S. Cherian,2 Prasanna Yergolkar3 and Shobha D. Chitambar1

Correspondence
Shobha D. Chitambar
drshobha.niv@gmail.com

1Enteric Viruses Group, National Institute of Virology, Pune, India
2Bioinformatics Group, National Institute of Virology, Pune, India
3National Institute of Virology, Bangalore Unit, Bengaluru, India

Acute flaccid paralysis (AFP) associated with coxsackievirus type B3 (CV-B3) of the species Enterovirus B is an emerging concern worldwide. Although CV-B3-associated AFP in India has been demonstrated previously, the genomic characterization of these strains is unreported. Here, CV-B3 strains detected on the basis of the partial VP1 gene in 10 AFP cases and five asymptomatic contacts identified from different regions of south-western India during 2009–2010 through the Polio Surveillance Project were considered for complete genome sequencing and characterization. Phylogenetic analysis of complete VP1 gene sequences of global CV-B3 strains classified Indian CV-B3 strains into genogroup GVI, along with strains from Uzbekistan and Bangladesh, and into a new genogroup, GVII. Genomic divergence between genogroups of the study strains was 14.4% with significantly lower divergence (1.8%) within GVI (n=12) than that within GVII (8.5%) (n=3). The strains from both AFP cases and asymptomatic contacts, identified mainly in coastal Karnataka and Kerala, belonged to the dominant genogroup GVI, while the GVII strains were recovered from AFP cases in north interior Karnataka. All study strains carried inter-genotypic recombination with the structural region similar to reference CV-B3 strains, and 5’ non-coding regions and non-structural regions closer to other enterovirus B types. Domain II structures of 5’ non-coding regions, described to modulate virus replication, were predicted to have varied structural folds in the two genogroups and were attributed to differing recombination patterns. The results indicate two distinct genomic compositions of CV-B3 strains circulating in India and suggest the need for concurrent analysis of viral and host factors to further understand the varied manifestations of their infections.

INTRODUCTION

Coxsackievirus type B3 (CV-B3) belongs to the species Enterovirus B, in the genus Enterovirus and family Picornaviridae. This infectious agent spreads through the faecal–oral route in humans. The virus is a non-enveloped icosahedral particle of 27–30 nm in diameter that contains a positive-sense, ssRNA molecule of ~7400 nt. Flanked by two highly structured non-coding regions (NCRs) at the 5’ and 3’ ends, a single large ORF of the genome encodes four structural (VP1, VP2, VP3 and VP4) and seven non-structural (2A, 2B, 2C, 3A, 3B, 3C and 3D) viral proteins (Pallansch & Roos, 2007; Racaniello, 2007).

Despite the closely related genomic frames, strains of the same CV-B3 type are known to cause a wide range of infections in humans from subclinical to severe forms. CV-B3 is a common causative agent of cardiac diseases, pancreatitis and acute flaccid paralysis (AFP) (Kaplan et al., 1983; Wong et al., 2011). Such diverse outcomes have been described to be due in part to the genetics/immune status of the host or the initial viral load during infection or the viral genetics or diversity in the viral population (Tracy et al., 2008). Among these, analysis of the viral genome is considered important for understanding the issue of disease causation and/or outcome (Tracy et al., 2008). However, approaches that are based on sequencing of the variable genomic region of structural proteins,
often used for typing of viral genomes, have been described to have limited ability to delineate the viral strains of the same genotype responsible for different clinical manifestations (Nasri et al., 2007). In contrast, the 5’NCR sequence and its folds have been reported to be useful to differentiate poorly virulent and cardiovirulent genomes of CV-B3 strains (Dunn et al., 2003; Tracy et al., 2008). The frequent recombination events in these viruses further

![Phylogenetic tree](image)

**Fig. 1.** Phylogenetic tree reconstructed on the basis of complete VP1 (nt 2453–3300) sequences of the strains of this study and CV-B3 reference strains. ■ Strains from AFP cases; □, strains from asymptomatic contacts of the present study (GVI and G VII). Reference strains are shown with GenBank accession numbers. CV-B3 strains belonging to other genogroups (GI–GIV) are condensed and the available data on the country and year of these strains are depicted in the table alongside the phylogenetic tree. The bar indicates genetic distance (nucleotide substitutions per site).

### Table 1. Strains included in this study

<table>
<thead>
<tr>
<th>CV-B3 strains from AFP cases</th>
<th>CV-B3 strains from asymptomatic contacts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain name</strong></td>
<td><strong>Sampling date</strong></td>
</tr>
<tr>
<td>NIV1021241LV380_P1</td>
<td>Jul 2010</td>
</tr>
<tr>
<td>NIV909331LV26_P2</td>
<td>Mar 2009</td>
</tr>
<tr>
<td>NIV909311LV33_P3*</td>
<td>Mar 2009</td>
</tr>
<tr>
<td>NIV909931LV53P4</td>
<td>May 2009</td>
</tr>
<tr>
<td>NIV9014321LV141P5</td>
<td>Jul 2009</td>
</tr>
<tr>
<td>NIV9023491LV157_P6</td>
<td>Sep 2009</td>
</tr>
<tr>
<td>NIV9099101LV198P7</td>
<td>May 2009</td>
</tr>
<tr>
<td>NIV9099741LV204_P8</td>
<td>May 2009</td>
</tr>
<tr>
<td>NIV9017801LV243_P9</td>
<td>Aug 2009</td>
</tr>
<tr>
<td>NIV1019241LV350P10</td>
<td>Jul 2010</td>
</tr>
</tbody>
</table>

*Denotes AFP index case/its household or neighbourhood asymptomatic contacts.
stresses the need for complete genome sequencing (Kyriakopoulou et al., 2015).

Although the CV-B3 strains associated with AFP have often been identified on the basis of VP1 gene sequencing and BLAST analysis (Oberste et al., 1999; Wong et al., 2011), phylogenetic characterization of these strains has rarely been carried out. A few recent studies based on VP1 have classified CV-B3 strains into five genogroups, GI–GV (Chu et al., 2010; Huang et al., 2014). Of these GI, GII and GIII strains were found to be prevalent in America and Europe, GIV strains were reported from Taiwan and GV strains were found predominantly in East Asia (Chu et al., 2010; Huang et al., 2014). The importation/exportation of the CV-B3 genogroups across countries is presumed to be due to frequent population exchange (Tao et al., 2012). In India, although a few studies have demonstrated an association of CV-B3 with AFP (Dhole et al., 2009; Laxmivandana et al., 2013; Rao et al., 2012), genomic characterization of such strains has not been reported to date.

The aim of this study was therefore to sequence and analyse the complete genomes of CV-B3 strains isolated from 10 AFP cases and five asymptomatic contacts identified during 2009–2010 in different geographical/meteorological zones [north interior Karnataka (NIK), south interior Karnataka (SIK), coastal Karnataka (CK) and Kerala] in the south-western states of India (Table 1). The study was further focused on examination of genomic variations, phylogeny, recombination patterns and 5’NCR secondary structures of these strains.

## RESULTS

### Genome organization of the CV-B3 strains

The genomes of the CV-B3 strains of this study were found to be 7399–7400 nt, with a G+C content of 47–48 mol% and consisting of a single long ORF of 6555 nt that corresponds to a precursor polyprotein of 2185 aa flanked by NCRs of 743–744 and 101 nt at the 5’ and 3’ ends, respectively.

### Phylogenetic analyses

Phylogenetic trees were reconstructed using full-length (nt 1–7400) and partial genomic sequences [5’NCR (nt 1–744), structural (P1: nt 745–3300), VP1 (nt 2453–3300) and non-structural (P2-P3-3’NCR: nt 3301–7400)] to investigate the relationship of the study strains with other CV-B3 strains and/or the closely related other enterovirus B (EV-B) prototype/reference strains identified by NCBI BLAST.

In the complete VP1 region, the study strains were found to cluster into two distinct genogroups, GVI and GVII (Figs 1 and S1, available in the online Supplementary Material). The nucleotide divergence between the two genogroups was found to be 15.9 % in the VP1 region (Table 2). These genogroups were also significantly distant from the genogroups GI/GII/GIII and GIV/GV, which comprised strains from non-Asian and East Asian countries, respectively. GVI included the study strains [CK (n=7), Kerala (n=3), NIK (n=1), SIK (n=1)], other Indian strains [Karnataka/NIK (2009, n=6); Table 2. Mean nucleotide and amino acid diversities (%) of the study strains

<table>
<thead>
<tr>
<th>Strains</th>
<th>VP1</th>
<th>P1</th>
<th>P2-P3</th>
<th>5’NCR</th>
<th>Complete genome/polyprotein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study CV-B3 vs other CV-B3*</td>
<td>14.8</td>
<td>2.5</td>
<td>15.6</td>
<td>2.1</td>
<td>19.1</td>
</tr>
<tr>
<td>Study CV-B3 vs other EV-B†</td>
<td>37.6</td>
<td>31.3</td>
<td>32.3</td>
<td>25.3</td>
<td>14.9</td>
</tr>
<tr>
<td>Between GVI and GVII study/CV-B3 strains</td>
<td>15.9</td>
<td>1.8</td>
<td>16.5</td>
<td>1.7</td>
<td>13.1</td>
</tr>
<tr>
<td>Within GVI study/CV-B3 strains</td>
<td>1.5</td>
<td>1</td>
<td>1.6</td>
<td>1</td>
<td>2.2</td>
</tr>
<tr>
<td>Within GVII study/CV-B3 strains</td>
<td>6.1</td>
<td>0.5</td>
<td>5.6</td>
<td>0.4</td>
<td>11.4</td>
</tr>
</tbody>
</table>

*Other CV-B3 (gb_AY896762 and gb_FJ357838).
†Other EV-B [CV-B4 (gb_KF878966), E-7 (gb_FJ460595) and EV-B74 (gb_JQ397329)].

In the complete VP1 region, the study strains were found to cluster into two distinct genogroups, GVI and GVII, and a new genogroup, GVII (Figs 1 and S1, available in the online Supplementary Material). The nucleotide divergence between the two genogroups was found to be 15.9 % in the VP1 region (Table 2). These genogroups were also significantly distant from the genogroups GI/GII/GIII and GIV/GV, which comprised strains from non-Asian and East Asian countries, respectively. GVI included the study strains [CK (n=7), Kerala (n=3), NIK (n=1), SIK (n=1)], other Indian strains [Karnataka/NIK (2009, n=6); Table 2. Mean nucleotide and amino acid diversities (%)

<table>
<thead>
<tr>
<th>Strains</th>
<th>Complete genome/polyprotein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study CV-B3 vs other CV-B3*</td>
<td>17.3</td>
</tr>
<tr>
<td>Study CV-B3 vs other EV-B†</td>
<td>20.8</td>
</tr>
<tr>
<td>Between GVI and GVII study/CV-B3 strains</td>
<td>14.4</td>
</tr>
<tr>
<td>Within GVI study/CV-B3 strains</td>
<td>14.8</td>
</tr>
<tr>
<td>Within GVII study/CV-B3 strains</td>
<td>15.9</td>
</tr>
</tbody>
</table>

*Other CV-B3 (gb_AY896762 and gb_FJ357838).
†Other EV-B [CV-B4 (gb_KF878966), E-7 (gb_FJ460595) and EV-B74 (gb_JQ397329)].

### Phylogenetic analyses

Phylogenetic trees were reconstructed using full-length (nt 1–7400) and partial genomic sequences [5’NCR (nt 1–744), structural (P1: nt 745–3300), VP1 (nt 2453–3300) and non-structural (P2-P3-3’NCR: nt 3301–7400)] to investigate the relationship of the study strains with other CV-B3 strains and/or the closely related other enterovirus B (EV-B) prototype/reference strains identified by NCBI BLAST.

In the complete VP1 region, the study strains were found to cluster into two distinct genogroups, GVI and GVII (Figs 1 and S1, available in the online Supplementary Material). The nucleotide divergence between the two genogroups was found to be 15.9 % in the VP1 region (Table 2). These genogroups were also significantly distant from the genogroups GI/GII/GIII and GIV/GV, which comprised strains from non-Asian and East Asian countries, respectively. GVI included the study strains [CK (n=7), Kerala (n=3), NIK (n=1), SIK (n=1)], other Indian strains [Karnataka/NIK (2009, n=6); Table 2. Mean nucleotide and amino acid diversities (%)]

### RESULTS

**Genome organization of the CV-B3 strains**

The genomes of the CV-B3 strains of this study were found to be 7399–7400 nt, with a G+C content of 47–48 mol% and consisting of a single long ORF of 6555 nt that corresponds to a precursor polyprotein of 2185 aa flanked by NCRs of 743–744 and 101 nt at the 5’ and 3’ ends, respectively.

**Phylogenetic analyses**

Phylogenetic trees were reconstructed using full-length (nt 1–7400) and partial genomic sequences [5’NCR (nt 1–744), structural (P1: nt 745–3300), VP1 (nt 2453–3300) and non-structural (P2-P3-3’NCR: nt 3301–7400)] to investigate the relationship of the study strains with other CV-B3 strains and/or the closely related other enterovirus B (EV-B) prototype/reference strains identified by NCBI BLAST.

In the complete VP1 region, the study strains were found to cluster into two distinct genogroups, GVI and a new genogroup, GVII (Figs 1 and S1, available in the online Supplementary Material). The nucleotide divergence between the two genogroups was found to be 15.9 % in the VP1 region (Table 2). These genogroups were also significantly distant from the genogroups GI/GII/GIII and GIV/GV, which comprised strains from non-Asian and East Asian countries, respectively. GVI included the study strains [CK (n=7), Kerala (n=3), NIK (n=1), SIK (n=1)], other Indian strains [Karnataka/NIK (2009, n=6); Table 2. Mean nucleotide and amino acid diversities (%)]

**Table 2. Mean nucleotide and amino acid diversities (%) of the study strains**

<table>
<thead>
<tr>
<th>Strains</th>
<th>Complete genome/polyprotein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study CV-B3 vs other CV-B3*</td>
<td>17.3</td>
</tr>
<tr>
<td>Study CV-B3 vs other EV-B†</td>
<td>20.8</td>
</tr>
<tr>
<td>Between GVI and GVII study/CV-B3 strains</td>
<td>14.4</td>
</tr>
<tr>
<td>Within GVI study/CV-B3 strains</td>
<td>14.8</td>
</tr>
<tr>
<td>Within GVII study/CV-B3 strains</td>
<td>15.9</td>
</tr>
</tbody>
</table>
Fig. 2 (cont.)

(b) P2-P3-3' NCR
Fig. 2 (cont.)

(c) 5′ NCR

Genomic analysis of CVB3 strains from India

http://jgv.microbiologyresearch.org 699
Fig. 2 (cont.)
(d) Complete genome

CV-B3 GIII

CV-B3 GVIII

CV-B3 GV

CV-B3 GI

Exploring the genetic diversity and evolution of the Echovirus serotype 9 (E9) genome, researchers have identified several unique genetic features. The complete genome of E9 is depicted here, highlighting the genetic variability within and across different strains.

E9 strains, such as EV93-10887-99, EV93-14103-00, and EV93-15936-01, have been sequenced and analyzed to understand their geographical and temporal distribution. The genomic sequences show a high degree of similarity, indicating a broad host range and potential for widespread transmission.

The complete genome of E9 is approximately 7500 bases long, with varying coding regions and non-coding segments. The coding regions are essential for viral replication and transcription, while the non-coding regions play a role in viral packaging and stability.

Analyzing the complete genome allows researchers to identify potential targets for vaccine development and antiviral therapies. The genetic diversity within the E9 genome also provides insights into the evolutionary pressures acting on the virus.

In conclusion, the complete genome analysis of E9 strain EV93-10887-99 provides a valuable tool for understanding the viral life cycle, host-virus interactions, and the epidemiology of E9 infections.
the north Indian state Uttar Pradesh (UP) (2008, \(n=3\)), and strains from Uzbekistan (1999, \(n=1\)) and Bangladesh (\(n=1\)). GVII included three study strains from NIK and also other Indian strains [SIK (2008, \(n=1\)), Kerala (2007, \(n=1\); 2008, \(n=1\)) and UP (2009, \(n=2\)) (Figs 1 and S1).

The phylogenetic tree based on the P1 region (Fig. 2a) also identified the CV-B3 genogroups (GI–GIII and GV–GVII) with a topology similar to that of the complete VP1-based phylogeny (Fig. 1). It was noted that, except for the VP1 region, no other gene sequences of the GIV strains were available in GenBank. The phylogeny of P2-P3-3’NCR (Fig. 2b) revealed that the study CV-B3 strains formed two independent clusters that were distant from the other CV-B3 strains but closer to the strains of other EV-B types. In contrast, the phylogeny of the 5’NCR (Fig. 2c) indicated that the two clusters of the study strains were distant from each other and showed relatedness with different groups of CV-B3/other EV-B strains. Although the topologies of the phylogenetic trees based on P2-P3-3’NCR and 5’NCR (Fig. 2b, c) were discrepant in comparison with that of the P1-based phylogeny (Fig. 2a), the majority of the CV-B3 strains were noted to retain their clustering in the same genogroup (Fig. 2). In the complete genome-based phylogenetic tree (Fig. 2d), the two clusters of the study strains were found to share monophyletic clustering with all other CV-B3 strains, as in the P1 region-based phylogeny, although with a different topology. All the CV-B3 strains, except for strain gb_AY896762 from Uzbekistan, could be distinguishable into the specific genogroups in the complete genome-based phylogenetic tree.

Although the phylogenetic trees of all the analysed regions reflected the occurrence of the two genogroups of the study strains [GVI (\(n=12\)) and GVII (\(n=3\))], no segregation of the strains from AFP cases, with respect to the strains from asymptomatic contacts, was noted (Figs 1 and 2).

**Recombination analysis**

To investigate further the incongruence observed in the phylogeny of the study strains in the different genomic regions (Fig. 2), recombination analysis was carried out separately for each of these strains in comparison with other phylogenetically close CV-B3 and other EV-B reference strains.

The similarity plot analysis revealed a mosaic recombination between the genome sequences of the CV-B3 study strains and other EV-B reference strains (Fig. S2). Furthermore, Bootscan analysis confirmed these recombination events (Fig. S2). Fig. 3(a, b) shows the recombination patterns of strains representative of GVI and GVII,
respectively. The P1 genomic region of all of the study strains showed high similarity to CV-B3 type only. In contrast, the P2-P3 region of the study strains showed high similarity/recombination events with different EV-B types (CV-B4, E-7 and EV-B74). Furthermore, in the non-structural region, all of the study strains within genogroup GVI displayed a similar pattern of recombination that was found to be different from the pattern detected in the GVII strains (Figs 3 and S2). In the 5’NCRs, all strains from GVI were identified to have low sequence homology (unclassified recombination) with any of the CV-B3 or other EV-B reference strains (Figs 3a and S2a), while the study strains from GVII showed marks of recombination with E-7 and EV-B74 strains (Figs 3b and S2b).

As multiple types were found to be similar in different regions of the genomes of the study strains, similarity plots and Bootscan graphs of the representative study strains were also constructed using the consensus sequences of (i) two CV-B3 strains (gb_AY896762 and gb_FJ357838) and (ii) three other EV-B strains [CV-B4 (gb_KF878966), E7 (gb_FJ460595) and EV-B74 (gb_JQ397329)] as putative parents, with a sequence of CV-A24 strain (gb_KF667361) as an outgroup. The occurrence of the mosaic recombination between the CV-B3 and other EV-B types was revealed with clear signals (Fig. 4).

**Mean nucleotide/amino acid divergence of the study strains**

Mean nucleotide/amino acid diversities of the study strains with other CV-B3 and EV-B (CV-B4, E-7 and EV-B74) reference strains are shown in Table 2.

At the complete genome/polyprotein, P1, P2-P3 and 5’NCR levels, the phylogenetically separated groups of study strains, GVI (n=12) and GVII (n=3), showed mean nucleotide/amino acid sequence divergence of 14.4/2.1, 16.5/1.7, 13.1/2.4 and 15 %, respectively. The mean genomic divergence (nt/aa) within the study strains from GVII was found to be more (8.5/1.3 %) than that detected within GVI (1.8/0.8 %) (Table 2). No specific nucleotide/amino acid mutation(s) in the genomes of strains from AFP cases, with respect to the genomes of strains from asymptomatic contacts, were noted.

---

**Fig. 4.** Simplot and bootscan graphs constructed for the complete genome sequences of representative study strains of (a) GVI and (b) GVII in comparison with the consensus sequences of CV-B3/other EV-B reference strains, with the sequence of CV-A24 strain as an outgroup, using Simplot v3.5.1 with a window size of 500 nt and a step size of 20 nt.
Analysis of predicted structures of 5’NCR domains of the study strains

The secondary structures of the domains DI–DVII of the 5’NCR of the study strains were predicted using the RNA-structure algorithm (Reuter & Mathews, 2010). The DII domain (nt 90–185), implicated in regulation of the virus replication, was found to differ in its structural fold in the two genogroups, GVI and GVII, of the study strains (Fig. S3). In this domain region, the study strains from GVI were found to differ at 28 nucleotide positions with respect to the GVII strains. Fig. 5(a, b) shows the predicted structure of the DII domain of a strain representative of GVI and GVII, respectively. The predicted secondary structures of all the other domains [DI (nt 1–89), DIII (nt 185–240), DIV (nt 241–450), DV (nt 451–580), DVI (nt 581–623) and DVII (nt 624–647)] of the 5’NCRs of the study strains showed no significant differences (Fig. S4).

DISCUSSION

The present study reports the genomic characterization of CV-B3 strains detected in AFP cases and their asymptomatic contacts identified during 2009–2010 from different regions of the south-western states of India. It is evident from the phylogenetic analysis of full VP1 that Indian CV-B3 strains diverged into two distinguishable genogroups, GVI and a new genogroup, GVII, with respect to globally representative CV-B3 strains (Figs 1 and S1). Notably, monophyletic clustering of CV-B3 strains associated with both AFP and asymptomatic contacts of this study was observed in the dominant genogroup, GVI. This is suggestive of transmission of the CV-B3 infection from an index case to its asymptomatic contacts (Table 1). Furthermore, the genetic relatedness of the NIK strain (NIV099351LV53P4) with the strains from CK (Figs 1 and 2) may also indicate the possible transmission of the prevalent strain of CK across the region. The frequent travelling of populations in different geographical regions is reported to cause dispersal of CV-B3 strains (Tao et al., 2012). This could thus explain the occurrence of Indian strains from different regions in the two genogroups.

Only the P1-based tree was found to possess a topology matching that of the VP1-based tree and also to distinguish all the available CV-B3 strains into the appropriate genogroups (Figs 1 and 2a). Despite the discrepant tree topologies in case of the other genomic regions (P2-P3-3’NCR and 5’NCR), identification of most of the CV-B3 strains including the study strains in the appropriate genogroups may denote a relatively conserved evolution of CV-B3 genomes within the genogroup (Fig. 2b, c). The observation of the CV-B3 Uzbekistan strain (gb_AY896762) as an outlier of GVI in the complete genome phylogeny (Fig. 2d) may be attributed to the different recombination pattern noted in
the non-structural region (Fig. 2b, c) and is in agreement with results reported previously (Lukashev et al., 2005).

The mean nucleotide/amino acid sequence divergence between the two genogroups of the study strains was found to be noteworthy (14.4/2.1 %) (Table 2). Within the genogroups, the GVII strains were more divergent (nt/aa) than GVI strains, specifically in the 5’NCR (3.6 vs 0.9 %) and P2-P3 protein region (11.4/1.8 vs 2.2/0.6 %). In the P1 region, however, the nucleotide sequence divergence within GVII strains was higher (5.6 %) than that of the GVI strains (1.6 %), and the amino acid divergence in the GVII strains (0.4 %) was found to be less than that of GVI (1 %) (Table 2).

The clustering of the study strains along with CV-B3 reference strains in the P1 region (Fig. 2a) but closer to other EV-B types in the P2-P3-3’NCR and 5’NCR (Fig. 2b, c) illustrated the occurrence of recombination events between the CV-B3 strains and other EV-B types in the non-structural genomic regions. Similarity plot and Bootscan analyses across the complete genomes of each of the study strains revealed multiple sites of potential recombination (Figs 3 and S2). The breakpoints of recombination were identified to be distributed in the 5’NCR and throughout the P2-P3 region but not in the P1 region. Thus, the similarity of the non-structural genomic region of the study strains with multiple types (CV-B4, E-7 and EV-B74) indicates its susceptibility to inter-typic recombination. The differences in the recombination patterns identified between the GVI and GVII strains of this study suggest that the recombination events might have played a role in the emergence of the novel genogroup GVII of CV-B3. It should be noted that only the P1 domain of all the study strains showed significantly higher nucleotide/amino acid identity with CV-B3 reference strains than with other EV-B reference strains (CV-B4, E-7, and EV-B74) (84.4/97.9 vs 67.7/74.7 %) (Table 2). In contrast, the P2-P3 domain of the study strains showed higher nucleotide/amino acid divergence with CVB3 reference strains than with other EV-B reference strains (CV-B4, E-7 and EV-B74) (19.1/3.9 vs 14.9/2.5 %) (Table 2). In the 5’NCR of GVI/GVII strains, the mean nucleotide diversity from other CV-B3 and EV-B strains was found to be 12.7/14.7 and 11.8/13.3 %, respectively (Table 2). These findings are in agreement with the concept of independent evolution of structural and non-structural genomic fragments of EV-B types (Lukashev et al., 2005).

Structural alteration in the variable domain, DII, of the 5’NCR of CV-B3 strains has been described to play a role in virus processes such as translation and replication, known to contribute to the viral virulence (Dunn et al., 2003; Prusa et al., 2014). The difference in the predicted structures of the DII of 5’NCRs in the study strains from genogroups GVI and GVII (Figs 5 and S3) may be attributed to different parental origins, as was noted from the recombination patterns (Figs 3 and S2), which might further reflect the influence of the 5’NCR in the evolution of these genogroups.

In the present study, no specific mutations were found in AFP-associated strains. Furthermore, the secondary structure of the 5’NCR DII region of all of the strains revealed no specific differences between strains from AFP cases and asymptomatic contacts (Figs 5 and S3), thus suggesting a lack of association of the structural features with neurological manifestations. Similar findings have been reported through characterization based on the genomic diversity and structural alterations of EV-A71 strains from severe and mild disease patients (Ortner et al., 2009; Wang et al., 2012). Thus, the manifestation of CV-B3 infection might be driven by the host–virus interaction. Other factors such as viral load during the infection and host immune status may also influence the outcome of CV-B3 infections and need to be examined (Casadevall & Pirofski, 2014; Kishimoto et al., 2015).

To summarize, to the best of our knowledge, this is the first report on complete genomic analysis and phylogenetic characterization of CV-B3 strains recovered from AFP cases and their asymptomatic contacts from different regions of India. The identification of two genogroups co-circulating in India, one of them newly described, highlights the genetic divergence of CV-B3 in this particular geographical region, although no specific viral genetic mutation(s)/marker(s) could be linked to AFP. In view of these findings, further studies to identify the quasispecies nature of these viral infections and host genetic markers and examination of the host–virus interaction pathways involved in progression to AFP manifestation will be crucial.

METHODS

Specimen selection. Fifteen CV-B3 strains were detected on the basis of partial VP1 gene sequences in 10 AFP cases and five asymptomatic contacts, notified during 2009–2010 through the Polio Surveillance Project and which tested negative for poliovirus at the National Polio Laboratory, Bangalore (Table 1) (Laxmivandana et al., 2013). All of the strains were isolated from human rhabdomyosarcoma cell cultures inoculated with stool specimens from the AFP patients/asymptomatic contacts, aged from <1 to 13 years, identified in different regions of the south-western states of India.

RNA extraction and reverse transcription (RT)-PCR. The viral RNA was extracted from the cell culture supernatants using a MagMAX RNA Isolation kit (Ambion), according to the manufacturer’s instructions, on an automated RNA extractor (Applied Biosystems). The viral RNA was reverse transcribed and amplified using 10 overlapping primer pairs (Table S1) and a SuperScript III RT-PCR system with Platinum Taq (Invitrogen). The 5’ and 3’ ends of the genomes were determined by an RNA ligase-mediated RACE kit (Ambion). The amplicons were separated by electrophoresis in 1.5 % agarose gels containing ethidium bromide and visualized under a UV transilluminator.

Nucleotide sequencing and genetic analyses. The excised PCR products were purified using a QIAquick gel extraction kit (Qiagen). Sequencing of the products was carried out using an ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems) on an automated DNA sequencer (ABI PRISM 3100 Genetic Analyzer; Applied Biosystems).

Analysis of the nucleotide sequences obtained in this study was carried out using MEGA version 6.0 software package (Tamura et al., 2013). The complete genome sequences of the strains of this study
and EV-B prototype/reference strains were aligned using CLUSTAL W. Phylogenetic trees were reconstructed on the basis of nucleotide sequences using the neighbour-joining method and maximum composite likelihood model with 1000 bootstrap replications. Percentage nucleotide/ amino acid identity/divergence among the sequences was estimated using the p-distance model.

Recombination analysis. The alignments were also analysed for recombination using the SimPlot version 3.5.1 software package (Lole et al., 1999). SimPlot and Bootscan graphs were built with a window size of 500 nt and step size of 20 nt and the Bootscan analysis was run with a neighbour-joining tree algorithm (Kimura distance model) and 100 pseudo-replicates.

Prediction of RNA secondary structure. RNA secondary structure prediction was performed with RNAstructure software (http://rna.urmc.rochester.edu/RNAstructure.html) (Reuter & Mathews, 2010) using the graphical user interface and default thermodynamic parameter options, which included calculation of all structures within 10% free energy of the most stable structure.

ACKNOWLEDGEMENTS

As this study involved the use of cell culture isolates of viruses recovered from stool specimens of AFP cases and their contacts that were investigated at the National Polio Laboratory, the requirement of informed consent was waived and the study was approved by the Institutional Human Ethics Committee as per the Indian Council of Medical Research (ICMR) Guidelines (ICMR, 2006). The authors thank the ICMR, New Delhi, for supporting the first author (R. L.) with her Research Fellowship, Dr D.T. Mourya, Director, NIV, Pune, for constant support during this study and Mr Atul M. Walimbe for providing data on AFP cases. India, especially all technical staff of the National Polio Laboratory, Bangalore, for providing data on AFP cases.

REFERENCES


