Cyclovirus in nasopharyngeal aspirates of Chilean children with respiratory infections

Tung Gia Phan,1,2 Vivian Luchsinger,3 Luis F. Avendaño,3 Xutao Deng1 and Eric Delwart1,2

1Blood Systems Research Institute, San Francisco, CA 94118, USA
2Department of Laboratory Medicine, University of California at San Francisco, San Francisco, CA 94118, USA
3Programa de Virología, Facultad de Medicina, Universidad de Chile, Santiago, Chile

Some respiratory tract infections remain unexplained despite extensive testing for common pathogens. Nasopharyngeal aspirates (NPAs) from 120 Chilean infants from Santiago with acute lower respiratory tract infections were analysed by viral metagenomics, revealing the presence of nucleic acids from anelloviruses, adenovirus-associated virus and 12 known respiratory viral pathogens. A single sequence read showed translated protein similarity to cycloviruses. We used inverse PCR to amplify the complete circular ssDNA genome of a novel cyclovirus we named CyCV-ChileNPA1. Closely related variants were detected using PCR in the NPAs of three other affected children that also contained anelloviruses. This report increases the current knowledge of the genetic diversity of cycloviruses whose detection in multiple NPAs may reflect a tropism for human respiratory tissues.

Cycloviruses, members of a proposed genus within the family Circoviridae, have a circular ssDNA genome of approximately 2 kb (Li et al., 2010). Genetically highly diverse cycloviruses were initially found in the faeces of Pakistani children with and without acute flaccid paralysis (Victoria et al., 2009), in wild chimpanzees (Li et al., 2010) and in tissues of farm animals including cows, goats, bats and chickens (Ge et al., 2011; Li et al., 2010, 2011). Unexpectedly, other cyclovirus species have also been detected in insects, namely dragonflies and cockroaches (Dayaram et al., 2013; Padilla-Rodriguez et al., 2013; Rosario et al., 2011). In 2013, a cyclovirus species (CyCV-CN) was found initially using viral metagenomics and then by PCR in 4% of cerebrospinal fluid (CSF) specimens from Vietnamese children with unexplained central nervous system disorder, but not in CSF from patients with non-neurological problems, as well as in 4.2% of faeces from healthy Vietnamese children (Tan et al., 2013). CyCV-CN DNA was also detected in a throat swab (Tan et al., 2013). In this study, 58% of faecal specimens from pigs and poultry in Vietnam were also positive for the same cyclovirus, suggesting possible sources of human infection (Tan et al., 2013). A related cyclovirus was also detected in 10% of CSF samples and 15% of serum samples from adult patients with paraplegia (leg paralysis) from Malawi (Smits et al., 2013).

Nasopharyngeal aspirates (NPAs) from Chilean children less than 2 years old with acute lower respiratory infections were tested for respiratory syncytial virus (RSV), adenovirus, parainfluenza virus 1–3 and influenza A and B viruses by indirect immunofluorescence assays and virus isolation (Avendaño et al., 2003). From 1998 to 2000, a mean of 29% of acute lower respiratory infections samples were positive for RSV (Avendaño et al., 2003). To initiate the characterization of the viruses in non-reactive NPA samples, viral particles were enriched by filtration, and unprotected DNA and RNA were digested using a combination of nucleases enzymes (Victoria et al., 2009). The remaining nucleic acids were then extracted using a MagMAX Viral RNA Isolation kit (Life Technologies), which recovers both RNA and DNA. A DNA library was constructed using a ScriptSeq v2 RNA-Seq Library Preparation kit (Epicentre), which amplifies both RNA and DNA, and sequenced using the Illumina MiSeq platform. Viral sequences were identified using translated protein sequence similarity searches to annotated viral proteins available in GenBank (using BLASTX) and results were mapped using the NCBI Virus Taxonomy browser (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?name=Viruses). The study was approved by the University of California at San Francisco committee on human research.

A total of 120 respiratory specimens from Chilean children were analysed in 12 pools of 10 specimens using one Illumina MiSeq run of 250 base paired-ends. Viral sequence reads were identified with amino acid similarity >95% to
known viruses. Most numerous sequences were from anelloviruses, followed by enterovirus C, betacoronavirus, bocavirus 1, RSV, human adenovirus 3, enterovirus B, human rhinoviruses A and C, human parainfluenza 3, adenov-associated virus, human pneumovirus, human rhinovirus B and human parechovirus sequences (Table 1). Anelloviruses were identified in all but one pool (Table 1). Anelloviruses have been reported previously in human respiratory secretions (Burián et al., 2011; Jartti et al., 2012). Anelloviruses are usually considered commensal viruses (Okamoto, 2009a), although increased prevalence was found in bronchoalveolar lavage of children with acute exacerbation idiopathic pulmonary fibrosis (Woottton et al., 2011) and acute respiratory diseases (Maggi et al., 2003), and in lung tissues of pigs infected with known respiratory pathogens (Rammohan et al., 2012). Anellovirus plasma load is also increased in advanced AIDS (Li et al., 2013) and in immunosuppressed patients following organ transplantation (De Vlaminck et al., 2013). Increased anellovirus loads may reflect increased replication in immune cells stimulated by chronic inflammation, rather than indicating a direct pathogenic role (De Vlaminck et al., 2013). Detection of RSV, parainfluenza 3 and adenovirus in four, two and one pools, respectively, was probably the result of viral loads being too low for detection by immunofluorescence assays and cell culture (Avendaño et al., 2003). Except for anelloviruses and adenov-associated virus, all other viruses found have been associated with respiratory symptoms.

One sequence from one sample pool showed significant similarity to cyclovirus proteins (BLASTX E-score of 2 x 10^-7 to dragonfly cyclovirus, GenBank accession no. KCS12919). The full circular cyclovirus genome, referred to as CyCV-ChileNPA1, was then amplified using inverse PCR with specific primers designed from the Illumina-derived short sequence and directly Sanger sequenced by primer walking. Putative ORFs in the cyclovirus genome were predicted using the NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/orffinder) requiring an initiation Met codon.

The complete circular genome of CyCV-ChileNPA1 was 1790 bases, with a G + C content of 43 mol%. This genome consisted of two major ORFs encoding replication-associated protein (Rep) and capsid protein (Cap). The intergenic region was 247 bases and encoded a putative stem–loop structure with a stem length of 13 bases, predicted by the Mfold program (Zuker, 2003). Similarly to other cycloviruses, the highly conserved nonamer (TAGTATTAC) was found in the loop (Fig. 1a).

The International Committee on the Taxonomy of Viruses has also proposed a threshold of 75% nucleotide identity over the entire genome and 70% amino acid identity for the capsid protein. Rep showed the closest match (65%) to that of dragonfly CyV-8, whilst Cap shared a lower identity of 30% with that of the same virus. The higher level of nucleotide divergence of cap relative to rep was also observed with sequence alignments of CyCV-ChileNPA1 with its closest relatives (Fig. 1b). Such a high level of sequence divergence indicated that CyCV-ChileNPA1 may be considered a new species within the genus Cyclovirus.

CyCV-ChileNPA1 shared conserved Rep motifs (Dayaram et al., 2013; Rosario et al., 2012). Analysis of the deduced amino acid Rep sequences of CyCV-ChileNPA1 revealed three rolling-circle replication motifs I–III: FTxNN (FTWHD), YCSKxGX (YCSKSGE) and HLQGxxNL (HLQGFGCSL), respectively (Fig. 2). In the N terminus of cyclovirus Rep, two consensus high-affinity DNA-binding specificity determinants (SPDs), TxR for SPD-region 1 and PxR for SPD-region 2, were present (Dayaram et al., 2013; Londoño et al., 2010). CyCV-ChileNPA1 showed a mutated VxR for SPD-region 1 of unknown functional consequence (Fig. 2). The C-terminal region of the CyCV-ChileNPA1 Rep protein possessed ATP-dependent helicase motifs Walker A, B and C, or GxxGTGKS (GPPGTGKS), VIIDFYGW and ITSN, respectively (Fig. 2).

Sequence alignment was performed using CLUSTAL_X (Saitou & Nei, 1987). A phylogenetic tree with 100 bootstrap resamples of the alignment datasets was generated using MEGA5 and the neighbour-joining method (Tamura et al., 2011). Bootstrap values (based on 100 replicates) for each node are given for values >70% (Fig. 1c). Phylogenetic analysis confirmed the presence of a highly diverse cyclovirus species.

To determine the prevalence of this virus, a nested PCR assay was designed and used to test all 120 NPA samples. Primers ChilenNPA-F1 (5'-TGGGTCAAGCTATTACTGGGAG-3') and ChilenNPA-R1 (5'-ACTGAATGTCCGTCGCC-3') were used for the first round of PCR, and primers ChilenNPA-F2 (5'-CAGTGCCATAGTACAGTGGCACA-3') and ChilenNPA-R2 (5'-CTCCCCCTACTCAAAAAGAACCTCGGCT-3') for the second round of PCR, resulting in an expected amplion of ~310 bp. The PCR conditions were as follows: denaturation at 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 53 or 55 °C (for the first or second round, respectively) for 30 s and 72 °C for 1 min, a final extension at 72 °C for 10 min, and then held at 4 °C. Amplicons were then sequenced directly for identification. Three additional cases (CyCV-ChileNPA2–4) were positive. The genomic sequences of CyCV-ChileNPA2–4 were positive for the new cyclovirus, yielding a prevalence of 3.3% in the studied population (4/120). The full genomes of these three viruses were then acquired by overlapping PCR. The genomic sequences of CyCV-ChileNPA2–4 shared a high nucleotide identity of >99%, showing two, five and six nucleotide mutations compared with the CyCV-ChileNPA1 genome, respectively. All of these point mutations were synonymous except R25S and K93E in the ORF of CyCV-ChileNPA4 Rep.

The four CyCV-ChileNPA PCR-positive samples were then reanalysed using the same metagenomics approach but individually tagged to identify other viruses in these four samples. A total of 2697 unique sequence reads were generated. Three unique CyCV-ChileNPA1 reads were
Table 1. Distribution of sequence reads to different viral types/species in 12 NPA pools from Chile

<table>
<thead>
<tr>
<th>Virus</th>
<th>NPA pool no. and no. of reads</th>
<th>Total reads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1  2  3  4  5  6  7  8  9  10 11 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1246 305 1381 713 1561 246 1775 509 1136 293 1383 673 1327 383 1139 854 664 165 621 987 1033 709 1462 032</td>
<td></td>
</tr>
<tr>
<td>Anelloviruses</td>
<td>247 1115 605 43 81 907 950 575 58 123 645</td>
<td>5347 2531 1198 454</td>
</tr>
<tr>
<td>Enterovirus C</td>
<td>1759 772 12</td>
<td>12 116</td>
</tr>
<tr>
<td>Beta coronavirus</td>
<td>1198</td>
<td>1198</td>
</tr>
<tr>
<td>Bocavirus 1</td>
<td>326 12 13 13 119 4</td>
<td>159</td>
</tr>
<tr>
<td>Respiratory syncytial virus</td>
<td>10</td>
<td>93</td>
</tr>
<tr>
<td>Human adenovirus 3</td>
<td></td>
<td>93</td>
</tr>
<tr>
<td>Enterovirus B</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Human rhinovirus A</td>
<td>5 10</td>
<td>11 11 37</td>
</tr>
<tr>
<td>Human rhinovirus C</td>
<td>8 2</td>
<td>2 12</td>
</tr>
<tr>
<td>Human parainfluenza 3</td>
<td>2 8</td>
<td>10</td>
</tr>
<tr>
<td>Adeno-associated virus</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Human metapneumovirus</td>
<td>6 2</td>
<td>8</td>
</tr>
<tr>
<td>Human rhinovirus B</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Human parechovirus</td>
<td>2</td>
<td>2 4</td>
</tr>
<tr>
<td>Cyclovirus</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

T. G. Phan and others
generated from the NPA sample in which it was originally detected. No other viral sequences were detected. A total of 6362 unique reads were also generated from the other three samples positive only by PCR for the cyclovirus. All three samples contained anellovirus sequences (a total of 456 reads) and no other close matches to mammalian viruses. Anelloviruses are highly prevalent viruses present in many anatomical sites of different mammals (Okamoto, 2009b).

**Fig. 1.** Details of the novel cyclovirus CyCV-ChileNPA1. (a) Genome organization and its stem–loop structure. The locations of the putative rep and cap genes are indicated by arrows. (b) Pairwise sequence alignments of CyCV-ChileNPA1 with its closest relatives. The sequence nucleotide similarity (%) is indicated by the height of each point along the y-axis. The x-axis shows the nucleotide positions in the complete genome. (c) Phylogenetic trees generated with Rep and Cap proteins (concatenated) of cycloviruses. Bars, amino acid substitutions per position.
and are generally considered commensal infections. Anello- 
viruses have also been found in a significant minority of 
cases of idiopathic pulmonary fibrosis and in cases of acute 
lung injury (Wootten et al., 2011), are at higher prevalence in 
plasma and nasopharyngeal samples of febrile versus non- 
febrile cases (McElvania TeKippe et al., 2012) and are 
generally increased in the plasma of immunosuppressed 
individuals such as advanced AIDS patients (Li et al., 2013) 
or transplant recipients (De Vlaminck et al., 2013). A porcine 
anellovirus (torque teno sus virus species 1) has also been 
associated with porcine respiratory disease complex where it 
might exacerbate infections caused by porcine circovirus 2 
and the arterivirus porcine reproductive and respiratory 
disease symptom virus (Rammohan et al., 2012).

The detection of cyclovirus DNA in different human 
samples, including faeces, blood and CSF, and in the 
muscle tissues of farm animals suggests that cycloviruses 
may cause systemic infections in mammals (Li et al., 2010; 
Smits et al., 2013; Tan et al., 2012). The detection of 
cyclovirus DNA in NPAs (upper respiratory tract) of 
children with lower tract respiratory problems raises the 
possibility of a role for these viruses in respiratory illnesses. 
Further investigations of the host and tissue tropism, the 
transmission route(s) and any physiological consequences of 
human cyclovirus infections and possible interactions 
with anelloviruses are required.

Acknowledgements
We acknowledge NHLBI grant R01 HL105770 to E. L. D. and the 
Blood Systems Research Institute for sustained support.

References
for respiratory syncytial virus in infants hospitalized for acute lower 
respiratory infection in Chile (1989 to 2000). J Clin Microbiol 41, 
4879–4882.

Burián, Z., Szabó, H., Székely, G., Gyurkovits, K., Pankovics, P., 
teno midivirus (‘small anelloviruses’) in nasopharyngeal aspirates and three 

Dayaram, A., Potter, K. A., Moline, A. B., Rosenstein, D. D., Marinov, 
M., Thomas, J. E., Breitbart, M., Rosario, K., Arguello-Astorga, G. R. & 
Varsani, A. (2013). High global diversity of cycloviruses amongst 

De Vlaminck, I., Khush, K. K., Strehl, C., Kohli, B., Luikart, H., Neff, 
N. F., Okamoto, J., Snyder, T. M., Cornfield, D. N. & other authors 
(2013). Temporal response of the human virome to immunosuppres- 

(2011). Genetic diversity of novel circular ssDNA viruses in bats in 


Li, L., Kapoor, A., Siklas, B., Bamiude, O. S., Wang, C., Shaukat, S., 
Multiple diverse circoviruses infect farm animals and are commonly 

Delwarte, E. (2011). Possible cross-species transmission of circoviruses 
and cycloviruses among farm animals. J Gen Virol 92, 768–772.

Li, L., Deng, X., Linsuwannon, P., Bangsberg, D., Bwana, M. B., Hunt, 
P., Martin, J. N., Deeks, S. G. & Delwart, E. (2013). AIDS alters the 

Londoño, A., Riego-Ruiz, L. & Argüello-Astorga, G. R. (2010). DNA- 
binding specificity determinants of replication proteins encoded by 
eukaryotic ssDNA viruses are adjacent to widely separated RCR 

Maggi, F., Pifferi, M., Tempestini, E., Fornai, C., Lanini, L., Andreoli, 
E., Vatteroni, M., Presciutti, S., Pietrobelli, A. & other authors 
(2003). TT virus loads and lymphocyte subpopulations in children 

McElvania TeKippe, E., Wylie, K. M., Deych, E., Sodergren, G., 


