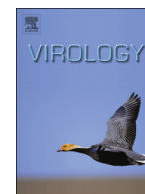




ELSEVIER

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Small circular single stranded DNA viral genomes in unexplained cases of human encephalitis, diarrhea, and in untreated sewage [☆]



Tung Gia Phan ^{a,b}, Daisuke Mori ^c, Xutao Deng ^a, Shaman Rajindrajith ^d, Udaya Ranawaka ^e, Terry Fei Fan Ng ^{a,b}, Filemon Bucardo-Rivera ^f, Patricia Orlandi ^g, Kamruddin Ahmed ^{c,h,i}, Eric Delwart ^{a,b,*}

^a Blood Systems Research Institute, San Francisco, CA 94118, USA

^b Department of Laboratory Medicine, University of California at San Francisco, San Francisco, CA 94118, USA

^c Department of Microbiology, Faculty of Medicine, Oita University, Yufu 879-5593, Japan

^d Department of Pediatrics, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

^e Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

^f Department of Microbiology, University of Leon, Nicaragua

^g Oswaldo Cruz Foundation, Manaus, Brazil

^h Research Promotion Institute, Oita University, Yufu 879-5593, Oita, Japan

ⁱ Department of Pathobiology and Medical Diagnostics, Faculty of Medicine, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia

ARTICLE INFO

Article history:

Received 3 February 2015

Returned to author for revisions

16 February 2015

Accepted 3 March 2015

Available online 1 April 2015

Keywords:

Cyclovirus

Gemycircularvirus

Encephalitis

Diarrhea

CRESS-DNA

ABSTRACT

Viruses with small circular ssDNA genomes encoding a replication initiator protein can infect a wide range of eukaryotic organisms ranging from mammals to fungi. The genomes of two such viruses, a cyclovirus (CyCV-SL) and gemycircularvirus (GemyCV-SL) were detected by deep sequencing of the cerebrospinal fluids of Sri Lankan patients with unexplained encephalitis. One and three out of 201 CSF samples (1.5%) from unexplained encephalitis patients tested by PCR were CyCV-SL and GemyCV-SL DNA positive respectively. Nucleotide similarity searches of pre-existing metagenomics datasets revealed closely related genomes in feces from unexplained cases of diarrhea from Nicaragua and Brazil and in untreated sewage from Nepal. Whether the tropism of the cyclovirus and gemycircularvirus reported here include humans or other cellular sources in or on the human body remains to be determined.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Cyclovirus is a recently proposed genus of the *Circoviridae* family with small ssDNA circular genome of approximately 2-kb (Li et al., 2010a). Cycloviruses are a sister clade to the circoviruses, a genus known to infect a wide variety of birds (Todd, 2004) and mammals (Li et al., 2013, 2011). A third genus, krikovirus, was recently proposed from genomes detected in mosquitoes (Garigliany et al., 2015) and bat feces (Li et al., 2010b; Lima et al., 2015). The *Circoviridae* family is part of a much more diverse group of viruses with circular, replication initiator protein (Rep) encoding, single stranded DNA (CRESS-DNA) genomes, found in a wide range of hosts and environment (Delwart and Li, 2012; Rosario et al., 2012b). CRESS-DNA viruses, encoding only Rep and capsid (Cap) proteins, have the smallest genomes of autonomously

[☆]The GenBank accession numbers for new cycloviruses and gemycircularviruses are, KP133075, KP151567, KP133076–KP133080.

* Corresponding author at: Blood Systems Research Institute, San Francisco, CA 94118, USA. Fax: +1 415 567 5899.

E-mail address: delwarte@medicine.ucsf.edu (E. Delwart).

replicating eukaryotic viruses. Cyclovirus genomes, initially found in the feces of Pakistani children with and without acute flaccid paralysis (Li et al., 2010a) have also been reported in the feces of bats (Ge et al., 2011; Li et al., 2010b), poultry (Li et al., 2010a, 2011; Tan le et al., 2013) and in meat samples of various farm animals in Pakistan and Nigeria (Li et al., 2010a, 2011). Cyclovirus genomes have also been detected in the abdomen of dragonflies and in the Florida cockroach (Dayaram et al., 2013; Padilla-Rodriguez et al., 2013; Rosario et al., 2011). In 2013 a cyclovirus was also reported in 4% of cerebrospinal fluid (CSF) from south and central Vietnamese children with unexplained central nervous system disorders, and in 4.2% of healthy children's feces (Tan le et al., 2013). Fifty eight percent of feces from nearby pigs were also positive for the same cyclovirus DNA (Tan le et al., 2013). A later study did not detect this virus in CSF of similar patients from northern Vietnam, Cambodia, Nepal and The Netherlands (Le et al., 2014). A recent study detected this cycloviral DNA in fecal samples from healthy children from Madagascar and from healthy pigs in Cameroon (Garigliany et al., 2014). In 2013 a related cyclovirus was found in 10% of CSF and 15% of sera from 58 paraplegia (leg paralysis) Malawian adult patients (Smits et al., 2013). These CSF associated

cyclovirus genomes phylogenetically clustered together along with Tunisian strains from the feces of children with non-polio acute flaccid paralysis (Li et al., 2010a, 2011). Recently another cyclovirus species was detected in 3% of nasopharyngeal aspirates from Chilean infants with unexplained acute lower respiratory tract infections (Phan et al., 2014). Detection of cyclovirus DNA in various tissue samples, human feces, CSF, blood, and respiratory secretion, and in farm animal meat suggests systemic infections although with the caveat that replication in mammalian cells and sero-conversion remains to be demonstrated. The detection of cyclovirus sequences by PCR in sewage water from cities in the US reflects the presence of cycloviruses in industrialized countries (Blinkova et al., 2009).

Recently, another distinct group of CRESS-DNA genomes was classified under the proposed genus name of Gemycircularvirus (Rosario et al., 2012a; Sikorski et al., 2013). The members of this proposed genus are also called myco-like viruses because of their overall genome similarity to those of Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1 (SsHADV-1), the first DNA virus found replicating in fungi (Yu et al., 2010). Other gemycircularvirus genomes have also been identified in feces of different animals (Sikorski et al., 2013; van den Brand et al., 2012), cassava infecting fungi (Dayaram et al., 2012), the body of insects (Dayaram et al., 2015; Ng et al., 2011; Rosario et al., 2012a), and sewage (Kraberger et al., 2015). Gemycircularvirus genomes were also recently reported in blood from a patient with multiple sclerosis and asymptomatic cattle (Lamberto et al., 2014).

Here we report on the detection of both cyclovirus and gemycircularvirus DNA in the CSF of encephalitis patients, in fecal samples of unexplained human diarrhea, as well as in untreated sewage from Kathmandu, Nepal.

Results

The genomes of novel cyclovirus and gemycircularvirus were identified by deep sequencing the CSFs from 62 Sri Lankan patients with unexplained encephalitis (*Materials and Methods*). These CSF samples had previously tested negative for a wide range of pathogens (*Materials and Methods*). Using a BLASTx cutoff of E scores $< 10^{-5}$ one pool of five CSF samples containing 1.42×10^6 sequence reads showed a single read of human enterovirus B, 247 reads related to cycloviruses, and 6 reads related to gemycircularvirus. Total nucleic acids of the individual samples within the pool containing these sequences were extracted and tested for the cyclovirus and gemycircularvirus sequences by PCR—identifying two individual CSF samples. The complete genomes of these two CRESS-DNA viruses were then acquired by inverse PCR and Sanger sequencing.

The circular genome of the cyclovirus, referred to as CyCV-SL, was 1710 bases, with a GC content of 41%, including the major ORFs for Rep and Cap (GenBank KJ831064). The intergenic region is 207-bp and contains a stem loop structure with a stem of 13-bp. Similar to other cycloviruses, the highly conserved nonamer [TAGTATTAC] was found in the loop (Fig. 1A). The Rep protein was 276-aa in length and showed the closest identity (71%) to that of cyclovirus-7 from the abdomen of a dragonfly (GenBank KC512919) with the next closest Rep from cyclovirus PK5006 found in human feces (GenBank GQ404844) at 70% identity. The Cap protein (230-aa) shared highest identity of 39% to dragonfly cyclovirus-7, with the next closest Rep from cyclovirus PK6197 found in human feces (GenBank GQ404848) at 36% identity. Analysis of the deduced amino acid Rep sequences of human CyCV-SL revealed three rolling circle replication motifs I-III, CFTWNN, IHLQG and YCSK, respectively [conserved amino acid in boldface] (Rosario et al., 2012b) (Fig. 1B). In the N-terminus of

the 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; Londono et al., 2010). The CyCV-SL Rep showed a mutated VxR for SPD-region 1 (Fig. 1B). The C-terminal region of the Rep possessed ATP-dependent helicase motifs Walker A (GPPGSGKS), B (IIDDF) and C (UTSN) [conserved amino acid in boldface] (Rosario et al., 2012b) (Fig. 1B).

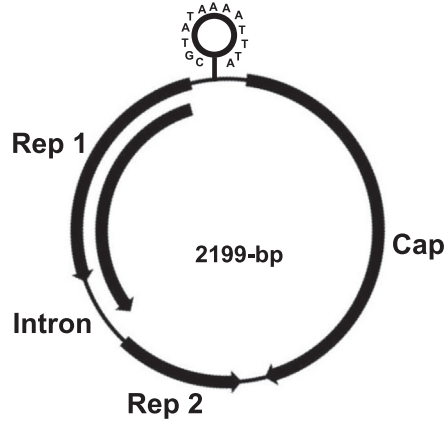
Phylogenetic analysis confirmed that CyCV-SL is distinct from other known members of the *Cyclovirus* genus (Fig. 1C). A recent ICTV proposal has set a genome identity of $> 80\%$ for members of the same circovirus or cyclovirus species. Our pair-wise genetic analysis revealed that CyCV-SL shared only 60% nucleotide identity over the entire genome and 39% amino acid identity for the capsid protein to the closest cyclovirus genome. CyCV-SL is therefore proposed as prototype of a novel species in the *Cyclovirus* genus.

The genome of a new gemycircularvirus which we named GemyCV-SL1 was similarly amplified and sequenced from another CSF sample yielding a 2199 bases genome, including major Rep1 (194-aa), minor Rep2 (112-aa) and Cap (322-aa) (GenBank KP133075). The intergenic region (127-nt) contained a stem loop structure having an atypical nonamer [TAAATTTA] (Fig. 2A) (Sikorski et al., 2013). Similar to other members of the *Gemycircularvirus* genus GemyCV-SL1 had a unique intron (140-nt) between two Rep ORFs. Intron splicing resulted in a single Rep ORF 320-aa in length. Rep and Cap proteins showed the highest identity (56% and 31%) to those of gemycircularvirus HCB18.215 found in healthy cattle blood (GenBank LK931483) (Lamberto et al., 2014). The spliced Rep protein in GemyCV-SL1 possessed three rolling circle replication motifs I-III (LLTYSG, THLHA and YACK) [conserved amino acid in boldface] (Rosario et al., 2012b) (Fig. 2B). Genetic analysis of the C-terminal region of the spliced Rep also detected three ATP-dependent helicase motifs Walker A (GPSRMGKT), B (IFDDL) and C (WISN) [conserved amino acid in boldface] (Rosario et al., 2012b) (Fig. 2B). Phylogenetic analysis confirmed GemyCV-SL1 to be related to gemycircularvirus HCB18.215 (Fig. 2C). GemyCV-SL1 therefore represents a new candidate species in the *Gemycircularvirus* genus.

Two nested PCR assays were developed to determine the prevalence of these viruses in the Sri Lankan CSF samples. In order to confirm the presence of these two novel viruses in the Sri Lankan CSF samples and exclude the possibility of local DNA contamination, the same two PCR protocols were also used by an independent lab at Oita University, Japan. This lab had previously extracted the same 62 CSF samples by QIAamp Viral RNA Mini Kit (Qiagen) and tested them for nucleic acids of known viruses (Mori et al., 2013). The presence of both CyCV-SL and GemyCV-SL1 DNA were confirmed by PCR in the same two samples. The Oita University lab also PCR screened for these two genomes using PCR in another set of 139 CSF samples previously extracted by QIAamp Viral RNA Mini Kit (Qiagen) from similar Sri Lankan patients with unexplained encephalitis. No other CyCV-positive sample was detected but GemyCV DNA was amplified in 2 out of these 139 nucleic acids extracts, yielding a total frequency of GemyCV DNA detection of 1.5% (3/201). The full genomes of these two GemyCVs were then acquired by inverse PCR and direct Sanger sequencing. Nucleotide alignment showed that the genomic sequences of GemyCV-SL2 and -SL3 (Genbank KP133076 and KP133077) showed three and six nucleotide mutations relative to the GemyCV-SL1 genome, respectively.

Nucleotide sequences closely related to CyCV-SL and GemyCV-SL were then sought using MegaBlast in 135 MiSeq and 60 pyrosequencing (454 Roche) runs previously generated in the San Francisco laboratory with other human, animals, and environmental samples. In one MiSeq run, four CyCV-SL sequence matches were found among 6.2×10^5 reads generated from one pool of five

A



B

GemyCV-SL1	M-TFHCNARYE FLTYPx CGELDPWKVSNHLGELRGECIVAREAHADGG xHxHA FVDFGRKFRSRRTSIFD
Cattle HCB18.215	.S..RFY...A...A...G...FTIVS.I.D.G...IG...S...R.Q...AAV..
	RCR motif 1 YAxK RCR motif 2
GemyCV-SL1	VDNHHPNVSATHSTPREGFDY YAxK OGDIVAGGLERPEQSDRVVVRSTGWHDIVAATSREEFFELLMADHPR
Cattle HCB18.215	.GGF...I.I.KRD.QVHY...I...V.....AADASAGVENKH.Q..L.PT.A...AVRDL...
	RCR motif 3
GemyCV-SL1	TLCTCFAQLSKYADWRYREDPEPYRSPDSVVIDASAFPELEWPHRALQPAHVRRGRRGQSLVLY GPSRM
Cattle HCB18.215	L.L.S.PS.E...H..V..D...HDPEYEF.T.Q...NQ.AEDYIGWDATV...K...T
	Walker-A Walker-B
GemyCV-SL1	GKTL WARS LGSHAYFGGLFSLQEETDNVDY IFDDI NGGIPFFPQYKWWLGHQGFYATDKYKGGKLVHW
Cattle HCB18.215	...M...AK.....E.SLED.K...V...JA...LK.....S...N.LD..S...K...AIK.
	Walker-A Walker-B
GemyCV-SL1	GRPA WISN KDPREEHGADYDNLNANCLFIYVDTPIAWEIE
Cattle HCB18.215	.K.C...EY.-ITE.V.Q..MDK..I.VE.T..LYS---
	Motif-C

C

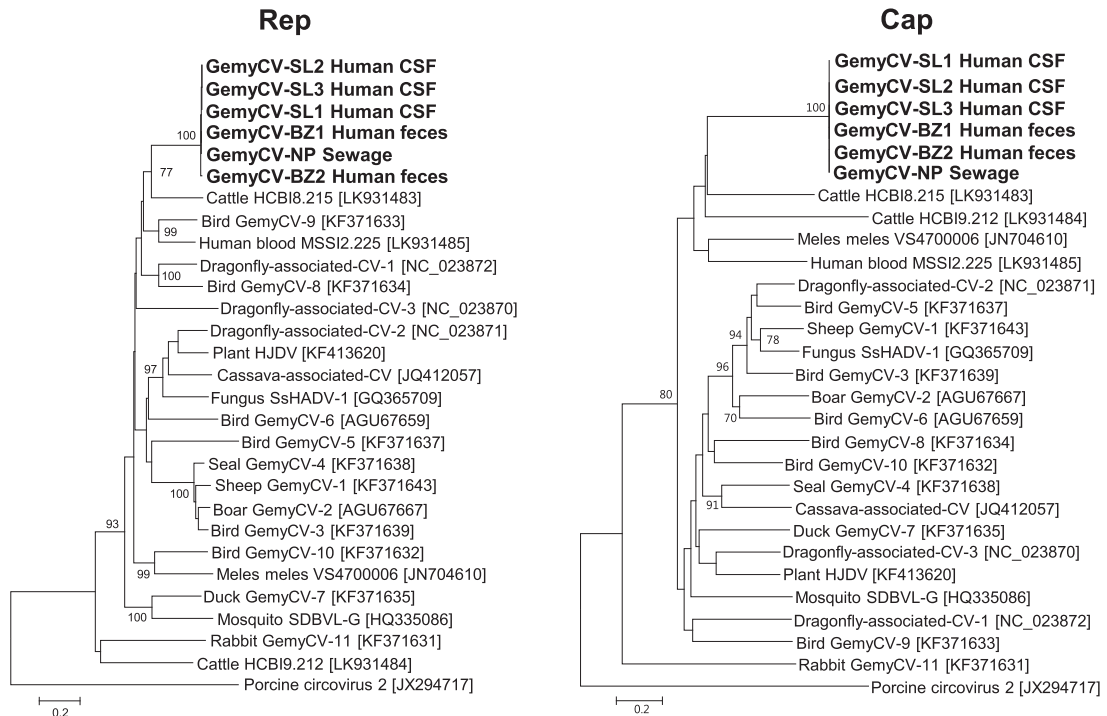


Fig. 2. New gemycircularvirus. (A) Genome organization and its stem loop structure. (B) Alignment of Rep proteins of the newly identified Sri Lankan gemycircularvirus and cattle gemycircularvirus. Conserved motifs were shown. (C) Phylogenetic trees generated with Rep and Cap proteins of gemycircularvirus detected in Sri Lanka patients with unexplained encephalitis, Brazilian patients with unexplained diarrhea, Nepalese sewage, and all genetically-related gemycircularviruses in the *Gemycircularvirus* genus. The scale indicates amino acid substitutions per position.

Table 1
Detection of cyclovirus and gemycircularvirus reads in deep sequencing of pools and individual samples.

Source	Genome sequenced	Country	PCR detection	Miseq reads	GemyCV reads	CyCV reads
CSF pool		Sri Lanka	+	1429342	6	247
CSF	CyCV-SL	Sri Lanka	+	322060		438
CSF	GemyCV-SL1	Sri Lanka	+	2925070	72	
CSF	GemyCV-SL2	Sri Lanka	+	866078	116	
CSF	GemyCV-SL3	Sri Lanka	+	1269398	14	
Feces pool		Nicaragua	+	625641		4
Feces	CyCV-NI	Nicaragua	+	753634		
Feces pool		Brazil	+	3214352	14	
Feces	GemyCV-BZ1	Brazil	+	3652720	4	
Feces	GemyCV-BZ2	Brazil	+	753614		

unexplained human diarrhea samples from Nicaragua (Bucardo et al., 2014). In another Miseq run, 14 GemyCV-SL sequences were found in 3.2×10^6 reads generated from one pool of five unexplained human diarrhea feces from Brazil. A total of 83 GemyCV-SL sequences were also found among 5915 reads generated by 454 pyrosequencing from raw sewage from Nepal (Ng et al., 2012). Total nucleic acids of individual feces samples within the Nicaraguan and Brazilian diarrhea sample pools were extracted and each was tested by PCR. One Nicaraguan and two Brazilian samples were found to be PCR positive. The DNA from these three samples were then used to generate the complete genomes by inverse PCR yielding Nicaraguan cyclovirus (CyCV-NI, 1710 bases, Genbank KP151567), and two Brazilian gemycircularvirus (GemyCV-BZ1 and GemyCV-BZ2, both 2199 bases, Genbank KP133078 and KP133079). The Nepalese sewage gemycircularvirus (GemyCV-NP, 2199 bases, Genbank KP133080) was also acquired by inverse PCR and Sanger sequencing. Nucleotide alignment showed the CyCV-NI genome to differ by 72 nucleotide mutations relative to CyCV-SL1 (95.8% identity). The genomes of GemyCV-BZ1 and -BZ2, and GemyCV-NP showed 32, 31, 29 nucleotide mutations to the GemyCV-SL1. The two Brazilian GemyCV genomes in two individual feces samples differed by 26 nucleotides.

To exclude the possibility of contamination from the DNA extraction method used we re-extracted all of the originally cycloviruses and gemycircularviruses positive CSF, feces and sewage samples (except GemyCV-SL2 and -SL3 as only previously extracted DNA was available) using phenol-chloroform extraction and confirmed the presence of cyclovirus and GemyCV by PCR.

To determine whether other viruses were also present in the samples containing cyclovirus and gemycircularvirus which might account for the diseased state of the patients their nucleic acids were individually deep sequenced and analyzed against all eukaryotic viruses (Table 1). The Sri Lankan CSF sample, in which the cyclovirus was first identified and sequenced, showed 438 CyCV-SL reads out of 3.2×10^5 reads. No other viral sequences were identified. The three individual Sri Lankan CSF samples positive for GemyCV by PCR all yielded GemyCV reads ($n=14-116$) and no other virus except for anelloviruses in two samples. No cyclovirus sequences were detected in the individual Nicaraguan feces positive by PCR possibly due to a low viral load as reflected by the few sequence reads ($n=4$) detected in the previously sequenced pool. 1185 human adenovirus 41 sequences read were detected in that fecal sample. For the two individual Brazilian fecal samples positive for GemyCV by PCR one yielded 4 GemyCV-BZ1 reads and the other none. No other eukaryotic viruses were detected (Table 1).

To investigate the possibility of a fungal origin for GemyCV, we used semi-nested PCR with universal fungal ITS2/5.8 S rDNA primers on the three GemyCV-positive CSF samples and five GemyCV-negative CSF samples from Sri Lanka. The three GemyCV-positive CSF samples were negative for fungal DNA while one of five GemyCV-negative CSF samples was positive. Sequence analysis of the amplicon indicated the presence of *Cryptococcus victorialis* DNA, a widely distributed fungal species (Danesi et al., 2014; Gallardo et al., 2014; Montes et al., 1999).

Discussion

We describe the genomes of a cyclovirus and gemycircularvirus in CSF from unexplained cases of encephalitis in Sri Lanka. Closely related viruses were also detected in unexplained human diarrhea samples from Nicaragua and Brazil as well as in untreated sewage from Nepal. In all but one case no other eukaryotic sequences were detected by deep sequencing indicating a potential role for these viruses in these patients' symptoms. In one sample adenovirus 41 was detected as a likely cause of the patient's diarrhea (Cruz et al., 1990; Sdiri-Loulizi et al., 2008, 2009).

While identifying CRESS-DNA genomes in CSF, a site considered sterile, supports the possibility of replication in the human host; however, reports confirming replication in mammalian cells or of sero-conversion to these viruses are still lacking. Because prior reports have detected distinct cycloviral genomes in the abdomen of insects (Dayaram et al., 2013; Padilla-Rodriguez et al., 2013; Rosario et al., 2011) alternative explanations should be considered to explain their detection in CSF. Cycloviruses may have a wide tropism with some species infecting mammals while others infect insects. Alternatively some cyclovirus species may be arboviruses capable of replication in both insects and mammals infected through insect bites. Cyclovirus DNA in mammalian fecal samples may also be the result of consuming food contaminated with cyclovirus as observed for chicken anemia virus DNA in human feces (Chu et al., 2012; Phan et al., 2012) and insect viruses in feces of insectivorous bats (Donaldson et al., 2010; Ge et al., 2012; Li et al., 2010b; Reuter et al., 2014). Cycloviruses in sites considered sterile such as CSF and plasma may also conceivably result from contamination during sample collection with the needle introducing a small skin plug into the collected fluid. The detection of cyclovirus DNA (Foulongne et al., 2012) as well as a gyrovirus, polyomaviruses, and papillomavirus DNA on human skin swabs attest to the frequent presence of numerous viral genomes on the skin surface (Ekstrom et al., 2013; Foulongne et al., 2012; Hannigan and Grice, 2013; Sauvage et al., 2011; Schowalter et al., 2010).

The origin of the gemycircularvirus DNA found here in CSF and recently described in human and bovine plasma (Whitley et al., 2014) therefore remains uncertain. To date the only cellular host identified is for the prototype of this group, SsHADV-1, the first DNA virus shown to replicate in fungi (Yu et al., 2010). The cellular hosts of all other gemycircularviruses, many detected in feces of mammals (Sikorski et al., 2013; van den Brand et al., 2012), fungi-infected plants (Dayaram et al., 2012), insects (Dayaram et al., 2015; Ng et al., 2011; Rosario et al., 2012a), and sewage (Krabberger et al., 2015) remains uncertain. As is the case for cycloviruses, the detection of gemycircularvirus genome in mammalian feces, blood, and CSF, may reflect genuine viral replication in humans or alternatively fungal infection releasing virus into the blood stream, fungi or fungi-infected plants in the diet, contamination from the surface of the skin during phlebotomy, or even contamination from particles floating in air (Whon et al., 2012). Whether any gemycircularvirus actually replicates in mammalian cells or

simply reflects the presence of other cellular organisms remains uncertain despite their genomes' detection in "sterile" sites.

Among CRESS-DNA viruses only members of the circovirus genus, such as pig and canine circoviruses (Li et al., 2013; Todd et al., 2001), have been shown to replicate in mammals. The cellular hosts of the large majority of CRESS-DNA encoded viruses, recently brought to light using deep sequencing methods, remain unknown (Delwart and Li, 2012; Rosario et al., 2012b). Further studies are required to assign cellular hosts to cycloviruses and gemycircularviruses detected in human feces and tissues.

Materials and methods

Clinical samples

CSF samples were collected from patients with encephalitis in Sri Lanka. These included 29 children aged between 2 and 144 months old and 33 adults between 15 and 72 years old. These samples had pre-tested negative for bacteria, herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus, dengue virus, Japanese encephalitis virus, rubella virus, west Nile virus, yellow fever virus, tick-borne encephalitis virus, nipah virus, measles virus, mumps virus, parainfluenza virus, respiratory syncytial virus, metapneumovirus, chikungunya virus, sindbis virus, semliki forest virus, eastern equine encephalitis virus, western equine encephalitis virus, poliovirus, coxsackievirus, echovirus, enterovirus, lyssaviruses, chandipura virus, bocavirus, rotavirus, astrovirus, norovirus, parechovirus, and adenovirus (Mori et al., 2013).

Viral metagenomics

The 62 CSF samples from Sri Lankan patients were first analyzed in minipools of 4–5 samples using a metagenomics approach. Total nucleic acids were extracted using a Maxwell[®] 16 automated extractor (Promega) from supernatant of CSF spun 15 min in a microfuge. cDNA and DNA were then generated by using random RT-PCR followed by the use of the Nextera[™] XT Sample Preparation Kit (Illumina) to construct a DNA library. One Illumina MiSeq run of 250 bases paired-end reads yielded an average of 2.56 million sequence reads from each pool of clinical samples. After quality control removal of duplicate reads and reads shorter than 50 bases followed by de novo assembly, viral sequences were identified through translated protein sequence similarity search (BLASTx) to annotated viral proteins available in GenBank's viral RefSeq database.

Genome acquisition of novel circular viruses

The complete genomes of circular viruses were identified and recovered by *de novo*-assembly, inverse PCR and Sanger sequencing. Putative ORFs in the circular genomes were identified using the NCBI ORF finder. The secondary structure of intergenic regions was predicted using the Mfold program (Zuker, 2003). All alignments and phylogenetic analyses were based on the translated amino acid sequences. Sequence alignment was performed using CLUSTAL X with the default settings (Saitou and Nei, 1987). Sequence identities were determined using BioEdit. Phylogenetic analyses were performed using Rep and Cap proteins of novel circular viruses and other genetically-close relatives. Phylogenetic trees with 100 bootstrap resamples of the alignment data sets were generated using the neighbor-joining method and visualized using the program MEGA version 5 (Tamura et al., 2011). Bootstrap values (based on 100 replicates) for each node are shown if > 70%.

PCR assays for cyclovirus and gemycircularvirus

In order to understand the frequency of cyclovirus and gemycircularvirus, we designed two sets of primers to amplify segments of the viral genomes and used nested-PCR to screen other CSF samples from Sri Lanka.

For cyclovirus, primers SLCyV-F1 (5'-TCC GTC CCA CCA TAG TCC TCT-3') and SLCyV-R1 (5'-AAG CAT CTC CAT AAC TCA ATC CAT CT-3') were used for the first round of PCR, and primers SLCyV-F2 (5'-TTG ATT GGT TTG TTG CTT TTG CTT CTT CT-3') and SLCyV-R2 (5'-GTA GCC AAG GAC ATC GCT CAG A-3') for the second round of PCR.

For gemycircularvirus, primers SLGemy-F1 (5'-GTG GTA ATG GTC GTC GGT ATT C-3') and SLGemy-R1 (5'-CCT CAT CAT TCG TAG TAA GCA ATC TCA-3') were used for the first round of PCR, and primers SLGemy-F2 (5'-AGT CCT GAA TGT TTC CAC TCG-3') and SLGemy-R2 (5'-CAA GCG TTC CCT CGA AAA TGA C-3') for the second round of PCR.

The PCR conditions in these two assays were: 95 °C for 5 min, 35 cycles 95 °C for 30 s, 50 °C (for the first or second round) for 30 s and 72 °C for 1 min, a final extension at 72 °C for 10 min, resulting in an expected amplicon of 255-bp for cyclovirus and 476-bp for gemycircularvirus.

PCR assay for fungi

Semi-nested PCR with universal primers for fungal amplification were performed. Primers ITS1 (5'-TCC GTA GGT GAA CCT GCG G-3') and ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') were used for the first round of PCR. Primers ITS86 (5'-GTG AAT CAT CGA ATC TTT GAA C-3'), and ITS4 were used for the second round of PCR. The PCR condition was followed as previously published, producing an expected amplicon of 280-bp (Ferrer et al., 2001).

Funding

This work was supported by NHLBI Grant R01 HL105770 and support from the Blood Systems Research Instituteto E. D. and a Research Fund at the Discretion of the President, Oita University (Grant no. 610000-N5021) to K.A.

References

- Blinkova, O., Rosario, K., Li, L., Kapoor, A., Slikas, B., Bernardin, F., Breitbart, M., Delwart, E., 2009. Frequent detection of highly diverse variants of cardiiovirus, cosavirus, bocavirus, and circovirus in sewage samples collected in the United States. *J. Clin. Microbiol.* 47, 3507–3513.
- Bucardo, F., Reyes, Y., Svensson, L., Nordgren, J., 2014. Predominance of norovirus and sapovirus in Nicaragua after implementation of universal rotavirus vaccination. *PLoS One* 9, e98201.
- Chu, D.K., Poon, L.L., Chiu, S.S., Chan, K.H., Ng, E.M., Bauer, I., Cheung, T.K., Ng, I.H., Guan, Y., Wang, A., Peiris, J.S., 2012. Characterization of a novel gyrovirus in human stool and chicken meat. *J. Clin. Virol.: Official Publ. Pan Am. Soc. Clin. Virol.* 55, 209–213.
- Cruz, J.R., Caceres, P., Cano, F., Flores, J., Bartlett, A., Torun, B., 1990. Adenovirus types 40 and 41 and rotaviruses associated with diarrhea in children from Guatemala. *J. Clin. Microbiol.* 28, 1780–1784.
- Danesi, P., Furnari, C., Granato, A., Schivo, A., Otranto, D., Capelli, G., Cafarchia, C., 2014. Molecular identity and prevalence of *Cryptococcus* spp. nasal carriage in asymptomatic feral cats in Italy. *Med. Mycol.* 52, 667–673.
- Dayaram, A., Opong, A., Jaschke, A., Haddfield, J., Baschiera, M., Dobson, R.C., Offei, S. K., Shepherd, D.N., Martin, D.P., Varsani, A., 2012. Molecular characterisation of a novel cassava associated circular ssDNA virus. *Virus Res.* 166, 130–135.
- 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 dragonflies. *J. Gen. Virol.* 94, 1827–1840.
- Dayaram, A., Potter, K.A., Pailles, R., Marinov, M., Rosenstein, D.D., Varsani, A., 2015. Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. *Infect. Genet. Evol.: J. Mol. Epidemiol. Evolut. Genet. Infect. Dis.* 30, 278–287.
- Delwart, E., Li, L., 2012. Rapidly expanding genetic diversity and host range of the Circoviridae viral family and other Rep encoding small circular ssDNA genomes. *Virus Res.* 164, 114–121.

- Donaldson, E.F., Haskew, A.N., Gates, J.E., Huynh, J., Moore, C.J., Frieman, M.B., 2010. Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *J. Virol.* 84, 13004–13018.
- Ekstrom, J., Muhr, L.S., Bzhalava, D., Soderlund-Strand, A., Hultin, E., Nordin, P., Stenquist, B., Paoli, J., Forslund, O., Dillner, J., 2013. Diversity of human papillomaviruses in skin lesions. *Virology* 447, 300–311.
- Ferrer, C., Colom, F., Frases, S., Mulet, E., Abad, J.L., Alio, J.L., 2001. Detection and identification of fungal pathogens by PCR and by ITS2 and 5.8 S ribosomal DNA typing in ocular infections. *J. Clin. Microbiol.* 39, 2873–2879.
- Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Gouilh, M.A., Pariente, K., Segondy, M., Burguiere, A., Manuguerra, J.C., Caro, V., Eloit, M., 2012. Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. *PLoS One* 7, e38499.
- Gallardo, G., Ruiz-Moyano, S., Hernandez, A., Benito, M.J., Cordoba, M.G., Perez-Navado, F., Martin, A., 2014. Application of ISSR-PCR for rapid strain typing of *Debaryomyces hansenii* isolated from dry-cured Iberian ham. *Food Microbiol.* 42, 205–211.
- Garigliani, M.M., Böstler, J., Jöst, H., Badusche, M., Desmecht, D., Schmidt-Chanasit, Cadar, D., 2015. Characterization of a novel circovirus in *Aedes vexans* mosquitoes from Germany: evidence for a new genus within the family *Circoviridae*. *J. Gen. Virol.* 96, 915–920.
- Garigliani Mutien-Marie, Hagen Ralf Matthias, Frickmann Hagen, May Jürgen, Schwarz Georg Norbert, Perse Amanda, et al., 2014. Cyclovirus CyCV-VN species distribution is not limited to Vietnam and extends to Africa. *Scientific Reports* 4, Article number: 7552. <http://dx.doi.org/10.1038/srep07552>.Ge.
- Ge, X., Li, J., Peng, C., Wu, L., Yang, X., Wu, Y., Zhang, Y., Shi, Z., 2011. Genetic diversity of novel circular ssDNA viruses in bats in China. *J. Gen. Virol.* 92, 2646–2653.
- Ge, X., Li, Y., Yang, X., Zhang, H., Zhou, P., Zhang, Y., Shi, Z., 2012. Metagenomic analysis of viruses from bat fecal samples reveals many novel viruses in insectivorous bats in China. *J. Virol.* 86, 4620–4630.
- Hannigan, G.D., Grice, E.A., 2013. Microbial ecology of the skin in the era of metagenomics and molecular microbiology. *Cold Spring Harb. Perspect. Med.* 3, a015362.
- Kraberger, S., Arguello-Astorga, G.R., Greenfield, L.G., Galilee, C., Law, D., Martin, D. P., Varsani, A., 2015. Characterisation of a diverse range of circular replication-associated protein encoding DNA viruses recovered from a sewage treatment oxidation pond. *Infect. Genet. Evol. J. Mol. Epidemiol. Evolut. Genet. Infect. Dis.* 31C, 73–86.
- Lamberto, I., Gunst, K., Muller, H., Zur Hausen, H., de Villiers, E.M., 2014. Mycovirus-like DNA virus sequences from cattle serum and human brain and serum samples from multiple sclerosis patients. *Genome Announc.* 2, e00848-14. <http://dx.doi.org/10.1128/genomeA00848-14>.
- Le, V.T., de Jong, M.D., Nguyen, V.K., Nguyen, V.T., Taylor, W., Wertheim, H.F., van der Ende, A., van der Hoek, L., Canuti, M., Crusat, M., Sona, S., Nguyen, H.U., Giri, A., Nguyen, T.T., Ho, D.T., Farrar, J., Bryant, J.E., Tran, T.H., Nguyen, V.V., van Doorn, H.R., 2014. Limited geographic distribution of the novel cyclovirus CyCV-VN. *Sci. Rep.* 4, 3967.
- Li, L., Kapoor, A., Slikas, B., Bamidele, O.S., Wang, C., Shaikat, S., Masroor, M.A., Wilson, M.L., Ndjango, J.B., Peeters, M., Gross-Camp, N.D., Muller, M.N., Hahn, B. H., Wolfe, N.D., Triki, H., Bartkus, J., Zaidi, S.Z., Delwart, E., 2010a. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J. Virol.* 84, 1674–1682.
- Li, L., McGraw, S., Zhu, K., Leutenegger, C.M., Marks, S.L., Kubiski, S., Gaffney, P., Dela Cruz Jr, F.N., Wang, C., Delwart, E., Pesavento, P.A., 2013. Circovirus in tissues of dogs with vasculitis and hemorrhage. *Emerg. Infect. Dis.* 19, 534–541.
- Li, L., Shan, T., Soji, O.B., Alam, M.M., Kunz, T.H., Zaidi, S.Z., Delwart, E., 2011. Possible cross-species transmission of circoviruses and cycloviruses among farm animals. *J. Gen. Virol.* 92, 768–772.
- Li, L., Victoria, J.G., Wang, C., Jones, M., Fellers, G.M., Kunz, T.H., Delwart, E., 2010b. Bat guano virome: predominance of dietary viruses from insects and plants plus few mammalian viruses. *J. Virol.* 84, 6955–6965.
- Lima, F.E., Cibulski, S.P., Dos Santos, H.F., Teixeira, T.F., Varela, A.P., Roehe, P.M., Delwart, E., Franco, A.C., 2015. Genomic characterization of novel circular ssDNA viruses from insectivorous bats in southern Brazil. *PLoS One* 10, e0118070.
- Londono, A., Riego-Ruiz, L., Arguello-Astorga, G.R., 2010. DNA-binding specificity determinants of replication proteins encoded by eukaryotic ssDNA viruses are adjacent to widely separated RCR conserved motifs. *Arch. Virol.* 155, 1033–1046.
- Montes, M.J., Belloch, C., Galiana, M., Garcia, M.D., Andres, C., Ferrer, S., Torres-Rodriguez, J.M., Guinea, J., 1999. Polyphasic taxonomy of a novel yeast isolated from antarctic environment; description of *Cryptococcus victoriae* sp. nov. *Syst. Appl. Microbiol.* 22, 97–105.
- Mori, D., Ranawaka, U., Yamada, K., Rajindrajith, S., Miya, K., Perera, H.K., Matsumoto, T., Dassanayake, M., Mitui, M.T., Mori, H., Nishizono, A., Soderlund-Venermo, M., Ahmed, K., 2013. Human bocavirus in patients with encephalitis, Sri Lanka, 2009–2010. *Emerg. Infect. Dis.* 19, 1859–1862.
- Ng, T.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. *J. Virol.* 86, 12161–12175.
- Ng, T.F., Willner, D.L., Lim, Y.W., Schmieder, R., Chau, B., Nilsson, C., Anthony, S., Ruan, Y., Rohwer, F., Breitbart, M., 2011. Broad surveys of DNA viral diversity obtained through viral metagenomics of mosquitoes. *PLoS One* 6, e20579.
- Padilla-Rodriguez, M., Rosario, K., Breitbart, M., 2013. Novel cyclovirus discovered in the Florida woods cockroach *Eurycotis floridana* (Walker). *Arch. Virol.* 158, 1389–1392.
- Phan, T.G., Li, L., O’Ryan, M.G., Cortes, H., Mamani, N., Bonkougou, I.J., Wang, C., Leutenegger, C.M., Delwart, E., 2012. A third gyrovirus species in human faeces. *J. Gen. Virol.* 93, 1356–1361.
- Phan, T.G., Luchsinger, V., Avendano, L.F., Deng, X., Delwart, E., 2014. Cyclovirus in nasopharyngeal aspirates of Chilean children with respiratory infections. *J. Gen. Virol.* 95, 922–927.
- Reuter, G., Pankovics, P., Gyongyi, Z., Delwart, E., Boros, A., 2014. Novel dicistrovirus from bat guano. *Arch. Virol.* 159, 3453–3456.
- Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., Breitbart, M., Varsani, A., 2012a. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J. Gen. Virol.* 93, 2668–2681.
- Rosario, K., Duffy, S., Breitbart, M., 2012b. A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Arch. Virol.* 157, 1851–1871.
- Rosario, K., Marinov, M., Stainton, D., Kraberger, S., Wiltshire, E.J., Collings, D.A., Walters, M., Martin, D.P., Breitbart, M., Varsani, A., 2011. Dragonfly cyclovirus, a novel single-stranded DNA virus discovered in dragonflies (Odonata: Anisoptera). *J. Gen. Virol.* 92, 1302–1308.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sauvage, V., Cheval, J., Foulongne, V., Gouilh, M.A., Pariente, K., Manuguerra, J.C., Richardson, J., Dereure, O., Lecuit, M., Burguiere, A., Caro, V., Eloit, M., 2011. Identification of the first human gyrovirus, a virus related to chicken anemia virus. *J. Virol.* 85, 7948–7950.
- Schwalter, R.M., Pastrana, D.V., Pumphrey, K.A., Moyer, A.L., Buck, C.B., 2010. Merkel cell polyomavirus and two previously unknown polyomaviruses are chronically shed from human skin. *Cell Host Microbe* 7, 509–515.
- Sdiri-Loulizi, K., Gharbi-Khelifi, H., de Rougemont, A., Chouchane, S., Sakly, N., Ambert-Balay, K., Hassine, M., Guediche, M.N., Aouni, M., Pothier, P., 2008. Acute infantile gastroenteritis associated with human enteric viruses in Tunisia. *J. Clin. Microbiol.* 46, 1349–1355.
- Sdiri-Loulizi, K., Gharbi-Khelifi, H., de Rougemont, A., Hassine, M., Chouchane, S., Sakly, N., Pothier, P., Guediche, M.N., Aouni, M., Ambert-Balay, K., 2009. Molecular epidemiology of human astrovirus and adenovirus serotypes 40/41 strains related to acute diarrhea in Tunisian children. *J. Med. Virol.* 81, 1895–1902.
- Sikorski, A., Massaro, M., Kraberger, S., Young, L.M., Smalley, D., Martin, D.P., Varsani, A., 2013. Novel myco-like DNA viruses discovered in the faecal matter of various animals. *Virus Res.* 177, 209–216.
- Smits, S.L., Zijlstra, E.E., van Hellemond, J.J., Schapendonk, C.M., Bodewes, R., Schurch, A.C., Haagmans, B.L., Osterhaus, A.D., 2013. Novel cyclovirus in human cerebrospinal fluid, Malawi, 2010–2011. *Emerg. Infect. Dis.* 19, <http://dx.doi.org/10.3201/eid1909.130404>.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739.
- Tan le, V., van Doorn, H.R., Nghia, H.D., Chau, T.T., Tu le, T.P., de Vries, M., Canuti, M., Deijs, M., Jebbink, M.F., Baker, S., Bryant, J.E., Tham, N.T., NT, B.K., Boni, M.F., Loi, T.Q., Phuong le, T., Verhoeven, J.T., Crusat, M., Jeeninga, R.E., Schultsz, C., Chau, N.V., Hien, T.T., van der Hoek, L., Farrar, J., de Jong, M.D., 2013. Identification of a new cyclovirus in cerebrospinal fluid of patients with acute central nervous system infections. *mBio* 4, e00231 00213.
- Todd, D., 2004. Avian circovirus diseases: lessons for the study of PMWS. *Vet. Microbiol.* 98, 169–174.
- Todd, D., McNulty, M.S., Adair, B.M., Allan, G.M., 2001. Animal circoviruses. *Adv. Virus Res.* 57, 1–70.
- van den Brand, J.M., van Leeuwen, M., Schapendonk, C.M., Simon, J.H., Haagmans, B. L., Osterhaus, A.D., Smits, S.L., 2012. Metagenomic analysis of the viral flora of pine marten and European badger feces. *J. Virol.* 86, 2360–2365.
- Whitley, C., Gunst, K., Müller, H., Funk, M., Zur Hausen, H., de Villiers, E.M., 2014. Novel replication-competent circular DNA molecules from healthy cattle serum and milk and multiple sclerosis-affected human brain tissue. *Genome Announc.* 2, e00849-14. <http://dx.doi.org/10.1128/genomeA00849-14>.
- Whon, T.W., Kim, M.S., Roh, S.W., Shin, N.R., Lee, H.W., Bae, J.W., 2012. Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere. *J. Virol.* 86, 8221–8231.
- Yu, X., Li, B., Fu, Y., Jiang, D., Ghabrial, S.A., Li, G., Peng, Y., Xie, J., Cheng, J., Huang, J., Yi, X., 2010. A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc. Natl. Acad. Sci. USA* 107, 8387–8392.
- Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415.