

## MOLECULAR DETECTION AND GENOME ANALYSIS OF CIRCOVIRUSES OF EUROPEAN EEL (*ANGUILLA ANGUILLA*) AND SICHEL (*PELECUS CULTRATUS*)

Réka BORZÁK\*, Boglárka SELLYEI, Csaba SZÉKELY and Andor DOSZPOLY

Institute for Veterinary Medical Research, Centre for Agricultural Research,  
Hungarian Academy of Sciences, P.O. Box 18, H-1581 Budapest, Hungary

(Received 2 December 2016; accepted 4 April 2017)

The prevalence and distribution of piscine circoviruses (CVs) were tested in a routine virus monitoring programme in Lake Balaton, Hungary. A high prevalence of European eel CV (EeCV) was found in the apparently healthy eel population (35.5%). The copy number of the viral DNA in different organs was determined by quantitative real-time PCR. The results suggested that some eel specimens were in active viraemic status despite their asymptomatic condition. Furthermore, a novel, previously undescribed CV was also detected in eel and sichel samples. Full genome characterisation confirmed that the virus represents a novel EeCV species (EeCV-2). The genome contains an integrated eel chromosome-derived fragment, suggesting that the original host of the virus was the eel and it probably emerged subsequently in the sichel by host switching. In some samples, an additional, 1,111-nt-long circular ssDNA was also observed involving a CV-like stem-loop structure and an ORF showing homology to CV capsid protein genes, without any sign of a replication initiator protein sequence.

**Key words:** *Circoviridae*, circovirus, European eel, sichel

Members of the family *Circoviridae* are non-enveloped, icosahedral viruses with a single-stranded circular DNA (ssDNA) genome ranging from about 1.3 to 2.3 kb in size. According to the current taxonomy release (Breitbart and Varsani, 2015) of the International Committee on Taxonomy of Viruses (ICTV), the family is composed of two genera, *Circovirus* and *Cyclovirus*. Circoviruses (CV) have an ambisense genome organisation containing two major open reading frames (ORF), arranged in opposite directions, encoding the replication initiator protein (Rep) and the capsid protein (Cap). A well-conserved stem-loop structure is located between the 5' ends of the main ORFs, which has an important role in viral replication (Biagini et al., 2011). CVs were first detected from swine and psittacines (Pass and Perry, 1984; Ellis et al., 1998) in asymptomatic infections (*Porcine circovirus 1*, PCV-1) or causing multiple severe clinical signs

---

\*Corresponding author; borzak.reka@agrar.mta.hu; Phone: 0036 (1) 467-4079

(*Porcine circovirus 2*, PCV-2, *Beak and feather disease virus*, BFDV) (Allan and Ellis, 2000; Todd, 2004).

Recently, due to the different metagenomic approaches and broad-spectrum PCR assays (Halami et al., 2008), enormous numbers of circovirus-related sequences have been described from different mammals including humans (Li et al., 2010; Ge et al., 2011; Lian et al., 2014), avian species (Todd et al., 2001a; Todd et al., 2001b; Hattermann et al., 2003), lower vertebrates (amphibians and fish) (Lőrincz et al., 2011; Tarján et al., 2014), invertebrate animals (Rosario et al., 2011; Marton et al., 2015), and also from environmental samples (Rosario et al., 2009; Zawar-Reza et al., 2014).

Following the description of the first fish CV from barbel fry (*Barbus barbus*) showing high mortality in a Hungarian fish farm, newly discovered CVs were detected in adult European catfish (*Silurus glanis*) carcasses and in a European eel (*Anguilla anguilla*) with cauliflower disease from Lake Balaton (Lőrincz et al., 2011; Lőrincz et al., 2012; Doszpoly et al., 2014). Furthermore, CV-like *rep* sequences have been detected by PCR in different fish species such as common bream (*Abramis brama*), round goby (*Neogobius melanostomus*), common roach (*Rutilus rutilus*), asp (*Aspius aspius*) and two Indian cyprinid fish species (*Labeo rohita* and *Catla catla*), although the amplification of the complete, circular genome failed from these samples, probably because the discovered CVs were endogenous viral elements (EVEs) in the host genomes (Fehér et al., 2013; Tarján et al., 2014).

Apart from these case reports, our knowledge about piscine CVs is very limited. The aim of this study was to provide data on the prevalence and distribution of fish CVs during a routine virus monitoring programme in Lake Balaton, Hungary.

## Materials and methods

### Sampling

A total of 163 fish specimens (Table 1) representing 14 species were collected throughout the year 2014 from Lake Balaton at different sampling sites (sampling permit number: EHVF/121-1/2014). After euthanasia with MS-222 the following organs were reserved for virus isolation and molecular studies: gills, liver, spleen, kidney, intestine, and brain were collected separately and one pool from these organs per fish was also made for rapid CV detection.

### Virus isolation

Virus isolation was attempted from pooled organ homogenates and from organs which showed high viral load by qPCR using EK-1 cell line (Chen et al., 1982), as described in detail by Doszpoly et al. (2014).

**Table 1**

The fish species collected for this study. The CV-positive samples are in round brackets, below the total number of the species per month. *In italics*: where only the partial rep sequences (*rep+*) were detected

2014	March	April	May	June	September
Perch ( <i>Perca fluviatilis</i> )	2	–	4	–	3
Common bream ( <i>Abramis brama</i> )	3	4 (1 <i>rep+</i> )	2	3	11 (1 <i>rep+</i> )
Sichel ( <i>Pelecus cultratus</i> )	3 (1 EeCV-2 + CLV)	–	5	1	8
Silver bream ( <i>Blicca bjoerkna</i> )	3	2	3	–	1
Pikeperch ( <i>Sander lucioperca</i> )	6	3	3	–	–
Volga pikeperch ( <i>Sander volgensis</i> )	1	–	6	–	–
Common carp ( <i>Cyprinus carpio</i> )	2	–	1	–	3
Tench ( <i>Tinca tinca</i> )	1	–	–	–	–
Asp ( <i>Aspius aspius</i> )	2	–	2	–	1
Prussian carp ( <i>Carassius gibelio</i> )	1	–	10	–	2
Eel ( <i>Anguilla anguilla</i> )	15 (3 EeCV-1) (3 <i>rep+</i> )	–	–	–	16 (7 EeCV-1 + 2 CLV) (1 EeCV-2 + CLV) (6 <i>rep+</i> )
Roach ( <i>Rutilus rutilus</i> )	–	1 (1 <i>rep+</i> )	–	–	15 (2 <i>rep+</i> )
Bleak ( <i>Alburnus alburnus</i> )	–	3	–	2	–
Brown bullhead ( <i>Ameiurus nebulosus</i> )	–	–	5	–	4

#### DNA extraction and PCR

DNA was extracted from approx. 25 mg tissue samples using DNeasy<sup>®</sup> Blood & Tissue Kit (Qiagen, Germany) according to the manufacturer's instructions. For the detection of circoviral DNA, a broad-spectrum nested PCR method (Halami et al., 2008) targeting a roughly 350-nt-long region of the *rep* gene was used on the pooled tissue samples. The PCRs were performed using DreamTaq DNA polymerase (ThermoFisher Scientific, USA) and 10× DreamTaq Buffer (including 20 mM MgCl<sub>2</sub>), 0.2 mM dNTP, 0.2 µM from each primer (the primers are listed in Table 2), molecular-grade water and 2 µl template DNA in a 25-µl

reaction on a Thermal Cycler Applied Biosystem 2720 (ThermoFisher Scientific, USA). The optimised cycling protocol for the first round of PCR included 5 min of incubation at 95 °C, followed by 45 cycles each consisting of 94 °C for 30 s, 46 °C for 1 min and 72 °C for 1 min, and a final incubation at 72 °C for 5 min. The cycling protocol for the nested PCR was identical; however, an annealing temperature of 56 °C was used in this case.

The remaining part of the circular genome was amplified with sequence-specific inverse nested PCR (the primers are listed in Table 2). The reaction mixture consisted of 10 µM of each (forward and reverse) primer, 1× PrimeSTAR Max Premix (Takara Bio, Otsu, Japan), nuclease-free water (Ambion, ThermoFisher Scientific, Waltham, MA USA), and 2 µl of the target DNA in 50 µl final volume. The reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, ThermoFisher Scientific, USA), and the following amplification profile was used: 98 °C initial denaturation for 3 min, followed by 45 cycles of 98 °C for 10 sec, 55 °C for 15 sec, 72 °C for 10 sec and a final extension at 72 °C for 3 min.

When the inverse nested PCR was unsuccessful, rolling circle amplification (RCA) was also applied for obtaining the whole circular genomes. For RCA reaction the TempliPhi 100 Amplification Kit (GE Healthcare, Germany) was used according to the manufacturer's instructions. The amplified product was cleaved by SmaI restriction endonuclease enzyme and transformed into competent *E. coli* cells using the CloneJET PCR Cloning Kit (ThermoFisher Scientific, USA).

#### *Quantitative, real-time PCRs*

Quantitative PCRs were designed targeting an approx. 100-nt-long region of the *cap* gene to estimate the viral load in each organ (gills, liver, spleen, kidney, intestine, and brain) separately to gain information about the tropism of the virus and the viraemic status of the fish. DNA was extracted from the tissue samples, quantified with Nanodrop, then 100 ng/µl stock solutions were made from each organ with molecular-grade water. The primers and probes were designed with the Primer3 software (Untergasser et al., 2012) and are listed in Table 2. The reactions were performed in an Applied Biosystems® StepOnePlus™ Real-Time PCR System instrument (Life Technologies, ThermoFisher Scientific, USA). The reaction mixture contained 1× Xceed qPCR Probe Mix (Institute of Applied Biotechnologies, Czech Republic), 0.4 µM of each primer, 0.2 µM of the probe, nuclease-free water (Ambion, ThermoFisher Scientific, USA) and 100 ng extracted DNA in 20 µl final volume. The programme consisted of an initial step of 70 °C for 5 min, 60 °C for 30 sec and 72 °C for 5 min for the conversion of viral ssDNA to dsDNA, then for the amplification the following programme was used: 95 °C for 3 min, followed by 40 cycles of 95 °C for 5 sec and 60 °C for 20 sec. The results were analysed by the StepOne Software v2.1 (Applied Biosystems, ThermoFisher Scientific, USA). For the standard curve: a 507-bp-long PCR product was amplified from the *cap* gene, containing the target sequence of the

qPCR. It was purified, and the concentration was determined. Using the molecular weight of this fragment and the Avogadro constant, the number of *cap* fragments per 1 µl of this purified solution was determined, then a tenfold serial dilution of these circoviral *cap* fragments was used as standard.

#### *Sequencing and sequence analysis*

The PCR products were excised from 1% agarose gel after electrophoresis, and purified with the Gel/PCR Fragment Extraction Kit (Geneaid Biotech, Taiwan). DNA fragments were sequenced directly with the inner primers (Table 2). The cloned fragments obtained by RCA were sequenced with the primers supplied in the CloneJET kit.

The sequencing reactions were performed with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, ThermoFisher Scientific, USA). The capillary electrophoresis was carried out by a commercial service provider on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, ThermoFisher Scientific, USA). The newly obtained nucleotide sequences were corrected and assembled using the MEGA5.2 software (Tamura et al., 2011). The ORFs were identified with the ORF Finder graphical analysis tool on the NCBI website. The size and orientation of the identified ORFs and putative viral genes were defined with online prediction programs at Softberry portal (<http://linux1.softberry.com/berry.phtml?topic=index&group=programs&subgroup=gfindv>).

Homology searches were carried out in GenBank by the use of different BLAST algorithms (BLASTn, BLASTx) at the NCBI server. MEGA 5.2 software was used to calculate the pairwise distances at nucleotide and at amino acid level applying Maximum Composite Likelihood and p-distance algorithms.

#### *Statistical analysis*

Statistical significance was calculated using Student's *t*-test with the program package GraphPad Prism 4.00. P values of 0.05 or less were considered statistically significant. Each presented qPCR result represents data derived from a minimum of 3 replicates per group. All error bars represent standard deviation (SD).

## **Results**

#### *Virus isolation*

The virus isolation was unsuccessful on the EK-1 cells. No cytopathic effect (CPE) was detected after 14 days in either the first or the second passage. qPCR was carried out on the homogenised, pooled organs and the consecutive passages. A tenfold decrease in the copy number was observed after each passage, confirming that virus propagation did not occur.

**Table 2**  
Primers and probes used in nested PCRs and qPCR assays and the obtained genome sequences

Nested PCR	PCR round	Primer forward (5'-3')	Primer reverse (5'-3')	Reference
All samples – CV detection	1	AGAGGTGGGTCTTCACNHTBAAYAA	AAGGCAGCCACCCTARAARTCTC	Halami et al. (2008)
	2	AGCAAGGAACCCCTCAYYTBCARG G	ACGATGACTTCNGTCTTSMARTCACG	
<b>Nested inverse PCR</b>				
EeCV-like eel samples (10)	1	GCGCTTGAGGATTCTCATTC	CAGATCGTTCTCTCTTCCCTT	KU951578-9
	2	GACTTTGGATGGAAGAAAGCC	CCTTGTTATGCTGGTCGTTG	
Eel (3) and sichel (1) samples	1	GCGCTTGAGGATTCTCATGC	CAGATCGTTCTCTCTTCCCTT	KU951580
	2	GACTTCGGATGGAAGAAGGA	CCTTGTTATGCTGGTCGTTG	
Nested PCR	PCR round	Primer forward (5'-3')	Primer reverse (5'-3')	Reference
<b>qPCR</b>	<b>Probe (5'-FAM – TAMRA-3')</b>			
EeCV-like	CCGTGCTACCTTGGAAGGCTTCC	GGGACCTGTGAAGGAGTGG	AACATACAGCCGGTTTAGCC	KU951578-9
Circular cap	TTCATTTGAACATCGCCTTGAGCAA	CTCGGTTTCATCGCTTCACTA	CCAATACCAAAACCCAAAGGTC	KU951581

**Table 3**

Detailed information about the partial rep sequences detected in this study. Nucleotide sequence identities to known CV or fish genome sequences are in round brackets

2014 / rep +	March	April	September
Common bream ( <i>Abramis brama</i> )		1 (99% <i>Roach circo-like seq</i> , <i>Tarján et al. (2014)</i> , KF358278)	1 (98% <i>Swan CV</i> , <i>Halami et al. (2008)</i> , EU056310)
Eel ( <i>Anguilla anguilla</i> )	3 2 (100% <i>EeCV-2</i> ) 1 (69% <i>BarCV</i> , 70% <i>carp genome</i> )	6 1 (99% <i>EeCV-1</i> ) 4 (98% <i>EeCV-2</i> ) 1 (97% <i>EeCV-2</i> , 95% <i>EeCV-1</i> )	
Roach ( <i>Rutilus rutilus</i> )		1 (69% <i>BarCV</i> , 70% <i>carp genome</i> )	2 (69% <i>BarCV</i> , 70% <i>carp genome</i> )

**Table 4**

Conserved nonanucleotide motifs and orientation of the Rep protein in fish circoviruses based on the study of Rosario et al. (2009). U = I, L, V, M, F, Y, W. The results of the present study are marked with an asterisk

	Nonanucleotide motif	Rep orientation	RCR motifs				SF3 motifs			
			I	II	III	Walker-A	Walker-B	Motif-C		
Circovirus	n t TAGTATTAC	+ strand	V I CFTLNN	PHLQG	S YCxK	PG GPSxGKS	I VM I LDDF	I UTSN		
Cyclovirus	TAGTATTAC	- strand	C W VFTLNN	P xHLQG	S YCxK	PP GxtGxGKS	I I VUDDF	N UTSe		
Barbel CV	TAGTATTAC	+	CFTLNN	PHLQG	YCTK	GDPGCSKS	IVDDF	VTSN		
Catfish CV	TAGTATTAC	+	VFTLNN	PHLQG	YCSK	GPPGCGKS	ILDDF	ITSN		
Eel CV	TAGTATTAC	+	CFTLNN	PHSQQ	YCSK	GIPGCGKS	ILDDF	ITSN		
Eel specimens*	TAGTATTAC	+	CFTLNN	PHSQQ	YCSK	GIPGCGKS	ILDDF	ITSN		
Eel (AN24) and sichei specimens*	TAGTATTAC	+	CFTLNN	PHLQG	YCSK	GIPGCGKS	IVDDF	ITSN		

### *Circovirus detection by PCR*

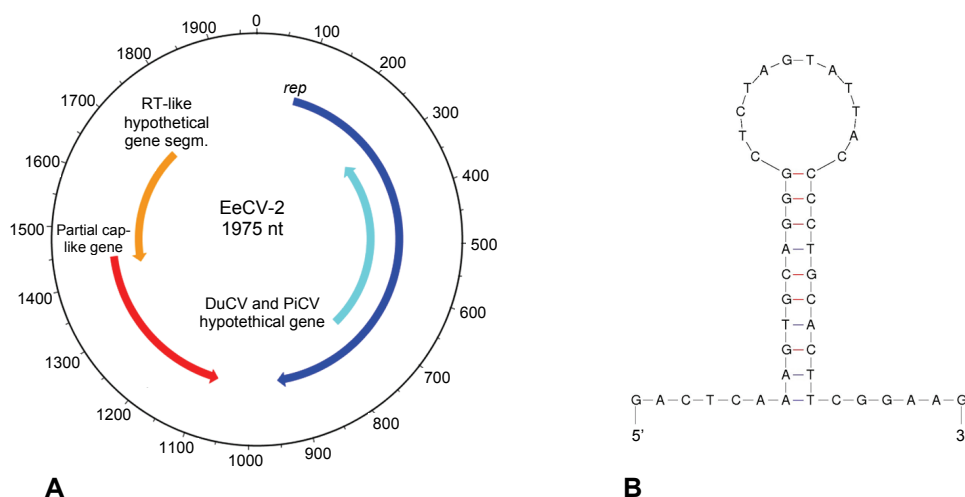
Circovirus-like sequences were detected in one sichel (*Pelecus cultratus*), two common bream (*Abramis brama*), three common roach (*Rutilus rutilus*) and twenty European eel (*Anguilla anguilla*) specimens, although the whole circular viral genome amplification by inverse PCR was successful only from the sichel sample and from 11 eel samples. Henceforth, only these samples were considered to be positive [6% (1/17) and 35.5% (11/31) prevalence in sichel and eel, respectively], while the rest were regarded as EVEs (Fehér et al., 2013; Tarján et al., 2014). A smaller circular DNA containing a CV-like *cap* but not *rep* genes was also revealed by RCA in the sichel sample and in three eel samples. Detailed information about the species containing any CV-related sequences can be found in Tables 1 and 3.

### *Genome sequence analysis*

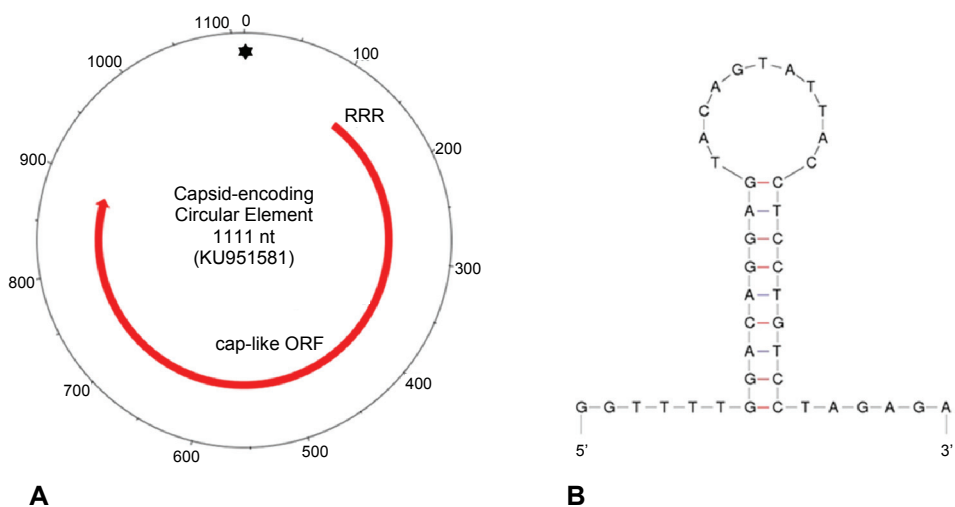
Sequence alignment from 10 eel samples encompassed a 1378-nt-long CV genome, presenting 99.9% nucleotide similarity and identical genomic structure to the previously described European eel circovirus (EeCV-1) (Dospoly et al., 2014).

The other complete genomes obtained from the sichel sample and from one eel sample were identical to each other (EeCV-2) with the size of 1975 nt and with an average G+C content of 47.2%. The genome organisation was ambisense with the presence of two major inversely arranged open reading frames (ORFs) that were responsible for encoding the Rep and the putative Cap (Fig. 1a). A CV-specific stem-loop structure with a 10-bp-long stem, a 12-nt-long loop structure and with the conserved nonanucleotide motif (5'-TAGTATTAC-3') was identified at the 5' intergenic region (Fig. 1b), which was much longer (~600 nt) than that in the previously reported genomes (29–207 nt) (Delwart and Li, 2012). The Rep consisted of 286 amino acids (aa) like that of EeCV-1. It contained the conserved domains of rolling circle replication (RCR) and of SF3 (superfamily III) viral helicase (Table 4). The highest similarity to the Rep of EeCV-1 was 85% aa identity. The 123 aa long putative Cap protein showed 39% and 35% aa identity with that of the *Human faeces associated circovirus* (HufaCV, previously named NG13) and the *European catfish circovirus* (EcatfishCV), respectively (Li et al., 2010; Lőrincz et al., 2012), but no homology to the EeCV-1 Cap was seen. Two additional ORFs were also observed in the same orientation as the *cap* gene. One of them was a 468-nt-long ORF overlapping the *rep* gene, which showed homology to a hypothetical gene in pigeon and duck CVs (Mankertz et al., 2000; Chen et al., 2006). The other, 312-nt-long ORF was located in the 5' intergenic region. It was presumed to be a host genome-derived sequence, since it was identical to a fragment of the reverse-transcriptase gene of the European eel (GenBank Acc. No.: AZBK01506574). This is the first description of a putative gene in CVs which shows obvious host origin.





**Fig. 1.** Genome organisation (A) and the stem-loop structure (B) of the novel CV genome (EeCV-2) from eel and sichel. A: The location of the stem-loop structure is indicated with a black asterisk. Abbreviations: DuCV – duck circovirus, PiCV – pigeon circovirus. The GenBank accession number is presented in round brackets



**Fig. 2.** Genomic organisation (A) and stem-loop structure (B) of the circular, monocistronic element, named Capsid-encoding Circular Element (CCE). A: The location of the stem-loop structure is indicated with a black asterisk. The GenBank accession number is presented in round brackets

Furthermore, a circular DNA (Capsid-encoding Circular Element, CCE) was detected in the EeCV-2 containing sichel and in three eel samples (one of them bearing EeCV-2 and two of them containing EeCV-1), its length proved to be 1,111 nt with a G+C content of 43.7% (Fig. 2a). A CV-specific stem-loop

structure was identified with a 9-bp-long stem and an 11-nt-long loop region, consisting of a conserved nonanucleotide motif (5'-CAGTATTAC-3') (Fig 2b). No *rep* gene was found in the sequence; however, a putative gene encoding Cap protein was identified with a size of 255 aa. Close to the N terminus of this protein, a 23-aa-long arginine-rich stretch (residues 1–23 aa, MRRTVTEDGPPA RRARPNSASRR) was detected. The most similar proteins with 29% and 26% aa identities were putative Cap proteins of ssDNA viruses from the aquatic invertebrates *Paphies subtriangulata* (Dayaram et al., 2015) and *Penaeus monodon* (Pham et al., 2014), respectively. The sequences of the complete genomes and the CCE were deposited in GenBank and assigned the accession numbers KU951578–KU951581.

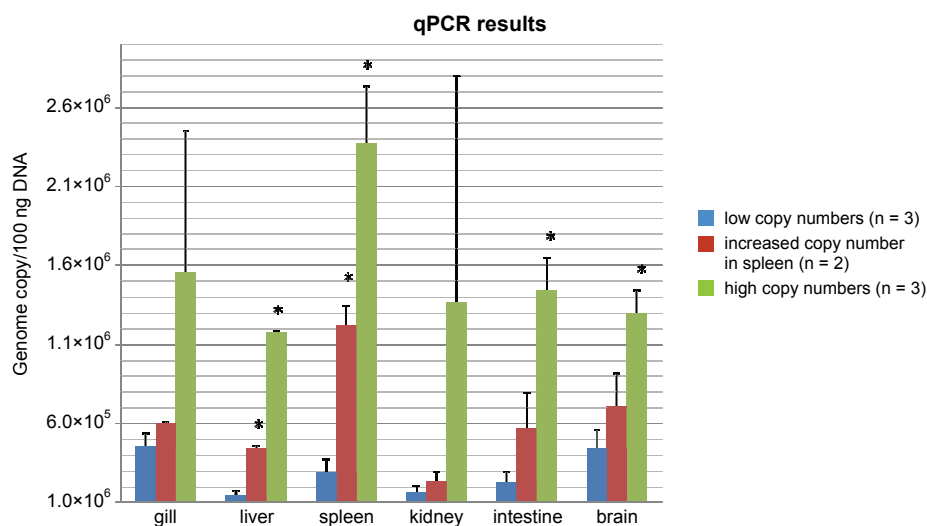


Fig. 3. Distribution of EeCV-1 by qPCR. The individuals were grouped in three tentative clusters, according to the low (less than  $5 \times 10^5/100$  ng DNA; blue columns,  $n = 3$ ), increased (red columns,  $n = 2$ ) and high copy numbers (more than  $10^6/100$  ng DNA; green columns,  $n = 3$ ) detected in the organs. The mean and the standard deviation values (SD) of the grouped samples are presented on the graph. Mean  $\pm$  SD compared by Student's *t*-test, \* $P < 0.001$

### Quantitative PCR

Absolute quantification of the EeCV-1 genome from the eel samples was performed in the gill, liver, spleen, kidney, intestine and brain tissues. The virus was detected in all of the examined tissues of eight eels having EeCV-1 infection (the detection limit was 6 copy numbers in 20- $\mu$ l reaction). Based on the detected amount of virus in the different organs, three tentative groups were delineated (Fig. 3). Low copy numbers ( $10^5 - 5 \times 10^5/100$  ng DNA) were detected in all organs of three eels. In the second group (two individuals) the viral load was sig-

nificantly higher only in the spleen ( $1.2 \times 10^6/100$  ng DNA), while in the third group (three eels) we found high viral load in all organs ( $1.2 \times 10^6 - 2.4 \times 10^6/100$  ng DNA).

### Discussion

In the present study, a high prevalence of EeCV-1 (35.5%) was detected in the European eel population of Lake Balaton, Hungary. Furthermore, the full genome sequence of a putative novel piscine CV (EeCV-2) and another circular ssDNA encoding Cap protein (CCE) were determined from two distantly related fish species (the European eel and the sibelius). The examined specimens did not show any pathological alterations, while an eel from which the EeCV-1 was first described, suffered from cauliflower disease (Doszpöly et al., 2014), and fish CVs were also detected in mass mortality events of barbel and European catfish populations (Lőrincz et al., 2011; Lőrincz et al., 2012). In fish, a direct connection between the diseases and CVs was never confirmed. However, CVs reported in other species frequently induce immunosuppression, which increases the host's susceptibility to other pathogens, causing several clinical signs (Todd, 2004). The EeCV-1-like genomes obtained from 10 eel samples differed only in two single-nucleotide substitutions from the original EeCV-1 genome, which is derived from an eel sample, collected in 2009 (Doszpöly et al., 2014). Although the single-nucleotide polymorphisms (SNPs) caused amino acid changes in both the Rep and the Cap proteins, the conserved sites (stem-loop structure, RCR, SF3) were not affected. The number of nucleotide substitutions was surprisingly low, considering that there was a 5-year difference between the collection dates (Doszpöly et al., 2014). In other studies on PCV-2, the mean substitution rate for the full-genome data set was  $1.21 \times 10^{-3}$  nucleotide substitutions per site per year (Firth et al., 2008).

The distribution of the viral DNA of EeCV-1 was determined in different organs of the eel samples. In previous studies (porcine and piscine CVs), for the quantification of the viral load, the *rep* was targeted (McIntosh et al., 2008; Doszpöly et al., 2014). As numerous integrated circoviral rep-like elements had been identified with this target in the lower vertebrate and invertebrate hosts recently (Belyi et al., 2010; Fehér et al., 2013; Marton et al., 2015), in this study the *cap* was targeted for the qPCR assay. Three groups were distinguished by the different viral copy numbers detected in the samples. Three eels belonged to the first group, where relatively low copy numbers were found in all organs ( $10^5 - 5 \times 10^5/100$  ng DNA) with the highest viral load detected in the gill, brain and spleen. In the second group we observed a significant increase of the viral load, especially in the spleen in two eels ( $1.2 \times 10^6/100$  ng DNA). In bony fishes (Osteichthyes), the spleen is one of the major lymphoid organs, which is composed of a system of splenic ellipsoids, melano-macrophage centres and lym-

phoid tissue. The main role of the ellipsoid system is capturing and trapping the antigens that enter the host (Ellis, 2001; Uribe et al., 2011). The higher viral load in the spleen may indicate an ongoing infection or a viraemic status when a high amount of virus is detected in the other organs, as well. The high viral load in the brain suggests that the virus can affect the central nervous system of the fish, causing behavioural changes. When the catfish CV was detected in 2011, infection trials were carried out on catfish fingerlings. Most of them showed abnormal swimming behaviour after a few months, but the virus could not be detected from the fish again (unpublished results). In the last group high copy numbers (more than  $10^6/100$  ng DNA) were seen in all organs in three eels, with the highest being detected in the spleen ( $2.4 \times 10^6/100$  ng DNA). Probably the individuals in this group were in viraemic state at the time of collection.

CVs are known as host-specific or narrow host range microorganisms (Todd et al., 2007). Nonetheless, here we describe representatives of a putative novel CV (EeCV-2) showing 100% nt identity, collected from a sichel and an eel sample separately throughout the year. To the best of our knowledge, this is the first report on a circoviral genome detected in different, distantly related fish species (members of different superorders). The *rep* gene of the novel CV had an 85% nucleotide identity to EeCV-1, but interestingly their *cap* genes did not show any homology. The nonanucleotide motif, the direction of the *rep* gene and the aa composition in the conserved motifs of the Rep protein confirmed the taxonomic position of the novel CV in the genus *Circovirus*. However, the Walker-A and -B motifs in all piscine CVs are slightly different from those of both the circo- and the cycloviruses of various hosts (Rosario et al., 2012). Like the corresponding region of the EeCV-1, the Cap protein of the EeCV-2 was relatively short, 123 aa long, although sequence similarity to the *cap* of EeCV-1 was not observed. Moreover, the CV-specific arginine-rich stretch at the N-terminus of the gene was not found. Nevertheless, 39% and 35% amino acid identity to the HufaCV and the EcatfishCV was determined. This result is consistent both with the previous report that presumed a close evolutionary origin of the fish CVs and the HufaCV (Dospoly et al., 2014) and with the novel ICTV release (2014.006a,hD) about their distinct taxonomic position from the other CVs. The unusually long 5' intergenic region of EeCV-2 contained a partial RT-like gene fragment, which showed 99% sequence identity to an RT gene of the European eel. This finding suggests the eel origin of the virus and a subsequent host switch to the sichel. Recombinational transfer of genetic material between ssDNA viruses and their hosts has already been reported in both directions (Martin et al., 2011). The integration of viral elements into the host genome seems to occur relatively frequently among CVs (Belyi et al., 2010; Fehér et al., 2013; Marton et al., 2015); however, this is the first report on an integrated host-derived sequence in the CV genome. The nucleotide identity of the complete genome of the novel virus and the EeCV-1 is 74.8%, while their Cap sequence similarity is less than

15%. Thus the novel CV fulfils the species demarcation criteria of the ICTV for circoviruses, hence the authors propose the *European eel circovirus 2* (EeCV-2) species designation for the newly detected virus. The EeCV-1 has not been yet accepted as a virus species by the ICTV since the small size of the putative Cap protein (114 aa) raises doubts about its ability to form operating capsid protein. The novel EeCV-2 also has a short *cap* gene (123 aa).

A circular, monocistronic ssDNA encoding a Cap protein of appropriate size (255 aa) was detected by RCA and was named Capsid-encoding Circular Element. Its stem-loop structure showed the highest similarity to the nonanucleotide motif of some avian and environmental CVs (Phenix et al., 2001; López-Bueno et al., 2009; Rosario et al., 2009), while the putative Cap protein showed similarity to that of the CVs of aquatic invertebrates. CV-specific arginine-rich stretch at the N-terminus of the capsid protein was also observed. Despite these similarities to the CVs, the size and the structure of this circular element were similar to a component of the multipartite, unidirectional plant-infecting nanoviruses (Gronenborn, 2004), suggesting the possible existence of multipartite circo-like viruses, members of a novel virus family constituting a missing link between circoviruses and nanoviruses (Rosario et al., 2012). Since both EeCV genomes encode a surprisingly short Cap protein, our hypothesis was that CCE expresses the real capsid, forming a multipartite virus with EeCVs. However, we were able to detect the CCE genome only in four samples (the eel and sibel samples containing EeCV-2 and two eel samples infected by EeCV-1). Moreover, in multipartite viruses, the non-coding regions of the genome components contain a highly conserved sequence, flanked by the same stretch of inverted repeats (Katul et al., 1997); such sequences were not found in the CCE. Therefore, these circular agents probably do not form a multipartite viral unit together with the EeCV-1 or EeCV-2. More likely the CCE could be a satellite virus using the EeCV-1 and -2 as a helper virus.

### Acknowledgements

The financial support (OTKA PD104315) provided by the Hungarian Scientific Research Fund is gratefully acknowledged. This study was supported by the GINOP-2.3.2-15-2016-00004 project: 'Establishing the sustainable angling-aimed management of Lake Balaton'. Andor Doszpoly and Boglárka Sellyei are recipients of the János Bolyai Research Scholarship awarded by the Hungarian Academy of Sciences. Thanks are also due to Dr. Kálmán Molnár and Ádám Varga from the Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences for their kind assistance in the sample collection. The authors express their thanks to colleagues from the Balaton Limnological Institute of the Hungarian Academy of Sciences and from the Balaton Fisheries Ltd., respectively, for the collection of fish.

## References

- Allan, G. M. and Ellis, J. A. (2000): Porcine circoviruses: a review. *J. Vet. Diagn. Invest.* **12**, 3–14.
- Belyi, V. A., Levine, A. J. and Skalka, A. M. (2010): Sequences from ancestral single-stranded DNA viruses in vertebrate genomes: the *Parvoviridae* and *Circoviridae* are more than 40 to 50 million years old. *J. Virol.* **84**, 12458–12462.
- Biagini, P., Bendinelli, M., Hino, S., Kakkola, L., Mankertz, A., Niel, C., Okamoto, H., Raidal, S., Teo, C. G. and Todd, D. (2011): Family Circoviridae. In: King, A. M. Q., Adams, M. J., Carstens, E. B. and Fields, B. S. (eds) *Virus Taxonomy*, IXth Report of the International Committee on Taxonomy of Viruses. Elsevier, Academic Press, London. pp. 99–123.
- Breitbart, M. and Varsani, A. (2015): Establish the genus *Cyclovirus* in the family *Circoviridae*; creating 28 new species in the family (2014.006b-eD). ICTV Taxonomy History for Cyclovirus, *Virus Taxonomy*: 2015 Release, EC 47, London, UK.
- Chen, C. L., Wang, P. X., Lee, M. S., Shien, J. H., Shien, H. K., Ou, S. J., Chen, C. H. and Chang, P. C. (2006): Development of a polymerase chain reaction procedure for detection and differentiation of duck and goose circovirus. *Avian Dis.* **50**, 92–95.
- Chen, S. N., Ueno, Y. and Kou, G. H. (1982): A cell line derived from Japanese eel (*Anguilla japonica*) kidney. *Proc. Natl. Sci. Coun. POC* **6**, 93–100.
- Dayaram, A., Goldstien, S., Argüello-Astorga, G. R., Zawar-Reza, P., Gomez, C., Harding, J. S. and Varsani, A. (2015): Diverse small circular DNA viruses circulating amongst estuarine molluscs. *Infect. Genet. Evol.* **31**, 284–295.
- Delwart, E. and 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.
- Doszpoly, A., Tarján, Z. L., Glávits, R., Müller, T. and Benkő, M. (2014): Full genome sequence of a novel circo-like virus detected in an adult European eel *Anguilla anguilla* showing signs of cauliflower disease. *Dis. Aquat. Organ.* **109**, 107–115.
- Ellis, A. E. (2001): Innate host defense mechanisms of fish against viruses and bacteria. *Dev. Comp. Immunol.* **25**, 827–839.
- Ellis, J., Hassard, L., Clark, E., Harding, J., Allan, G., Willson, P., Strokappe, J., Martin, K., McNeilly, F., Meehan, B., Todd, D. and Haines, D. (1998): Isolation of circovirus from lesions of pigs with postweaning multisystemic wasting syndrome. *Can. Vet. J.* **39**, 44–51.
- Fehér, E., Székely, C., Lőrincz, M., Cech, G., Tuboly, T., Singh, H. S., Bányai, K. and Farkas, S. L. (2013): Integrated circoviral rep-like sequences in the genome of cyprinid fish. *Virus Genes* **47**, 374–377.
- Firth, C., Charleston, M. A., Duffy, S., Shapiro, B. and Holmes, E. C. (2009): Insights into the evolutionary history of an emerging livestock pathogen: porcine circovirus 2. *J. Virol.* **83**, 12813–12821.
- Ge, X., Li, J., Peng, C., Wu, L., Yang, X., Wu, Y., Zhang, Y. and Shi, Z. (2011): Genetic diversity of novel circular ssDNA viruses in bats in China. *J. Gen. Virol.* **92**, 2646–2653.
- Gronenborn, B. (2004): Nanoviruses: genome organisation and protein function. *Vet. Microbiol.* **98**, 103–109.
- Halami, M. Y., Nieper, H., Müller, H. and Johne, R. (2008): Detection of a novel circovirus in mute swans (*Cygnus olor*) by using nested broad-spectrum PCR. *Virus Res.* **132**, 208–212.
- Hattermann, K., Schmitt, C., Soike, D. and Mankertz, A. (2003): Cloning and sequencing of Duck circovirus (DuCV). *Arch. Virol.* **148**, 2471–2480.
- Katul, L., Maiss, E., Morozov, S. Y. and Vetten, H. J. (1997): Analysis of six DNA components of the faba bean necrotic yellows virus genome and their structural affinity to related plant virus genomes. *Virology* **233**, 247–259.
- Li, L., Kapoor, A., Slikas, B., Bamidele, O. S., Wang, C., Shaukat, S., Masroor, M. A., Wilson, M. L., Ndjanga, J. B., Peeters, M., Gross-Camp, N. D., Muller, M. N., Hahn, B. H., Wolfe, N.

- D., Triki, H., Bartkus, J., Zaidi, S. Z. and Delwart, E. (2010): Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. *J. Virol.* **84**, 1674–1682.
- Lian, H., Liu, Y., Li, N., Wang, Y., Zhang, S. and Hu, R. (2014): Novel circovirus from mink, China. *Emerg. Infect. Dis.* **20**, 1548–1550.
- López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A. and Alcamí, A. (2009): High diversity of the viral community from an Antarctic lake. *Science* **326**, 858–861.
- Lőrincz, M., Cságola, A., Farkas, S. L., Székely, C. and Tuboly, T. (2011): First detection and analysis of a fish circovirus. *J. Gen. Virol.* **92**, 1817–1821.
- Lőrincz, M., Dán, Á., Láng, M., Csaba, G., Tóth, A. G., Székely, C., Cságola, A. and Tuboly, T. (2012): Novel circovirus in European catfish (*Silurus glanis*). *Arch. Virol.* **157**, 1173–1176.
- Mankertz, A., Hattermann, K., Ehlers, B. and Soike, D. (2000): Cloning and sequencing of columbid circovirus (CoCV), a new circovirus from pigeons. *Arch. Virol.* **145**, 2469–2479.
- Biagini, P., Lefeuvre, P., Golden, M., Roumagnac, P. and Varsani, A. (2011): Recombination in eukaryotic single stranded DNA viruses. *Viruses* **3**, 1699–1738.
- Ihász, K., Lengyel, G., Farkas, S. L., Dán, Á., Paulus, P., Bányai, K. and Fehér, E. (2015): Ubiquitous circovirus sequences raise challenges in laboratory diagnosis: the case of honey bee and bee mite, reptiles, and free living amoebae. *Acta Microbiol. Immunol. Hung.* **62**, 57–73.
- McIntosh, K. A., Harding, J. C., Parker, S., Krakowka, S., Allan, G. and Ellis, J. A. (2008): Quantitative polymerase chain reaction for Porcine circovirus-2 in swine feces in a Porcine circovirus disease-affected commercial herd and a nonaffected commercial herd. *Can. Vet. J.* **49**, 1189–1194.
- Pass, D. A. and Perry, R. A. (1984): The pathology of psittacine beak and feather disease. *Aust. Vet. J.* **61**, 69–74.
- Pham, H. T., Yu, Q., Boisvert, M., Van, H. T., Bergoin, M. and Tijssen, P. (2014): A circo-like virus isolated from *Penaeus monodon* shrimps. *Genome Announc.* 2:e01172-13.
- Phenix, K. V., Weston, J. H., Ypelaar, I., Lavazza, A., Smyth, J. A., Todd, D., Wilcox, G. E. and Raidal, S. R. (2001): Nucleotide sequence analysis of a novel circovirus of canaries and its relationship to other members of the genus *Circovirus* of the family *Circoviridae*. *J. Gen. Virol.* **82**, 2805–2809.
- Rosario, K., Duffy, S. and Breitbart, M. (2009): Diverse circovirus-like genome architectures revealed by environmental metagenomics. *J. Gen. Virol.* **90**, 2418–2424.
- Rosario, K., Duffy, S. and Breitbart, M. (2012): 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. and Varsani, A. (2011): Dragonfly cyclovirus, a novel single-stranded DNA virus discovered in dragonflies (Odonata: Anisoptera). *J. Gen. Virol.* **92**, 1302–1308.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011): MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.
- Tarján, Z. L., Péntzes, J. J., Tóth, R. P. and Benkő, M. (2014): First detection of circoviruses-like sequences in amphibians and novel putative circoviruses in fishes. *Acta Vet. Hung.* **62**, 134–144.
- Todd, D. (2004): Avian circovirus diseases: lessons for the study of PMWS. *Vet. Microbiol.* **98**, 169–174.
- Todd, D., Scott, A. N., Fringuelli, E., Shivraprasad, H. L., Gavier-Widen, D. and Smyth, J. A. (2007): Molecular characterization of novel circoviruses from finch and gull. *Avian Pathol.* **36**, 75–81.
- Todd, D., Weston, J., Ball, N. W., Borghmans, B. J., Smyth, J. A., Gelmini, L. and Lavazza, A. (2001a): Nucleotide sequence-based identification of a novel circovirus of canaries. *Avian Pathol.* **30**, 321–325.

- Todd, D., Weston, J. H., Soike, D. and Smyth, J. A. (2001*b*): Genome sequence determinations and analyses of novel circoviruses from goose and pigeon. *Virology* **286**, 354–362.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. (2012): Primer3 – new capabilities and interfaces. *Nucleic Acids Res.* **40**, 115.
- Uribe, C., Folch, H., Enriquez, R. and Moran, G. (2011): Innate and adaptive immunity in teleost fish: a review. *Veterinarni Medicina* **56**, 486–503.
- Zawar-Reza, P., Argüello-Astorga, G. R., Kraberger, S., Julian, L., Stainton, D., Broady, P. A. and Varsani, A. (2014): Diverse small circular single-stranded DNA viruses identified in a freshwater pond on the McMurdo Ice Shelf (Antarctica). *Infect. Genet. Evol.* **26**, 132–138.