Full genome sequence of a novel circo-like virus detected in an adult European eel *Anguilla anguilla* showing signs of cauliflower disease

Andor Doszpoly1,*, Zoltán L. Tarján1, Róbert Glávits2, Tamás Müller3, Mária Benkő1

1Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, PO Box 18, 1581 Budapest, Hungary

2Veterinary Diagnostic Directorate, National Food Chain Safety Office, Hungaria krt. 21, 1143 Budapest, Hungary

3Department of Fish Culture, Szent István University, 2100 Gödöllő, Hungary

ABSTRACT: An adult European eel *Anguilla anguilla*, showing typical signs of the so-called cauliflower disease, was subjected to pathological and molecular virological examinations. Samples taken from internal organs and the polypoid proliferative tissue from the mouth were examined by PCR for the detection of several viruses. Positive results were obtained with a nested PCR targeting the *rep* gene of circoviruses. Analysis of the partial *rep* sequence indicated the presence of a putative novel circovirus, but attempts to isolate it remained unsuccessful. The missing part of the genome was acquired by an inverse nested PCR with 2 specific primer pairs, designed from the newly determined *rep* sequence, followed by genome walking. The circular full genome was found to consist of 1378 nt (GenBank accession no. KC469701). Two oppositely oriented open reading frames (ORFs) were present, of which one was unambiguously identified as a circoviral *rep* gene. However, the predicted product of the other ORF, though it is a clear positional counterpart of the *cap* genes, showed no obvious homology to any known circoviral capsid proteins. A stem-loop-like element in the intergenic region between the 5’ ends of the ORFs was also found. Phylogenetic calculations indicated that the novel virus belongs to the genus *Circovirus* in the family *Circoviridae*. The relative amount of the viral DNA in the organ samples was estimated by quantitative real-time PCR. The results suggested that the examined fish was caught in an active viremic state, although the role of this circovirus in the etiology of the cauliflower diseases could not be ascertained.

KEY WORDS: Cauliflower disease · Stomatopapilloma · Fish circovirus · Complete genome · Phylogeny

INTRODUCTION

The cauliflower disease (stomatopapilloma or orocutaneous papillomatisis) of the European eel *Anguilla anguilla* was first described at the beginning of the 20th century (Wolf 1988). Sporadic occurrence of the disease, with no specificity for age and size of the affected fish, has been reported mostly from the tributary regions of northern European rivers (Delves-Broughton et al. 1980, Wolf 1988). The benign epidermal neoplasms usually grow on the head region, but lesions may occur on other parts of the body (Peters 1975). The papillomas consist of fibro-epithelium and are sometimes pigmented. The cause of the tissue proliferation is unknown. Different viruses, resembling birnaviruses, orthomyxoviruses, and rhabdoirus-like viruses, have been isolated from eels showing signs of the disease.
doviruses, have been detected in or isolated from eels with cauliflower disease, but experimental reproduction of the clinical signs has not been successful (Nagabayashi & Wolf 1979, Schwanz-Pfitzner et al. 1984, Ahne & Thomsen 1985, Ahne et al. 1987). Because of the seasonality of the tumor development, the effect of chemical pollutants and water temperature changes were assumed to contribute to the manifestation of the disease (Peters 1975).

The European eel, presently a Critically Endangered species (Freyhof & Kottelat 2010), was introduced to Lake Balaton (the largest lake in Central Europe, located in western Hungary) regularly from 1961 to 1991, except in 1985 and between 1989 and 1990. As a result, more than 83 million elvers (glass eel stage, at which feeding begins) were stocked in Lake Balaton. By the end of the 1980s, the lake had become overpopulated with eel, and there were 2 massive periods of eel mortality. These deaths were probably due to infection by the nematode worm Anguillicoloides crassus in 1991 and 1995 (Molnár et al. 1991). Thereafter, the Hungarian government prohibited the introduction of additional eels. At present, the eel stock in Lake Balaton is composed of very old individuals, of which even the youngest are >23 yr old.

In this work, the detection and full genome analysis of a novel circo-like virus are described. The virus was found in different organs of an eel from Lake Balaton which exhibited typical signs of the so-called cauliflower disease.

Circoviruses (CVs) are small, naked DNA viruses. The icosahedral virions, with a diameter of 12 to 26 nm, contain a circular single-stranded DNA (ssDNA) genome. The genome size of the CVs studied to date was found to range between 1.7 and 2.3 kb. The viral capsid is composed of 1 structural protein (capsid protein; Biagini et al. 2011). Targeted screenings with a sensitive consensus nested PCR (Halami et al. 2008) resulted in the recognition of a large and still sharply increasing number of novel CVs and circo-like viruses in an incredible diversity of specimens including natural and waste waters, faecal samples from different mammals (including bats and humans), as well as in several invertebrates and on the surface of a variety of raw meat products (Blinkova et al. 2009, Li et al. 2010, 2011). Up until a few years ago, only birds and swine had been known as frequent hosts for CVs. In the past several years, the vertebrate host spectrum has widened to include fish and amphibians (Lőrinicz et al. 2011, 2012, Tarján et al. 2014). However, in the majority of these cases, the eventual pathological role of the detected CVs is unknown.

**MATERIALS AND METHODS**

**Sample origin**

During a routine limnological survey, a ~50 cm long adult eel with cauliflower-like growths on both sides of its mouth was caught in Lake Balaton on 21 September 2009. After euthanasia, a necropsy was performed. Small pieces from the labial tissue proliferation, the gills, heart, liver, spleen, and kidney were collected in duplicate. One part of the samples was fixed in Bouin’s fixative, embedded in paraffin, sectioned (at 4–5 µm), stained with hematoxylin and eosin, and viewed by light microscopy according to standard procedures. The other set of samples were frozen and used for virus isolation and molecular studies.

**Virus isolation**

Virus isolation was attempted on the EK-1 cell line (Chen et al. 1982). The pooled organ homogenates were diluted to a 10% (w/v) suspension in an L-15 medium (Gibco) complemented with antibiotics (penicillin 300 U ml⁻¹, streptomycin 300 µg ml⁻¹). The suspension was centrifuged at 2000 × g (10 min), and the supernatant was immediately transferred into a new tube. Three parallel inoculations (500 µl suspension per flask) were made in 25 cm² flasks of EK-1 monolayers at 80% confluency. The flasks were incubated at 3 temperatures (15, 19, and 22°C) and checked for the appearance of cytopathic effect (CPE) daily.

**PCR**

For PCR, the nucleic acid extraction was made from approximately 25 mg of tissue as described in detail by Dân et al. (2003). The tissues were homogenized in 1× TE buffer, and 100 µl from each organ suspension was digested with Proteinase K (20 mg ml⁻¹) in the presence of sarcosyl (10%). After incubation with guanidine-hydrochloride (8 M) for 1 h at room temperature, the DNA was precipitated with ethanol. The presence of adenoviral DNA in the organ samples was tested by a sensitive nested PCR with consensus primers targeting the DNA-dependent DNA polymerase gene (Wellehan et al. 2004). The detection of herpes-, irido-, or poxviral DNA, a broad spectrum PCR method was used (Hanson et al. 2006). Demonstration of the circoviral DNA was
Doszpoly et al.: New circo-like virus genome sequence

attempted by a widely used and very efficient consensus nested PCR described by Halami et al. (2008). For the amplification of the missing part of the putative circular genome of the newly detected CV, 2 specific primer pairs were designed and used in a nested inverse PCR. The primer sequences were as follows: outer forward: 5’-GCC CTT GAG GAT TCT CAT TC-3’; outer reverse: 5’-CAG ATC GTT CCT CTT CCC TT-3’; inner forward: 5’-GAC TTT GGA TGG AAG AAG CC-3’; inner reverse: 5’-CCT TGT TAT GCT GGT CGT TG-3’.

The PCR program consisted of an initial denaturation step of 98°C for 5 min, followed by 45 cycles of 98°C for 30 s, 56°C for 30 s, 72°C for 60 s, and a final elongation cycle of 72°C for 3 min. The reaction mixture consisted of 34 µl of distilled water, 10 µl of 5× HF buffer (Phusion, Thermo Scientific), 0.5 µl of Phusion enzyme, 1 µl (50 µM) of each (forward and reverse) primer, 1.5 µl of dNTP solution of 10 mM concentration, and 2 µl of the target DNA in a final volume of 50 µl. The reactions were performed in a T1 Thermocycler (Biometra). The results of the PCRs were analyzed by electrophoresis in agarose gels.

Quantitative, real-time PCRs (qPCRs) for determining the relative amount of the viral DNA in 5 organs (the labial proliferative tissue, the gills, liver, spleen, and kidney) were carried out in an Applied Biosystems® StepOnePlus™ Real-Time PCR System instrument (Life Technologies). Two specific primers, suitable for the amplification of a 100 bp product, were designed. The reaction mixture contained 25 µl 2x PrimeSTAR Max Premix (Takara Bio), 17.5 µl distilled water, 2.5 µl EvaGreen™ Dye (Biotium), 1 µl of each primer (forward: 5’-AGG CAA CGA CCA GCA TAA CA-3’; reverse: 5’-AGT CGT CGA TGC AGG CCA AG-3’), and 3 µl of the target DNA in a final volume of 50 µl. The program consisted of an initial denaturing at 98°C for 5 min, followed by 40 cycles of 98°C for 10 s, 55°C for 5 s, and 72°C for 10 s. The beta-actin gene (with forward 5’-ACC GGT ATC GTC ATG GAC TC-3’; and reverse 5’-CGT CAG GGT CT TAT CAT CAG GT-3’ primers) was used as an internal standard. The results were analyzed by the StepOne Software v2.1 (Applied Biosystems).

Sequencing and sequence analyses

All PCR products were excised from the gels, purified with the QIAquick Gel Extraction Kit (Qiagen), and sequenced directly with the inner primers. The sequencing reactions were performed with the use of the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Electrophoresis was carried out by a commercial service provider on an ABI PRISM 3100 Genetic Analyzer. The amplification product from the inverse PCR, encompassing more than 1000 bp, was cloned with the use of the CloneJET Kit (Fermentas), and sequenced with the primers supplied with the kit. The newly obtained nucleotide sequences were compared with their homologs in GenBank by using different BLAST algorithms at the NCBI portal. The genome was assembled manually and confirmed with the use of the Staden Package as described elsewhere (Doszpoly & Shchelkunov 2010). The size and orientation of open reading frames (ORFs) and putative genes were examined after a 6-frame translation of the genomic DNA with the use of the JavaScript DNA Translator 1.1 program (Perry 2003).

Phylogeny inference

Phylogenetic calculations were performed online at the Mobyle portal (http://mobyle.pasteur.fr/cgi-bin/portal.py) of the Pasteur Institute (Paris) using the distance matrix analysis (Protdist) with the Jones-Taylor-Thornton matrix, and then Fitch calculations were performed with global rearrangements. The tree topology was tested by bootstrap analysis (Seqboot/1000 samplings, Protdist, Fitch, Consense).

RESULTS

Gross and microscopic pathology

On both sides of the mouth, we observed soft, polypoid, raspberry-colored masses (~1 cm in diameter; Fig. 1a). The fish was slightly emaciated. The only other gross lesions were restricted to the heart, where a small, flat, grayish, oval patch (~5 mm in diameter) was revealed on the epicardial surface which extended into the myocardium (Fig. 2a). Light microscopy did not reveal lesions in the gills or kidney. The labial polypoid masses were benign papillomas composed of proliferating connective tissue and covered by a multilayered epithelium. Proliferation of the malpighian cells upon a narrow base of connective tissue without an invasion down into the underlying dermis was also seen (Fig. 1b). The grayish spot on the heart was a benign myoma composed of collections of cross-striated muscle fibers and surrounded by
connective tissue (Fig. 2b). Additionally, vacuolization in the hepatocytes and multifocal hemosiderin deposition in the red pulp of the spleen were revealed.

**Virus isolation**

Rounding and detachment of the cells appeared on 12, 14, and 17 d post-inoculation in the tissue cultures incubated at 22, 19, and 15°C, respectively. In the negative controls, no CPE was observed. A few days later, the CPE was complete. However, in the second and third passages, no CPE was observed. PCR testing of samples from the successive passages for the detection of the CV or certain double-stranded DNA viruses (adeno-, herpes-, and iridoviruses) also remained negative.

**PCR and sequencing**

All PCRs performed for the detection of different large DNA viruses gave negative results. However, the circovirus PCR resulted in specific products from every examined organ (including the gills, liver, spleen, and kidney) as well as from the labial proliferative tissue. Interestingly, the gills, spleen, and labial tissue all initially displayed a heavy longitudinal smear, and discrete bands of amplicons were obtained after only 10× dilution of the target DNA solutions. The size of the identical amplicons was found to be 303 bp after editing out the primer sequences. By homology search, the conservative region of the circoviral replication-associated protein (Rep) was identified. The inverse nested PCR yielded an 1192 bp DNA fragment. It was first sequenced directly and then cloned and sequenced again for a
better quality outcome. This fragment contained the whole cap gene, the rest of the rep gene with the 2 intergenic regions.

**Genome sequence analysis**

The complete genome of the putative eel CV was found to encompass 1378 nt with an average G+C content of 48.7%. The full genome sequence was deposited in the GenBank database (accession number KC469701). The genome organization was found to be somewhat divergent from one typical of CVs, inasmuch as the putative cap gene proved to be significantly shorter (Fig. 3). Nonetheless, 2 oppositely oriented major ORFs which flanked a putative stem-loop element were found. The stem-loop element, situated in the 5' intergenic region, possessed a 12 bp stem. Its loop region consisted of 12 nucleotides with the conserved circoviral nonamer sequence (TAG TAT TAC), as shown in Fig. 4. The deduced product of the rep gene was predicted to consist of 286 amino acid (aa) residues containing the conserved RNA helicase domain. Its closest homolog in GenBank, with 51% aa identity, was the corresponding region of the CV recently described in the wels catfish Silurus glanis (Lőrincz et al. 2012). The other ORF, supposedly corresponding to the gene of the capsid protein (Cap), was predicted to code for 114 aa only. This protein did not show convincing homology to any proteins in GenBank, although similar to the Cap proteins of other known CVs, a 32 aa long arginine-rich stretch (RRR VYR KRS NRR PIR NCQ RRY RRP IRR ERN NR) was found close to the N terminus.

Quantitative PCR

With qPCR, the virus was successfully detected in all 5 of the examined organs. The statistical mean of 5 parallel qPCRs for the viral genome resulted in the following cycle threshold (Ct) values: liver 23.6, spleen 23.9, gills 26.0, kidney 27.5, stomatopapilloma 28.1. This means that the relative amount of the viral DNA in the examined organs is the following: 1 unit in the tissues of the stomatopapilloma, 1.25 unit in the kidney, 20 units in the gills, 130 units in the spleen, and 157 units in the liver.

Phylogenetic analysis

The phylogeny reconstruction made by the distance matrix analysis on complete deduced Rep sequences from 27 CVs is presented in Fig. 5. The alignments consisted of 229 aa. The newly detected putative European eel CV appeared in the clade of the Circovirus genus within the Circoviridae family.
as a sister group of the wels catfish CV (Lőrincz et al. 2012). However, the monophyly of fish CVs was not confirmed.

**DISCUSSION**

In the present work, the full genome sequence of a putative novel CV was determined. The virus originated from a European eel showing typical signs of the so-called cauliflower disease. We propose to name this virus the European eel circovirus (EeCV). The isolation of the virus on an eel kidney cell line failed. The CPE, observed after the first inoculation, did not appear again in the consecutive passages. We speculate that the cell degeneration was probably due to a direct cytotoxic effect by some material or unknown, uncharacterized virus that was not detectable by our PCRs.

Apart from the heart myoma and the stomatopapillomas, which were probably large enough to interfere with feeding, no other significant pathological alterations were revealed. It seems that the examined eel was caught in a viremic state, since every examined organ contained circoviral DNA. The ini-

![Fig. 4. Structure and sequence of the putative stem-loop of the European eel circovirus](image)

![Fig. 5. Phylogenetic tree reconstruction based on the distance matrix (JTT model) of the deduced amino acid sequences (229 aa) of the rep genes of single-stranded DNA viruses. Bootstrap values are shown on the branches. Circoviruses (CVs), fish CVs within the family Circoviridae, and cycloviruses (CyVs) are designated by different colored lines on the tree (black, blue, and red, respectively). GuCV: gull CV; PiCV: pigeon CV; StCV: starling CV; FiCV: finch CV; RaCV: raven CV; CaCV: canary CV; BKFD-1 (-2): beak and feather disease virus 1 (2); MDuCV: muscovy duck CV; MaCV: mallard CV; SwCV: swan CV; GoCV: goose CV; CanCV: canine CV; PCV-1 (-2): porcine CV1 (2); NG13: human stool-associated CV; BaCV: barbel CV; CfCV: catfish CV; EeCV: European eel CV; Fwc: Florida woods cockroach-associated CyV; Df-5 (-4, -1): dragonfly CyV 5 (4, 1); GF-4c: bat CyV; Chimp11: chimpanzee CyV 11; Bt-5: bat feces CyV; TN25: human feces CyV](image)
tial PCR with undiluted samples of the stomatopapilloma, gills, and spleen resulted in a heavy smear. This signaled the possibility of too high concentration of the total DNA. Indeed, the 10-fold dilution of the same samples produced discrete bands of the expected size.

Thanks to the significant improvement in the efficiency of detection methods (Halami et al. 2008), an incredible increase in the number of the known small circular ssDNA viruses has occurred in the second half of the past decade. The known host range of CVs (including mainly birds and swine) has also grown rapidly (Delwart & Li 2012). In addition to several mammals, a number of invertebrate hosts have also been found to harbor seemingly specific CVs. The first CVs from lower vertebrates have been published most recently. These viruses, detected in barbel Barbus barbus and wels catfish Silurus glanis, have been characterized by full genome analyses (Lárincz et al. 2011, 2012). Additional putative piscine CVs have been discovered by PCR in other fishes, as well as in 2 species of amphibians. However, these CVs have yet to be confirmed by complete genomic sequences (Fehér et al. 2013, Tarján et al. 2014). Although the barbel and wels catfish CVs have been found during investigations for the cause of increased mortality, the role of these viruses in disease could not be confirmed.

In this case, a direct connection between the presence of the putative circoviral DNA and the cauliflower disease could not be revealed. The porcine and most avian CVs are known to have an immunosuppressive effect (Todd 2004), which may exacerbate the pathogenicity of certain concurrent infectious agents. Nonetheless, it is also possible that CVs can cause generalized infection in individuals whose immune system is temporarily or permanently impaired. Ng et al. (2009) drew similar conclusions when they examined a novel small ssDNA virus in sea turtles with fibropapillomas.

The taxonomy of EeCV is somewhat ambiguous. According to the phylogeny reconstruction based on the Rep sequence, the piscine (barbel and wels catfish) CVs seem to form a common clade with an interesting small circular ssDNA virus. This common clade is formed with NG13, which was recently discovered by screening human fecal samples collected in Nigeria (Li et al. 2010). The NG13 virus had originally been described as an outlier to the proposed new genus Cyclovirus (Li et al. 2011); however, its exact taxonomic place is still unclear. The nonamer sequence of the stem-loop of NG13 is identical to that of the CVs, but its genome organization resembles that of the cycloviruses (CyVs) (Delwart & Li 2012). Our analysis shows that NG13 and fish CVs might share a close evolutionary origin. It seems that the EeCV does not form a monophyletic group with the other 2 fish CVs. A possible reason for such a result could be the incomplete taxon representation. Perhaps the full genome analyses of additional novel CVs from fishes and other lower vertebrates will help discern the phylogenetic relationships more precisely in the future.

The family Circoviridae is currently facing a radical revision. It has been proposed that the genus Gyrovirus be moved to the family Anelloviridae. Furthermore, the establishment of a new genus, Cyclovirus, within the Circoviridae has also been proposed (Biagini et al. 2013). However, at present there are numerous circovirus-like viruses with unknown ancestry.

Based on its genome organization, the putative EeCV most closely resembled the members of the genus Circovirus. It encodes a Rep protein on the virus sense strand and a putative Cap protein in the opposite direction. The 2 additional smaller ORFs exhibit a clear homology to the hypothetical genes of yet unknown function, described in pigeon (Mankertz et al. 2000) and duck CVs (Chen et al. 2006), and in the beak and feather disease virus (Niagro et al. 1998). These homologies further confirm the place of EeCV in the same genus. Furthermore, the stem-loop structure, which is involved in the initiation of the viral genome replication (Steinfeldt et al. 2001), had a typical circoviral nonamer sequence in EeCV, identical to that of the fish CVs, porcine CV1, pigeon, gull, and finch CVs (Li et al. 2010). The lengths of the intergenic regions between the rep and cap genes also show differences between CVs and CyVs (Li et al. 2010). These regions in EeCV are also characteristic for CVs. Its closest evolutionary relative is the wels catfish CV. Interestingly, the exceptionally short cap gene of EeCV did not show obvious homology to its counterparts in any known CVs or CyVs, although we did identify a characteristic feature of the circoviral Cap proteins, namely the arginine-rich stretch close to the N-terminus.

It is intriguing to speculate about a recombination event between a progenitor and an unknown circular virus resulting in EeCV. Among ssDNA viruses (even originating from distantly related virus families), recombination events occur frequently (Gibbs & Weiller 1999, Martin et al. 2011). Moreover, a recombination between a circo-like virus and ssRNA virus has also been hypothesized (Diemer & Stedman 2012). At this time, however, the origin of the cap gene of EeCV re-
remains enigmatic since the BLAST searches did not find any reliable homology to any known sequences among the environmental samples.

According to the proposed new rules of the Circoviridae Study Group of the International Committee on Taxonomy of Viruses, the comparative analysis of the complete genomes would replace the earlier practice of comparing the cap sequences in the establishment of new species (Biagini et al. 2013). Considering the demarcation criteria, EeCV unequivocally represents a novel species.

We must emphasize that our qPCR results are based on the examination of only 1 specimen collected at 1 time point. Thus, these data should be considered as preliminary results. The qPCR confirmed the results of the conventional PCR, inasmuch as all examined organs contained the viral DNA, suggesting that the examined eel was caught in a viremic state. The relative amount of the viral DNA may indicate that viral replication takes place in the liver and spleen.

To examine the pathogenicity of EeCV in experimental infections, the isolation of the virus would be essential. Screening eels, especially those showing clinical signs of cauliflower disease, for the presence of circoviral DNA should also be continued.

Acknowledgements. We are indebted to G. Reaney for a critical review of the manuscript. The financial support (OTKA PD104315) provided by the Hungarian Scientific Research Fund and Mohamed bin Zayed Species Conservation Fund (project no. 12252178) is gratefully appreciated. A.D. and T.M. are recipients of the János Bolyai Research Fund (project no. 12252178) is gratefully appreciated. The financial support from the Hungarian Scientific Research Fund and Mohamed bin Zayed Species Conservation Fund of the International Committee on Taxonomy of Viruses. The critical review of the manuscript. The financial support from the Hungarian Scientific Research Fund and Mohamed bin Zayed Species Conservation Fund of the International Committee on Taxonomy of Viruses. We must emphasize that our qPCR results are considered as preliminary results. The qPCR confirmed the results of the conventional PCR, inasmuch as all examined organs contained the viral DNA, suggesting that the examined eel was caught in a viremic state. The relative amount of the viral DNA may indicate that viral replication takes place in the liver and spleen.

To examine the pathogenicity of EeCV in experimental infections, the isolation of the virus would be essential. Screening eels, especially those showing clinical signs of cauliflower disease, for the presence of circoviral DNA should also be continued.

LITERATURE CITED


Gibbs MJ, Weillier GF (1999) Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. Proc Natl Acad Sci USA 96:8022−8027


Li L, Kapoor A, Slikas B, Bamidele OS and others (2010) Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. J Virol 84:1674−1682

Li L, Shao T, Bamidele OS, Masroor MA, Kunz TH, Zaidi SZ, Delwart E (2011) Possible cross-species transmission of circoviruses and cycloviruses among farm animals. J Gen Virol 92:768−772


Doszpoly et al.: New circo-like virus genome sequence

- Peters G (1975) Seasonal fluctuations in the incidence of epidermal papillomas of the European eel Anguilla anguilla L. J Fish Biol 7:415–422

Editorial responsibility: V. Gregory Chinchar, Jackson, Mississippi, USA

Submitted: November 28, 2013; Accepted: February 7, 2014
Proofs received from author(s): April 18, 2014