Gene 529 (2013) 288–295

Contents lists available at ScienceDirect

Gene

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

# Identification, genotyping, and molecular evolution analysis of duck circovirus

Zhilong Zhang [a,b,1], Renyong Jia [a,b,*,1], Yanyan Lu [a,b,1], Mingshu Wang [a,b,1], Dekang Zhu [a,b], Shun Chen [a,b], Zhongqiong Yin [b], Xiaoyue Chen [a,b], Anchun Cheng [a,b,**]

[a] Institute of Preventive Veterinary Medicine, Sichuan Agricultural University, Chengdu 611130, China
[b] College of Veterinary Medicine, Sichuan Agricultural University, 46 Xinkang Road, Ya'an, Sichuan 625014, China

ARTICLE INFO

*Article history:*
Accepted 6 July 2013
Available online 6 August 2013

*Keywords:*
Duck circovirus
Phylogenetic
Genotype
Evolution
Recombination
Selective pressure

ABSTRACT

Duck circovirus (DuCV) is a contagious immunosuppressive virus affecting many duck species, which is responsible for multiple outbreaks in poultry industries worldwide. In this study, the first DuCV isolate GH01 was identified in Sichuan by PCR, which shared a high level of nucleotide identity (81.8–99.4%) with sequences of other DuCV isolates available in GenBank. Comparative phylogenetic and pairwise sequence comparison analyses indicated that DuCV could be divided into two genotypes (DuCV-1 and DuCV-2) and six subtypes (1a, 1b, 1c, 2a, 2b and 2c) based on the complete genome sequence. The results revealed that both DuCV-1 and DuCV-2 had evolved from the same ancestor but undergone divergent evolution. Interestingly, phylogenetic analyses indicated that three isolates were classified into a cluster DuCV-2a using complete DuCV genome sequence and *cap* gene, except *rep* gene. Recombination analyses revealed that DuCV-2a arose from recombination between DuCV-1a and DuCV-2b isolates within the *rep* genes, and the recombination events mainly occur both in non-structural protein coding region and structural protein coding region. In addition, the mechanisms of recombination supporting the genetic variability in DuCV isolates were investigated. Likewise, selective pressure indicated that purifying selection had been a major driving force in maintaining diversity among the DuCV isolates. Because eradicating the virus from commercial ducks is impossible, it is necessary to take effective control measures and implement them throughout the world.


## 1. Introduction

The duck circovirus (DuCV) is a species tentatively placed in the genus *Circovirus* in the family Circoviridae, and the virion is icosahedral, non-enveloped, and 15 to 16 nm in diameter (Hattermann et al., 2003). DuCV was originally reported in two female 6-week-old Mulard ducks from a German farm. All had a feathering disorder and poor body condition (Hattermann et al., 2003; Soike et al., 2004). The virus had been reported in Hungary (Fringuelli et al., 2005), Taiwan (Chen et al., 2006), US (Banda et al., 2007) and Mainland China (Jiang et al., 2008; Zhang et al., 2009). DuCV has been detected in Muscovy, Mule and Pekin ducks, in which it causes stunting and feather abnormalities. Although controversial, lymphoid depletion predisposes the host to immunosuppression, and disease progression is further complicated by co-infections with other bacterial and viral pathogens (Soike et al., 2004; Xiang et al., 2012).

DuCV consists of a single-stranded circular DNA genome which contains about 1988–1996 nucleotides (nts) and two major open reading frames (ORFs) (Hattermann et al., 2003). One is ORF1, the *rep* gene, which encodes the replication associated protein required for the initiation of viral replication. The other is ORF2, the *cap* gene, which encodes the viral structural and virulence-associated protein involve in the host immune response. The intergenic regions contain a stem-loop, which is considered the site of the initiation of replication of viral DNA (Banda et al., 2007; Hattermann et al., 2003).

To date, despite the fact that a large number of complete DuCV genomes have been submitted to the GenBank database, no extensive phylogenetic studies have been performed on DuCV. This is probably attributable to the high level of homology among DuCV genomes. Currently, DuCV is prevalent in China, where it is associated with considerable economic losses in the poultry sector. The presence of DuCV in Chinese duck population has been reported, and several genotyping and genetic diversity studies have been performed on DuCV based on published complete genome sequences (Fu et al., 2011; Wang et al., 2011). Although there is a consensus that DuCVs can be divided into two major genotypes, but no evaluation of the utility of the *cap* and the *rep* genes as molecular markers. Furthermore,

*Abbreviations:* DuCV, duck circovirus; PCR, polymerase chain reaction; nt, nucleotide; PASC, pairwise sequence comparison; PRE, potential recombination event; dN, non-synonymous substitutions; dS, synonymous substitutions; ML, maximum-likelihood; NJ, neighbor-joining; MP, maximum parsimony.

* Correspondence to: R. Jia, Institute of Preventive Veterinary Medicine, Sichuan Agricultural University, Chengdu 611130, China. Tel.: +86 028 86291176; fax: +86 028 86291176.
** Correspondence to: A. Cheng, Institute of Preventive Veterinary Medicine, Sichuan Agricultural University, Chengdu 611130, China.
*E-mail addresses:* zhangzhilong0518@163.com (Z. Zhang), jiary@sicau.edu.cn (R. Jia), lu_yanyan5@163.com (Y. Lu), mshwang@163.com (M. Wang), zdk24@sicau.edu.cn (D. Zhu), sophia_cs@163.com (S. Chen), yinzhongq@163.com (Z. Yin), yfsy@sicau.edu.cn (X. Chen), chenganchun@vip.163.com (A. Cheng).
[1] These authors contributed equally to this work.

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.07.028

evolutionary studies are of increasing importance in genetic variability and viral pathogens.

Recombination is a major force for virus to gain genetic diversity, and plays a significant role in virus evolution together with selective pressure. Previous studies have provided evidences of natural recombination and adaptive selection in several members of the family Circoviridae, including torque teno virus (Manni et al., 2002), porcine circovirus (Ma et al., 2007; Olvera et al., 2007), and beak and feather disease virus (Heath et al., 2004; Julian et al., 2013). These results imply that recombination and selection pressure might also contribute to the duck circovirus in the molecular evolutionary.

In this study, we present comparative phylogenetic and exhaustive pairwise sequence comparison analyses to classify the DuCV isolates worldwide. Our findings revealed that a discordant cluster arose from recombination between DuCV-1a and DuCV-2b isolates within the *rep* genes using base-by-base comparison and recombination detection program. Moreover, we found five potential recombination events (PREs) from inter-genotypic and intra-genotypic recombination, mainly located in the *rep* gene and *cap* gene. Likewise, a second evolution mechanism, selective pressure, supporting the genetic variability of DuCV isolates.

## 2. Materials and methods

### 2.1. Sampling, screening and complete genome amplification of DuCV

Fifty tissue samples were collected from dead and diseased ducks at four previously DuCV-free duck farms in Guanghan, Sichuan. All sampled ducks were found to have the main clinical symptoms that manifested stunting and feather abnormalities. The collected tissue samples were subjected to nucleic acid extraction using standard phenol/chloroform/isoamyl alcohol protocols. The experimental procedures were performed in strict accordance with the Guidance Suggestions for the Care and Use of Laboratory Animals, and approved by the National Institute of Animal Health Animal Care and Use Committee of Sichuan Agricultural University (approval number 2012-032).

All samples were screened for DuCV by PCR using primers 5′-*CTCGAG*TGAACCCGGTGAACTGACC-3′ (1153 nt–1171 nt) and 5′-*GAATTC*ATGCGACGCAGCACCTATC-3′ (1924 nt–1906 nt). The product was analyzed on 1% agarose gel electrophoresis ultraviolet imaging. Positive sample was determined with 772 bp amplified product. From sample positive by PCR, the complete genome was amplified using primers 5′-TCC*GGATCC*GAAAAATCCAAATAC-3′ (1756 nt–1781 nt) and 5′-CCC*GGATCC*GGAACTGGACCAAC-3′ (1768 nt–1746 nt). The PCR product was purified with Easy Pure Quick Gel Extraction Kit (Tiangen Biotech) and ligated with pMD18-T Easy Vector System (Takara, Shanghai). To avoid misleading results caused by PCR artifacts, three random clones for each isolate were independently sequenced (Takara, Shanghai).

### 2.2. Sequencing analysis and construction of phylogenetic tree

After sequencing, the derived DNA sequences were compared to other sequences deposited in GenBank database using BLAST program to preliminarily confirm that all positive samples were true DuCV. Then, the open reading frames (ORFs) were identified according to the major criteria (Ettinger et al., 2012), a minimum length of 60 bp, an ATG start codon, and less than 60% overlap with adjacent ORFs using DNAstar 7.1 program. Complete DuCV sequence was submitted to the GenBank database. To investigate the phylogenetic inferences and molecular evolutionary analyses, all available complete DuCV genome, *cap*, and *rep* genes are obtained from GenBank or our study (Supplementary Table S1). The GoCV sequence was included in the tree as an outgroup. To determine the ancestral status of DuCV within the Circoviridae family, ten randomly selected DuCV genomes were aligned with representatives of the Circoviridae family, including porcine circovirus, goose circovirus, canary circovirus, beak and feather disease virus, and starling circovirus (Supplementary Table S2). Multiple sequence alignments were performed using ClustalX1.83 (Thompson et al., 1997). The degree of diversity among sequences was determined using molecular evolutionary genetics analysis (MEGA5) software (Tamura et al., 2011) and analyzing the alignment of the complete genome, *cap*, and *rep* genes through the neighbor-joining (NJ) method with p-distance (Grau-Roma et al., 2008). The neighbor-joining (NJ), maximum likelihood (ML), and maximum parsimony (MP) methods were used to construct the phylogenetic trees. However, only one tree is shown. Bootstrap values were estimated for 1000 replicates.

### 2.3. Genotype definition

The complete and partial sequences of DuCV were thought to be a good phylogenetic and epidemiologic marker for DuCV (Wang et al., 2011). p-Distance values of the two genotypes of DuCVs were calculated separately using the pairwise sequences comparison results of their complete genome (Munir et al., 2012; Varsani et al., 2011). p-Distance/frequency histograms were constructed to allow the determination of potential cut-off values to distinguish different genotypes of DuCV, and only the values that agree with the phylogenetic tree were considered as the exact cut-off values. The neighbor-joining method for simplified ultrametric tree-drawing and a bootstrap resampling of 1000 replications were performed. Frequency distribution analysis was conducted using Microsoft Excel, as described previously (Grau-Roma et al., 2008; Segalés et al., 2008).

### 2.4. Recombination among DuCV sequences

Systematic screening for the presence of potential recombination, likely parent sequences, and localization of possible recombination breakpoints were performed using nucleotide alignments and recombination detection program v4 (RDP4) (Martin et al., 2010). The aligned sequences were examined using six recombination detection methods (RDP, GeneConv, BootScan, MaxChi, Chimaera and SiScan) as implemented in RDP4. The six methods all used the following general settings: window size = 20, highest acceptable p-value = 0.001 and Bonferroni correction. Recombination events were judged to be credible if they were detected by three or more methods with significant p-values and a base-by-base comparison (Hesse et al., 2008) was used to find further proof of the recombination events.

### 2.5. Selection pressure on the DuCV

The existence of selective pressures along the genome was assessed first by calculating the difference between non-synonymous substitution (dN) and synonymous substitution (dS) rates for the aligned *cap* and *rep* genes. If the rate of dN is higher than the dS, positive selection ($dN - dS > 0$) is said to be operating. Alternatively, if the rate of dN is lower than the dS ($dN - dS < 0$), the gene is said to be under purifying selection (Hughes and Nei, 1989). We employed the programs MEGA5 (Tamura et al., 2011) to compare the rate for the aligned *cap*, *rep*, *rep*-ORF3 and ORF3 genes (as ORF3 was completely within *rep* gene and coded on the minus strand, the rates for *rep* gene were analyzed both with and without inclusion of the ORF3 region). As an additional measure of genetic complexity, entropy was calculated. Because entropy takes into account the number of distinct viral variants, it is less susceptible than genetic distance to variation as a result of hypermutation (Wolinsky et al., 1996). The amino acid sequence entropy was calculated using BioEdit and plotted against the difference between dN and dS.