Infection, Genetics and Evolution 28 (2014) 333–338

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Genome characterization of a novel chicken picornavirus distantly related to the members of genus *Avihepatovirus* with a single 2A protein and a megrivirus-like 3′ UTR

Ákos Boros [a], Péter Pankovics [a], Ádám Adonyi [a], Tung Gia Phan [b,c], Eric Delwart [b,c], Gábor Reuter [a,*]

[a] Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Pécs, Hungary
[b] Blood Systems Research Institute, San Francisco, CA, USA
[c] University of California, San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 6 August 2014
Received in revised form 22 October 2014
Accepted 27 October 2014
Available online 1 November 2014

Keywords:
Picornavirus
Recombination
IRES
3′ UTR
Pathogen
Avian

ABSTRACT

The members of the genus *Avihepatovirus* and related picornaviruses (“Aalivius”) of ducks, turkey and chickens possess identical 2A peptide composition including three functionally unrelated 2A peptides which is a characteristic genome feature of these monophyletic avian picornaviruses. The complete genome of a novel picornavirus provisionally named Orivirus A1 (KM203656) from a cloacal sample of a 4-week-old diarrheic chicken (*Gallus gallus domesticus*) distantly related to members of genus *Avihepatovirus* was characterized. The study strain contains a type-II-like IRES, a single 2A protein of unknown function unrelated to the 2A proteins of avihepatoviruses and a long 3′ untranslated region (UTR) with multiple repeated sequence motifs followed by an AUG-rich region. The repeated sequences of the 3′ UTR show significant identity to the “Unit A” sequences of the phylogenetically distant megriviruses. The presence of a novel single 2A and the megrivirus-like “Unit A” motifs suggest multiple recombination events in the evolution of this novel picornavirus.

1. Introduction

Members of family *Picornaviridae* consist of small viruses with single-stranded RNA genome of positive polarity and currently belong to 46 species grouped into 26 genera and several candidate species waiting for official classification (Knowles et al., 2012; Adams et al., 2013; www.picornaviridae.com). Picornavirus genomes possess some general features like (i) the capsid protein-encoding P1 followed by the 2BC–P3 non-structural proteins, (ii) the (predominantly) $3C^{pro}$-dependent polyprotein-processing scheme, (iii) the 3′ untranslated region (UTR) followed by the poly(A) tail, (iv) the presence of a 5′ Internal Ribosomal Entry Site (IRES), and (v) a terminal VPg. These features are characteristics of all currently known picornavirus strains (Palmenberg et al., 2010).

The genus *Avihepatovirus* and related genera (*Avisivirus* and the proposed “Aalivirus”) contain several monophyletic avian picornaviruses identified from domestic ducks – *Anas plathyrinchos* (Duck Hepatitis A viruses – DHAVs of genus *Avihepatovirus* and “Aalivirus A1” – AalV-A1 – of genus “Aalivirus”), domestic turkey – *Meleagris gallopavo* (Avisivirus A1 of genus *Avisivirus*) and from domestic chickens – *Gallus gallus domesticus* (Chicken picornavirus 2 and 3 which were possibly belonged to genus *Avisivirus*). These picornaviruses possess a Leaderless genome, uncleaved VP0 capsid protein and a multicistronic 2A genome region which build up of an DxExNPGP “Ribosome-skipping” site-containing aphthovirus-like 2A1, P-loop NTPase-containing 2A2 and an Hbox-NC-type 2A3 in a row supporting the common origin of these viruses (Tseng and Tsai, 2007; Wang et al., 2014; Boros et al., 2013; Lau et al., 2014).

In this study, a novel picornavirus from a broiler chicken provisionally named as Orivirus A1 (OrV-A1) was identified and characterized. OrV-A1 was distantly related to the members of genera *Avihepatovirus* and “Aalivirus”. Orivirus contains a type-II-like IRES, a 3-3-4 genome organization pattern with a single 2A protein-coding genome region and a long 3′ untranslated region (UTR) with a megrivirus-like organization (multiple repeated “Unit A” motif followed by an AUG-rich region) (Boros et al., 2014).

2. Materials and methods

A single cloacal sample (Pf-CHK-1) was collected from a 4-week-old broiler chicken (*Gallus gallus domesticus*) from a small

* Corresponding author at: Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Szabadság út 7, H-7623 Pécs, Hungary. Tel.: +36 06 (72) 514 979.
E-mail address: reuter.gabor@ddr.antsz.hu (G. Reuter).

http://dx.doi.org/10.1016/j.meegid.2014.10.025
1567-1348/

poultry colony in Orosháza, Hungary in April 2013, showed clinical signs of diarrhea. The sample was subjected to a viral metagenomic analysis using sequence independent random RT-PCR amplification of viral-particle protected nucleic acids. The viral cDNA library was constructed using ScriptSeqTM v2 RNA-Seq Library Preparation Kit (Epicentre) and sequenced by the Miseq Illumina platform, as described previously (Phan et al., 2013). For the determination of the complete picornavirus genome (chicken/Pf-CHK1/2013/HUN) and for the verification of the metagenomic contigs, long-range and conventional RT-PCR amplification, 5′/3′ rapid amplification of cDNA ends (RACE) and dye-terminator sequencing were used as previously described (Boros et al., 2011, 2012b). The metagenomic contigs served as templates for the virus-specific primer design. The primers used for the RT-PCR based amplification and sequencing of the genome fragments were available on request. Generic $3D^{pol}$ primers (PfCHK1-3D-F: 5′ CTT ATC GMA TGG TSA TGG GTG A 3′ and PfCHK1-3D-R: 5′ CCY CKG CAC CAC TGG ATC TT 3′) were designed based upon the $3D^{pol}$ sequences of chicken/Pf-CHK1/2013/HUN and DHAVs for screening the available cloacal samples ($N = 11$) collected from apparently healthy chickens.

The nucleotide (nt) and amino acid (aa) pairwise alignments and identity calculations were performed by the BioEdit software (version 7.1.3.0) (Hall, 1999) using the in-built Clustal W algorithm (Thompson et al., 1994). For the secondary structure analysis of the 2B protein, the PSIPRED Protein Sequence Analysis Workbench with the PSIPRED v3.3 prediction method was used (Jones, 1999). The aa phylogenetic trees were constructed using the neighbor-joining method based on the Jones–Taylor–Thornton matrix-based model of MEGA software (ver. 6.06) (Tamura et al., 2013). Bootstrap values (based on 1000 replicates) for each node are shown if >50%.

The secondary RNA structure of the 5′ UTR-IRES as well as the structure of the 3′ UTR was predicted by the Mfold program, visualized using the VARNA software ver. 3.9 and the Corel Draw Graphic Suite (Ver. 12) (Zucker, 2003; Darty et al., 2009).

3. Results

The *in silico* analysis of the metagenomic sequence reads resulted seven non-overlapping sequence contigs covering 37% of a picornavirus genome which was related to DHAV-2 strain 04G (GenBank ID: EF067923) as the closest relative (<55% amino acid identity) identified by GenBank BLASTx search. The 7037-nucleotide(nt)-long RNA genome of the chicken picornavirus strain chicken/Pf-CHK1/2013/HUN, (GenBank ID: KM203656) was predicted to have a 3-3-4 genome organization pattern: 5′ UTR-P1(VP0-VP3-VP1)-P2(2A-2B-2C)-P3(3A-3B-3C-3D)-3′ UTR (Fig. 1). The G + C content (49.81%) of the entire genome is significantly higher than the DHAVs (42.50–43.70) (Tseng and Tsai, 2007).

The genome regions of P1 (2124 nt; 708 aa), P2 (1596 nt; 532 aa) and P3 (2205 nt; 734 aa) show the closest sequence relationship to different DHAV and AalV-A1 strains (Table 1). The potential proteolytic cleavage sites of chicken/Pf-CHK1/2013/HUN were predicted based on the aa alignment with the three genotypes of DHAVs (DHAV1 strain R85952, GenBank ID: NC_008250; DHAV-2 strain 04G, GenBank ID: EF067923 and DHAV-3 strain AP-04114, GenBank ID: DQ812093). The mapped cleavage sites together with the length of the genome regions were shown in Fig. 1.

The sequence analysis of the N-terminal end of the viral polyprotein did not support the presence of leader protein. The N-terminal P1 genome region predicted to encode three capsid proteins: VP0 without an identifiable N-terminal myristoylation motif (GxxxS/T, *X* = variable aa), VP3 and VP1. Similar capsid structure was found among the members of genus *Avihepatovirus*, *Avisivirus* and "Aalivirus" (Tseng and Tsai, 2007; Boros et al., 2013; Lau et al., 2014; Wang et al., 2014).

The P2 genome region predicted to encode only three (2A, 2B and 2C) mature peptides. The single 93-aa-long 2A with currently unknown function possesses neither the DxExNPGP "ribosome-skipping" motif, nor the GxxGxGKS motif of P-loop NTPase-type 2A nor the Hbox/NC motif. The 2A peptides containing these three motifs are characteristic genome features of related picornaviruses of genera *Avihepatovirus*, *Avisivirus* and "Aalivirus" (Tseng and Tsai, 2007; Boros et al., 2013; Wang et al., 2014). The 115-aa-long 2B peptide was predicted to contain two transmembrane helixes (data not shown) which suggests the pore formation capability of this peptide similar to other picornaviral 2Bs (Palmenberg et al., 2010). Based on the presence of all three functional motifs (A–C) the 2C protein of the study strain – similar to the other picornaviral 2C proteins – probably belongs to the class III helicases (Fig. 1) (Hales et al., 2007).

The P3 genome region encodes four (3A to 3D) viral peptides, a single $3B^{VPg}$ peptide with a conservative Y (tyrosine) at the 3rd position followed by the picornaviral 3C proteinase and 3D RNA-dependent RNA polymerase in which the conserved aa motifs were recognizable (Fig. 1) (Gorbalenya et al., 1989).

The phylogenetic analysis was performed using the aa sequences of the complete P1, 2C and 3D genome regions of chicken/Pf-CHK1/2013/HUN and the representative members of the *Picornaviridae* family (P1, Fig. 2A) and the closest relatives of the study strain (2C and 3D trees, Fig. 2B and C). All three phylogenetic trees show the consequent but distant relationship of chicken/Pf-CHK1/2013/HUN to the members of genera *Avihepatovirus*, *Avisivirus* and "Aalivirus" (Fig. 2).

The predicted length of the 5′ UTR of chicken/Pf-CHK1/2013/HUN was 734 nt based on the presence of the last in-frame AUG initiation codon found in Kozak-context (GaaGCC$\underline{A}_{735}\underline{UG}$, conserved nts are in uppercase). Significant (86%) sequence identity was found between the 5′ UTR (from nt pos. 377 to 434) of the study strain and the 5′ UTR (from nt position 276 to 333) of turkey avisivirus strain USA-IN1 (GenBank ID: KC614703) using GenBank BlastN. This region contains the apical part of domain I of the type-II-like IRES of avisiviruses (Boros et al., 2013). Based upon this sequence identity and the predicted secondary structure of the 3′ part of 5′ UTR, the study sequence has a potential type-II-like IRES where all five major core domains from H to L identified in the type-II IRES of encephalomyocarditis virus (EMCV, genus *Cardiovirus*) were recognizable (Yu et al., 2011) (Fig. 3A). Beside the structural similarity, the binding sites of pyrimidine tract binding protein (PTB) and translation initiation factor eIF4G as well as sequence motifs essential for type-II IRES function i.e. GNRA, RAAA motifs identified in EMCV were all recognizable in the IRES of the study strain (Kaminski et al., 1995; Kolupaeva et al., 1998; López de Quinto and Martínez-Salas, 1997) (Fig. 3A).

The 378-nt-long 3′ UTR of chicken/Pf-CHK1/2013/HUN contains three consecutively repeated conserved sequence motif called "Unit A" which were identified first among the chicken and turkey megriviruses as well as pigeon mesiviruses (*Megrivirus*) which were phylogenetically distant from the study virus (Fig. 3B and C) (Boros et al., 2014). The "Unit A" sequences are involved in the build-up of a multi-loop structure in the 3′ UTR of chicken/Pf-CHK1/2013/HUN followed by 99-nt-long (from nt position 251) AUG-rich region where the cytosine content is low (7%) (Fig. 3B). The AUG-rich region builds up a single stem-loop where the apical loop region could form a pseudoknot with the loop of the third "Unit A" (Fig. 3B).

The study virus was present in six (54%) of the 11 chicken cloacal samples using RT-PCR method, which could indicate that chickens are the natural hosts of this virus.