Archives of Virology
https://doi.org/10.1007/s00705-018-3761-6

ANNOTATED SEQUENCE RECORD

Complete coding sequence of a novel picorna-like virus in a blackbird infected with Usutu virus

Steven Van Borm[1] · Mieke Steensels[1] · Elisabeth Mathijs[1] · Claude Kwe Yinda[2] · Jelle Matthijnssens[2] · Bénédicte Lambrecht[1]

Received: 18 December 2017 / Accepted: 22 January 2018

Abstract

Using random high-throughput RNA sequencing, the complete coding sequence of a novel picorna-like virus (a 9,228-nt contig containing 212,202 reads) was determined from a blackbird (*Turdus merula*) infected with Usutu virus. This sequence shares only 36% amino acid sequence identity with its closest homolog, arivirus 1, (an unclassified member of the order *Picornavirales*), and shares its dicistronic genome arrangement. The new virus was therefore tentatively named "blackbird arilivirus" (ari-like virus). The nearly complete genome sequence consists of at least 9,228 nt and contains two open reading frames (ORFs) encoding the nonstructural polyprotein (2235 amino acids) and structural polyprotein (769 amino acids). Two TaqMan RT-qPCR assays specific for ORF1 confirmed the presence of high levels of this novel virus in the original sample. Nucleotide composition analysis suggests that blackbird arilivirus is of dietary (plant) origin.

As in other European countries, passive surveillance demonstrated the circulation of Usutu virus in blackbirds and other wild and captive birds in Belgium in 2016 and 2017 [3, 11]. The original purpose of this study was to determine the complete genome sequence of an Usutu virus strain from a blackbird that had been found dead after displaying neurological symptoms in August 2017 in Brussels and had tested positive by real-time RT-PCR. A pool of brain, lung and intestinal tissue was homogenized (10% [wt/vol]) in 1x phosphate-buffered saline and pretreated by passage through a 0.45-µM-pore-size filter and nuclease treatment [10]. RNA was extracted using a NucleoSpin RNA Virus kit, (Macherey Nagel), followed by DNase treatment (TURBO DNase, Thermo Fisher Scientific). cDNA was synthesized using SuperScript IV reverse transcriptase (Thermo Fisher Scientific) and random hexamer primers, followed by synthesis of double-stranded cDNA using the NEBNext mRNA Second Strand Synthesis Module (New England Biolabs). A sequencing library was prepared using a Nextera XT kit (Illumina) and sequenced using a MiSeq Reagent Kit version 3 (Illumina) with 2 × 300-bp paired-end sequencing. Raw sequence data were trimmed using Trim Galore! (q = 30, l = 50, paired; https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). A random paired subset (2 × 50,000 reads) was used for *de novo* assembly using SPAdes v3.9.0 [1] and IVA v1.0.0. [5]. Contigs longer than1000 nt were identified in a BLASTn search, followed by a BLASTx translated search if no nucleotide-level similarities were found. The full trimmed dataset was mapped to relevant contigs using BWA-0.6.2 [8]. In addition to a 10,998-bp contig representing the complete Usutu virus genome (46,925 reads; average coverage, 848x, reported elsewhere), a 9,228-bp contig (212,202 reads; average coverage, 4681x; Fig. 1) represented a novel virus. This contig shared only 36% amino acid sequence identity with arivirus 1, a highly divergent member of the order *Picornavirales* that was discovered recently in human diarrhea samples [4]. This sequence is likely to represent the nearly complete genome. Two TaqMan RT-qPCR assays specific for the putative ORF1 of this picorna-like virus (Supplementary Table 1) confirmed the presence of high levels of viral

Handeling Editor: Zhenhai Chen.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-018-3761-6) contains supplementary material, which is available to authorized users.

✉ Steven Van Borm
Steven.vanborm@coda-cerva.be

[1] Directorate Viral Diseases, Veterinary and Agrochemical Research Center CODA-CERVA, Brussels, Belgium

[2] Laboratory of Viral Metagenomics, Department of Microbiology and Immunology, Rega Institute, KU Leuven, Leuven, Belgium

Published online: 13 February 2018

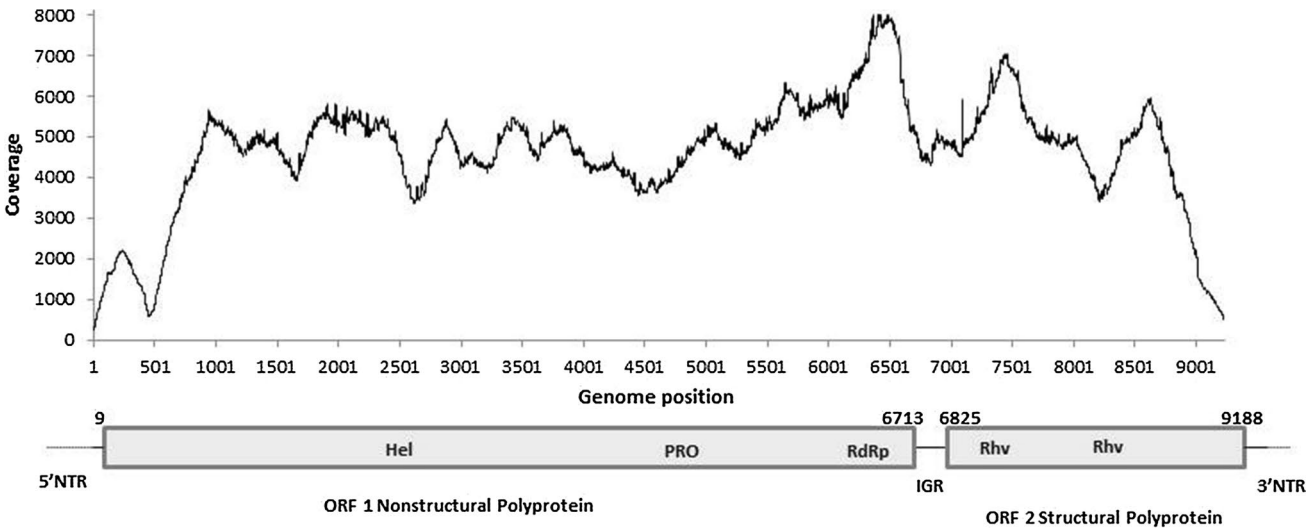

Fig. 1 Genome coverage and organization of blackbird arilivirus. Hel, RNA helicase; PRO, 3C cysteine protease; RdRp, RNA-dependent RNA polymerase; Rhv, picornavirus capsid protein

nucleic acid (Ct, 20.07-22.21) in the original sample. A retrospective screening of 47 archived blackbird samples collected in 2016-2017 failed to consistently identify additional positive cases. Only eight weakly positive brain and intestinal samples (Ct, 32.92-38.48) were identified using the RT-qPCR assay targeting the RdRp domain, while the assay targeting the protease domain was negative for all tested samples. However, the samples available for retrospective testing mostly consisted of brain tissue, which may have contributed to the low detection rate. The complete coding sequence of this virus, which we have tentatively named "blackbird arilivirus" (ari-like virus), has been deposited in the GenBank database under accession number MG662717.

Two ORFs, separated by a short intergenic region of 11 nt, were identified and found to be homologous to the putative arivirus nonstructural polyprotein ORF (2235 amino acids) and the structural polyprotein ORF (788 amino acids), respectively. Conserved sequence elements were predicted using NCBI's conserved domain database [9], Pfam [2] and Phyre2 [6]. ORF1 displays the typical picornaviral replication block layout

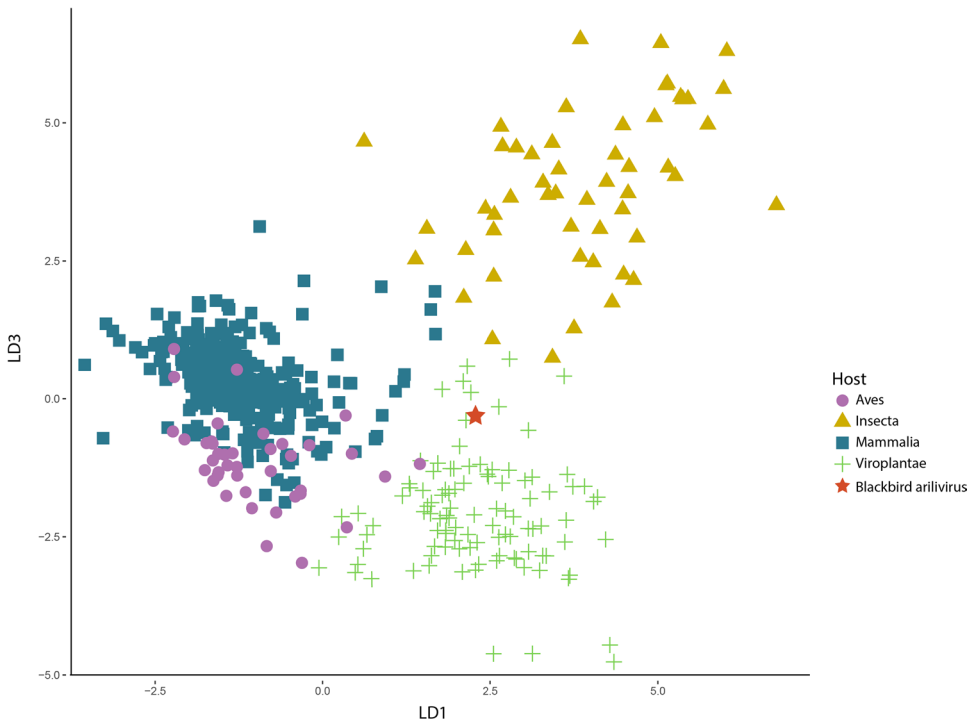

Fig. 2 Linear discriminant analysis for the classification of viral sequences into host groups using the most influential factors. Points represent values for individual sequences

with the predicted domains showing 49, 30 and 46% amino acid sequence identity to the RNA helicase (Hel), proteinase (PRO), and RNA-dependent RNA polymerase (RdRp) non-structural protein domains, respectively, of arivirus 1. ORF2 displays two Rhv-like picornavirus capsid domains (showing 46 and 33% amino acid sequence identity, to arivirus 1) (Fig. 1). The order *Picornavirales* is a rapidly increasing taxonomic unit of positive-stranded RNA viruses with a wide range of hosts including invertebrates, plants and vertebrates [7]. We performed a nucleotide composition analysis as described previously [12], including in our database members of the order *Picornavirales* that infect avian hosts. This analysis suggested that blackbird arilivirus is probably of plant origin (Fig. 2).

Acknowledgements DVM. F. Hénin kindly collected the bird and circumstantial information. The expert technical assistance of O. Ozhelvaci, S. H. Mahibullah and A. Ausloos is much appreciated. Raw sequence data were generated by the Biotechnology and Molecular Biology Platform of the Scientific Institute for Public Health (Brussels, Belgium).

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477
2. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230
3. Garigliany M, Linden A, Gilliau G, Levy E, Sarlet M, Franssen M, Benzarti E, Derouaux A, Francis F, Desmecht D (2017) Usutu virus, Belgium, 2016. Infect Genet Evol 48:116–119
4. Holtz LR, Cao S, Zhao G, Bauer IK, Denno DM, Klein EJ, Antonio M, Stine OC, Snelling TL, Kirkwood CD, Wang D (2014) Geographic variation in the eukaryotic virome of human diarrhea. Virology 468–470:556–564
5. Hunt M, Gall A, Ong SH, Brener J, Ferns B, Goulder P, Nastouli E, Keane JA, Kellam P, Otto TD (2015) IVA: accurate de novo assembly of RNA virus genomes. Bioinformatics 31:2374–2376
6. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858
7. Koonin EV, Wolf YI, Nagasaki K, Dolja VV (2008) The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nat Rev Microbiol 6:925–939
8. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760
9. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226
10. Rosseel T, Ozhelvaci O, Freimanis G, Van Borm S (2015) Evaluation of convenient pretreatment protocols for RNA virus metagenomics in serum and tissue samples. J Virol Methods 222:72–80
11. Van Borm S, Lambrecht B, Vandenbussche F, Steensels M (2017) Complete coding sequence of Usutu virus strain *Gracula religiosa*/U1609393/Belgium/2016 obtained from the brain tissue of an infected captive common hill myna (*Gracula religiosa*). Genome Announc 5(12):e00042-17
12. Yinda CK, Zell R, Deboutte W, Zeller M, Conceicao-Neto N, Heylen E, Maes P, Knowles NJ, Ghogomu SM, Van Ranst M, Matthijnssens J (2017) Highly diverse population of Picornaviridae and other members of the Picornavirales, in Cameroonian fruit bats. BMC Genom 18:249