Virus Genes (2019) 55:815–824
https://doi.org/10.1007/s11262-019-01703-w

ORIGINAL PAPER

Identification of a distinct lineage of aviadenovirus from crane feces

Yahiro Mukai[1,2] · Yuriko Tomita[3] · Kirill Kryukov[4] · So Nakagawa[4] · Makoto Ozawa[5,6,7] · Tsutomu Matsui[8] · Keizo Tomonaga[1,2,9] · Tadashi Imanishi[4] · Yoshihiro Kawaoka[10,11] · Tokiko Watanabe[3] · Masayuki Horie[1,12]

Received: 12 April 2019 / Revised: 29 August 2019 / Accepted: 3 September 2019 / Published online: 23 September 2019

Abstract

Viruses are believed to be ubiquitous; however, the diversity of viruses is largely unknown because of the bias of previous research toward pathogenic viruses. Deep sequencing is a promising and unbiased approach to detect viruses from animal-derived materials. Although cranes are known to be infected by several viruses such as influenza A viruses, previous studies targeted limited species of viruses, and thus viruses that infect cranes have not been extensively studied. In this study, we collected crane fecal samples in the Izumi plain in Japan, which is an overwintering site for cranes, and performed metagenomic shotgun sequencing analyses. We detected aviadenovirus-like sequences in the fecal samples and tentatively named the discovered virus crane-associated adenovirus 1 (CrAdV-1). We determined that our sequence accounted for approximately three-fourths of the estimated CrAdV-1 genome size (33,245 bp). The GC content of CrAdV-1 genome is 34.1%, which is considerably lower than that of other aviadenoviruses. Phylogenetic analyses revealed that CrAdV-1 clusters with members of the genus *Aviadenovirus*, but is distantly related to the previously identified aviadenoviruses. The protein sequence divergence between the DNA polymerase of CrAdV-1 and those of other aviadenoviruses is 45.2–46.8%. Based on these results and the species demarcation for the family *Adenoviridae*, we propose that CrAdV-1 be classified as a new species in the genus *Aviadenovirus*. Results of this study contribute to a deeper understanding of the diversity and evolution of viruses and provide additional information on viruses that infect cranes, which might lead to protection of the endangered species of cranes.

Keywords Adenovirus · Aviadenovirus · Crane · Feces · Metagenomics

Introduction

Virus research has long focused mainly on viruses that cause disease in humans, as well as those that are economically important to animals and plants, and thus a large number of viruses that do not cause disease or cause mild symptoms in hosts have been overlooked. Considering the entire virosphere, presently identified viruses are believed to be just the tip of the iceberg: for example, a previous study estimated that only 0.005% of viruses have been discovered thus far [1]. Therefore, for a deeper understanding of the diversity and evolution of viruses, it is important to continue searching for unidentified viruses.

Unbiased detection is important for exploring the virosphere. Deep sequencing is a promising and unbiased approach to detect novel viruses in organisms or from environmental samples [2]. Several reports have shown that this approach is also useful for noninvasive investigation of viruses in animal-derived samples such as fecal specimens from wild animals [3]. Indeed, many viruses that are probably non-pathogenic or mildly pathogenic have been detected in animal samples by deep sequencing [2].

The Izumi plain (Fig. 1) is an overwintering site for many birds, including cranes and ducks. Over 10,000 cranes spend the winter in this plain, including the hooded crane (*Grus monacha*), white-naped crane (*G. vipio*), sandhill crane (*G. canadensis*), common crane (*G. grus*), Siberian crane (*G. leucogeranus*), and demoiselle crane (*G. virgo*) [4]. *G. monacha* and *G. vipio*, which are the

Edited by William Dundon.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11262-019-01703-w) contains supplementary material, which is available to authorized users.

✉ Tokiko Watanabe
tokikow@ims.u-tokyo.ac.jp

✉ Masayuki Horie
horie.masayuki.3m@kyoto-u.ac.jp

Extended author information available on the last page of the article

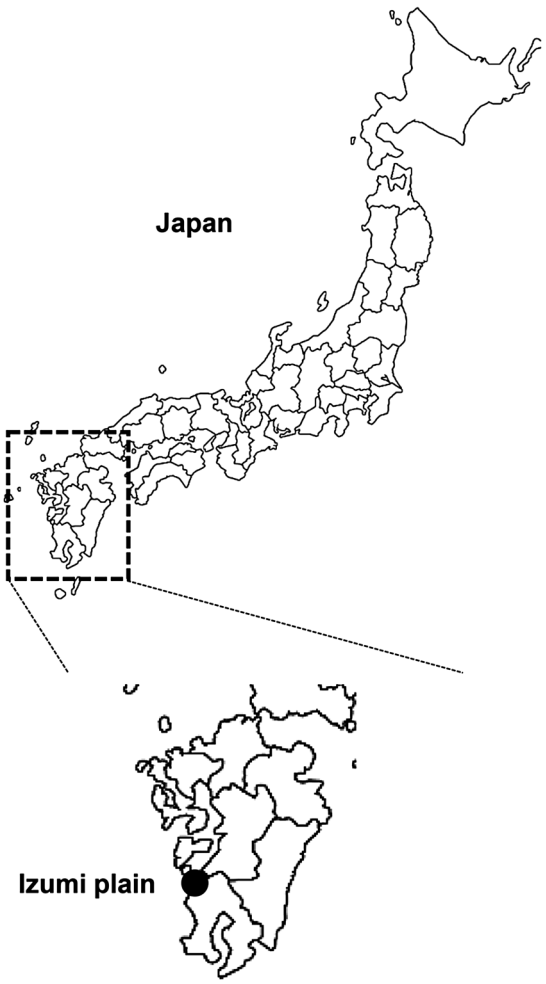

Fig. 1 Location of the Izumi plain. The location of the Izumi plain is indicated by the black circle

major crane species in the plain, are also listed as vulnerable species in the International Union for Conservation of Nature (IUCN) Red List (https://www.iucnredlist.org).

Although surveillance of cranes for known viruses such as influenza A viruses [5–8] has been conducted at this site, the viruses targeted have been limited. Thus, unbiased virus surveillance in this site could lead to the discovery of novel viruses.

In this study, to identify novel viruses, we collected crane fecal samples from this site and analyzed them by deep sequencing. We found a novel adenovirus and tentatively named it crane-associated adenovirus 1 (CrAdV-1). CrAdV-1 clusters phylogenetically with members of the genus *Aviadenovirus*, but is distantly related to known viruses in the genus. We propose that this virus be considered a novel species in the genus *Aviadenovirus*. Results of this study can aid in our understanding of the diversity and evolution of viruses.

Materials and methods

Sample collection

On 20 November 2016, six fecal samples were collected in the Izumi plain, Japan. Feces were obtained with disposable plastic spoons and introduced into sterile 2-ml sample tubes, which were placed on ice. The fecal samples were transported to the laboratory in a cooler with ice packs for storage at – 80 °C.

Metagenomic analysis of fecal samples

DNA was extracted from fecal samples using QIAamp DNA Stool Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. Library construction and deep sequencing were performed on the paired-end Illumina HiSeq sequencing platform (HiSeq 2500, 2 × 150 bp) by GENEWIZ (USA). The obtained sequences were registered in the DDBJ Sequence Read Archive (DRA) (accession numbers DRR173096-DRR173101).

Identification of the host species

To identify the host origin of the fecal samples, 34,400 bird barcode sequences of the mitochondrial gene, cytochrome oxidase subunit 1 (COI), were downloaded from the Barcode of Life Data (BOLD) System [9]. Using the entire sequence from each sample, a BLASTn search [10] was performed against the downloaded sequences. The sensitivity parameters used were as follows: -evalue 1e−10 -dbsize 3,200,000,000. The hits were then filtered to retain only those with a bit score of at least 210, and species associated with sequences showing the highest bit score were assumed to be the host species from which each feces sample was obtained.

Identification of virus-like sequences in the deep sequencing data

The sequence reads were individually preprocessed using fastp [11] and then co-assembled using metaSPAdes 3.11.1 [12] with multiple k-mers (55, 65, and 77). The resultant contigs were used for the downstream analyses. The resultant contigs with equal to or greater than 1000 nucleotides were extracted using SeqKit [13], which were used for the following three-step virus-like sequence similarity searches. First, using the extracted contigs as queries, an MMseqs2 search [14] was performed with the RefSeq viral protein database as a target database (downloaded from NCBI Viral Genomes, https://www.ncbi.nlm.nih.gov/genome/viruses/).

Table 1 Primers used in this study

Name	Sequence (5′ to 3′)	Note
YM_ta_1	CTCTCCAGTTCTCATTGGTCTTGTTG	Amplicon-seq
YM_ta_2	GGTGCAAGATCGGTTGCTCC	Amplicon-seq
YM_ta_3	GGGTGGTAGTGGTGCATTTGC	Amplicon-seq
YM_ta_4	GCTACTGGAGCTGCACGTG	Amplicon-seq
YM_ta_5	AGATCAGCTTAATCTACTCAGGTACCTG	Amplicon-seq
YM_ta_6	TCCCATATCTAACACCCAACTGTCTC	Amplicon-seq
YM_ta_7	CGCCCAGCAAATTGGAGAAATAGA	Amplicon-seq
YM_ta_8	GAAGCTTGCGAGGATGTCTCC	Amplicon-seq
YM_ta_9	CCTCCACCTTCAACACCATCC	Amplicon-seq
YM_ta_10	GTCTGGAACTTCTACATGTGAAGGTG	Amplicon-seq
YM_ta_11	AAAATGGTGGTCCAATAGATAGTGATGC	Amplicon-seq
YM_ta_12	GGTTATTCATGGCTGGCATGGATTAC	Amplicon-seq
YM_ta_13	CCAACTGCAGGACATGCTTCAC	Amplicon-seq
YM_ta_14	CTACGTCAGTCAGTGGGTGTATACTC	Amplicon-seq
YM_ta_15	TCCCACAATTTCATTTGAATGTTGCAAC	Amplicon-seq
YM_ta_16	AGCACATCCAAACCCAAGAACAG	Amplicon-seq
YM_ta_17	ATGTGTGTTAACCCAAACATCACCAC	Amplicon-seq
YM_ta_18	GGTTGTGGACTTGAAGGAAGAGG	Amplicon-seq
HKB212	GGTCATTGTCCTTTTGCAGC	Nested-PCR
HKB213	TCTGCTAATGCTTCCACGTC	Nested-PCR
HKB214	CCACAGTACTTGGGCCTTAC	Nested-PCR
HKB215	GTCCACACTAATCTGCCACC	Nested-PCR

In response to the sequence similarity search results, the second MMseqs2 search was conducted against the RefSeq protein database. Finally, a BLASTx [10] search was performed against the NCBI nr database using sequences with the best hits for viral sequences in the second MMseqs2 searches as queries. The query sequences with the best hits for viral proteins were regarded as virus-like contigs and analyzed in detail manually.

Amplicon-seq

PCR and amplicon sequencing were conducted to determine the CrAdV-1 genome sequence. PCR was performed with 30 ng of DNA using Phusion Green Hot Start II High-Fidelity DNA Polymerase (Thermo Fisher Scientific, F537S) and the primers listed in Table 1. The PCR conditions were as follows: an initial denaturation at 98 °C for 30 s; 30 cycles of 98 °C for 10 s, 66 °C for 30 s, and 72 °C for 2 min 30 s; and a final extension at 72 °C for 10 min. The bands obtained with expected sizes were purified and then mixed. A library was then prepared from the mixed amplicons using a NEBNext Ultra II DNA Library Prep Kit (NEB, E7645S) and sequenced on the Illumina NextSeq platform.

The obtained sequence reads were preprocessed using fastp [11], and then 1% of the preprocessed reads were extracted in SeqKit [13]. The extracted reads were assembled using SPAdes [15], and the resulting contigs were analyzed by BLASTn searching against the adenovirus-like contigs identified in the first metagenomic analysis. The obtained amplicon-seq data and final contig were deposited in DDBJ (accession numbers DRR173095 and LC469780, respectively).

Gene annotation

A BLASTx search was performed against RefSeq protein sequences of viruses in the genus *Aviadenovirus* (taxid: 10552) using the partial genome sequence of CrAdV-1 as a query with the options word size 2 and e-value threshold 1e−4. The results of the BLASTx analysis were used for manual annotations. Splicing sites were deduced from those present in the genomes of other aviadenoviruses.

Mapping metagenomic reads to the CrAdV-1 contig

The preprocessed reads were individually mapped to the CrAdV-1 contig using Bowtie 2 version 2.3.4.3 [16], and the numbers of mapped reads were counted in BamTools 2.5.1 [17].

Nested-PCR for detection of crane-associated adenovirus 1

The first PCR was carried out using 120 ng of DNA as a template with Phusion Green Hot Start II High-Fidelity DNA Polymerase and primers HKB212 and HKB213 (Table 1). The PCR conditions were as follows: a first denaturation at 98 °C for 30 s; 30 cycles of 98 °C for 10 s, 61 °C for 30 s, and 72 °C for 40 s; and a final extension at 72 °C for 10 min. The PCR products were purified using innuPREP PCRpure Kit (Analytik Jena). The second PCR was performed using 1 out of 50 µl of each eluate as a template with Green Hot Start II High-Fidelity DNA Polymerase and primers HKB214 and HKB215 (Table 1). The PCR products were analyzed in a MultiNA DNA 1000 (Shimadzu, Japan) microchip electrophoresis system.

Molecular evolutionary analyses

Phylogenetic analyses were performed using DNA polymerase or hexon gene sequences obtained in this study and those listed in Supplementary Table 1. The deduced amino

acid sequences were aligned using the L-INS-i algorithm in MAFFT v7.407 [18] and then phylogenetically analyzed. The multiple alignment files are available in Supplementary Materials. Phylogenetic relationships of the DNA polymerase and hexon genes were inferred using the maximum likelihood method with LG + G + I and LG + G + F models, respectively, that were determined by the Find DNA/Protein models function in MEGA X [19]. The reliabilities of the trees were assessed by 100 bootstrap replicates.

Pairwise distances between the DNA polymerase sequence of CrAdV-1 and other aviadenoviruses were computed in MEGA X using the *p*-distance model.

Results

Determination of the origins of samples from deep sequencing reads

To explore viruses that infect cranes, we collected six samples that were visually identified as crane feces in the Izumi plane in Japan (Fig. 1). We extracted DNA from the samples and performed deep sequencing on an Illumina HiSeq platform. We obtained approximately 30.5–40.5 M read pairs for each sample.

To investigate the host species from which each fecal sample was obtained, we compared the sequence reads of each sample with bird barcodes (the cytochrome oxidase subunit 1 gene in the BLOD database). The results of this analysis are summarized in Supplementary Table 2. The most common BLAST hit from sequence reads of the six samples (S85, S86, S87, S89, S90, and S91) was *G. monacha* (hooded crane), which is the predominant crane species in the Izumi plain [4]. Although *G. grus* was the second most common species identified from the six samples, only a few common cranes were observed in the Izumi plain [4]. Therefore, we assumed that each of the six fecal samples was obtained from the *Grus* genus and likely all were obtained from *G. monacha*.

Detection of adenoviral sequences in crane fecal samples by deep sequencing

To detect viral sequences in the fecal samples, we performed DNA-seq analyses (Fig. 2a). We first co-assembled all the sequence reads. We extracted the resulting cross-contigs that were greater than or equal to 1,000 nucleotides in length and performed homology search-based analyses using these contig sequences. We found that the best BLASTx hits for the four contigs were for aviadenovirus proteins (data not shown), suggesting that these contigs were derived from aviadenoviruses in the crane feces. Interestingly, the tentatively annotated DNA polymerase of the detected virus-like sequences showed relatively low amino acid identities (approximately 46–52%) with sequences of known adenoviruses, strongly suggesting that these sequences may be derived from a novel adenovirus. We tentatively named this virus CrAdV-1 and further analyzed the sequence.

Determination and analysis of the crane-associated adenovirus 1 genome sequence

Next, we attempted to fill the gaps between the contigs to determine the CrAdV-1 genome sequence (Fig. 2a). To do so, we designed primers that bound to the identified contigs and performed PCR and an amplicon-seq analysis using one of the CrAdV-1 positive samples (sample ID 293; Table 2) as a template. We assembled 1% of the preprocessed sequence reads, which was optimal for assembly (data not shown), and obtained an adenovirus contig of 33,245 nucleotides in length with an average coverage of approximately 292×. Based on genome sizes of other aviadenoviruses (about 41–45 kb), this contig probably covers approximately three-fourths of the CrAdV-1 genome.

We next analyzed the features of the CrAdV-1 genome. The GC content of the genome is very low (34.1%) compared to that of other aviadenoviruses (44.7–66.9%) (Table 3). CrAdV-1 genome contains at least 25 open reading frames homologous to those in other aviadenoviruses (Fig. 2b), and the central region of the CrAdV-1 genome is almost identical to those of other aviadenoviruses (Fig. 2b). Some aviadenoviruses have been reported to encode only one fiber gene, whereas others encode two fiber genes [20]. CrAdV-1 apparently encodes one fiber gene (Fig. 2b). We detected a tandem repeat at the right end of the identified contig (Fig. 2b). To exclude the possibility of misassembly, we also performed PCR and direct Sanger sequencing for this region, and confirmed that the repeats are indeed present in the CrAdV-1 genome (data not shown). This repeat consists of 297-bp repetition units showing 98.7% nucleotide identity.

Molecular evolutionary analyses of CrAdV-1

To gain insights into evolutionary relationships among adenoviruses, we performed phylogenetic analyses of DNA polymerase and hexon genes. The DNA polymerase tree showed that CrAdV-1 clusters with members of the genus *Aviadenovirus* but is distantly related to known aviadenoviruses (Fig. 3a). The hexon gene tree showed a topology

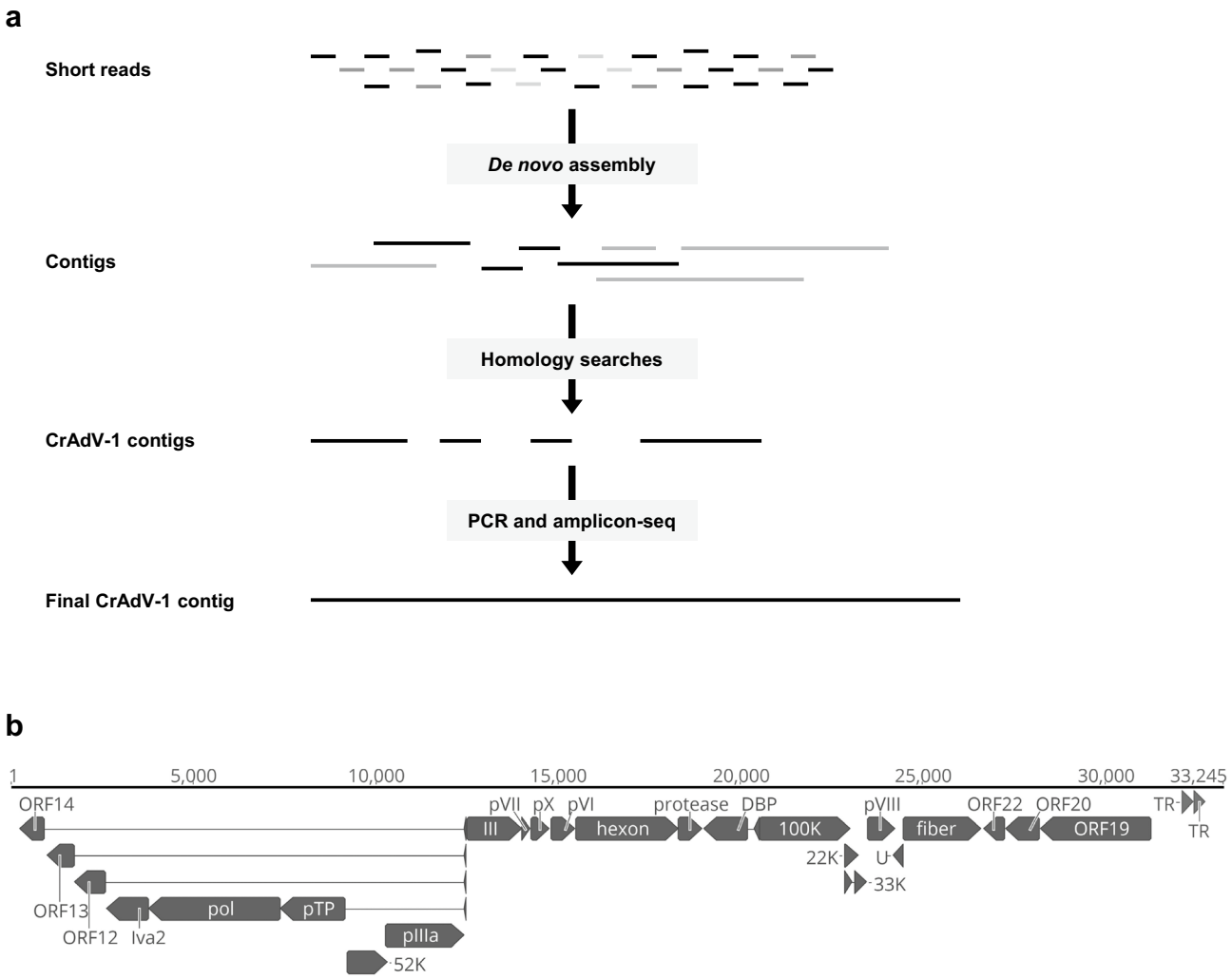

Fig. 2 Genome organization of CrAdV-1. **a** Strategy to determine the CrAdV-1 genome. **b** CrAdV-1 genome organization. The numbers show nucleotide positions in the CrAdV-1 contig (accession number LC469780). The boxes indicate open reading frames, and the thin lines show putative introns. *TR* tandem repeat

Table 2 Crane fecal samples used in this study

Sample ID	DRA accession number	Number of mapped read pairs	PCR
293	DRR173095	174	+
302	DRR173096	0	−
331	DRR173097	4	+
332	DRR173098	0	−
345	DRR173099	4	+
346	DRR173100	2016	ND

ND not done

similar to that of the DNA polymerase tree (Fig. 3b). We also calculated pairwise distances between CrAdV-1 and other aviadenovirus sequences, and showed that these distance ranged approximately 45–47% (Table 4). These results revealed that CrAdV-1 is an aviadenovirus distantly related to known aviadenoviruses.

Prevalence of crane-associated adenovirus 1 in the collected fecal samples

To investigate the prevalence of CrAdV-1 in the six fecal samples collected, we first mapped the metagenomic sequence reads from each sample to the CrAdV-1 genome. We found that two of the six samples (samples 293 and 346) contained many sequence reads that mapped to the CrAdV-1 genome (Table 2), suggesting that they were derived from CrAdV-1-infected individuals. Another two samples (samples 302 and 332) did not contain any reads that mapped to

Table 3 G+C contents of aviadenoviruses

Virus name	Accession number	GC%
Crane-associated adenovirus 1	LC469780	34.2
Goose adenovirus 4	NC_017979.1	44.7
Duck adenovirus 2	NC_024486.1	46.1
Turkey adenovirus 4	NC_022612.1	48.5
Pigeon adenovirus 2	NC_031503.1	48.6
Turkey adenovirus 5	NC_022613.1	51.6
Psittacine aviadenovirus B	NC_039032.1	52.0
Fowl adenovirus D	AC_000013.1	53.8
Fowl adenovirus D	NC_000899.1	53.8
Fowl adenovirus A	AC_000014.1	54.3
Fowl adenovirus A	NC_001720.1	54.3
Fowl adenovirus C	NC_015323.1	54.6
Fowl adenovirus 5	NC_021221.1	56.5
Fowl adenovirus 6	NC_038332.1	57.9
Fowl adenovirus E	NC_014969.1	57.9
Pigeon adenovirus 1	NC_024474.1	63.8
Turkey adenovirus 1	NC_014564.2	66.9

the CrAdV-1 genome. The remaining two samples (331 and 345) contained two read pairs that mapped to the CrAdV-1 genome.

To clarify whether the samples really contained CrAdV-1 DNA, we performed a nested-PCR analysis on five of the six samples, with primers targeting the CrAdV-1 DNA polymerase gene (Due to a shortage of the sample, 346 was not included in this analysis). We detected bands of the expected size from three samples: 293, 331, and 345 (Fig. 4). These results revealed that at least four samples, 293, 331, 345, and 346, were positive for CrAdV-1 DNA.

Discussion

In this study, we analyzed fecal samples from cranes in the Izumi plain of Japan and discovered a novel adenovirus distantly related to known adenoviruses. To the best of our knowledge, this study is the first to detect an adenovirus in crane-related samples. Importantly, the DNA polymerase gene of CrAdV-1 is distantly related to other aviadenoviruses: the pairwise distances of DNA polymerase amino acid sequences from CrAdV-1 and other aviadenoviruses are 45.2–46.8% (Table 4), meeting the species demarcation criterion for adenoviruses [21] (also available at https://talk.ictvonline.org/ictv-reports/ictv_9th_report/dsdna viruses-2011/w/dsdna_viruses/93/adenoviridae). In addition, the nucleotide composition of CrAdV-1 (34.2% G+C) is drastically different from that of other aviadenoviruses (44.7–66.9% G+C) (Table 3). Thus, CrAdV-1 is an aviadenovirus that is distinct from other aviadenoviruses. Based on these data, CrAdV-1 is a novel species in the genus *Aviadenovirus*.

Fig. 3 Phylogenetic tree of adenoviruses. Maximum likelihood trees were reconstructed using amino acid sequences deduced from the adenovirus DNA polymerase (**a**) and hexon (**b**) genes. The black circles indicate CrAdV-1. Bootstrap values less than 70% were omitted. Scale bars indicate amino acid substitutions per site

The discovery of CrAdV-1 indicates the existence of unidentified aviadenoviruses. The phylogenetic distances between CrAdV-1 and other aviadenoviruses are large relative to those between known aviadenoviruses (Fig. 3), suggesting that viruses that fill the phylogenetic gaps between them may exist, as previously proposed [2]. Therefore, further studies should be undertaken to discover novel virus species in the genus *Aviadenovirus*.

CrAdV-1 showed extremely low GC content (34.2% G+C) compared to the other aviadenoviruses (44.7–66.9% G+C). A previous study has shown that GC content is positively correlated with the genome size of adenoviruses [22]. Although genome sizes do not vary extensively among members of the genus *Aviadenovirus*, it would be still interesting to identify the whole genome sequence of this virus.

There is a tandem repeat sequence at the right end of the identified contig (Fig. 2b). Tandem repeats are also present in the genomes of several other aviadenoviruses [23–26]. Interestingly, the tandem repeats are all located in a similar region: they are present at the right side of the genomes. The biological roles of these tandem repeats are so far unknown [27]. Unfortunately, we could not determine the genome ends, and therefore were not able to assess the coding capacity of this region. Further sequencing and molecular biological analyses are needed to understand the biological function of the tandem repeat.

Here, we showed the effectiveness of using deep sequencing to identify the origins of fecal samples. In the Izumi plain, several species of cranes and ducks are living together [5]. We collected presumed crane feces based on their sizes, but this presumption may have been inaccurate. However, our deep sequencing analyses showed that the samples were indeed derived from cranes, and probably from *G. monacha* specifically (Supplementary Table 2). Although we cannot completely exclude the possibility that the viral nucleic acids were contaminants from other sources, our data strongly suggest that CrAdV-1 was present in the feces of *G. monacha*. Note that *G. monacha* might not be a host for CrAdV-1 because we cannot completely exclude the possibility that the virus merely infected the diet of the cranes. However, considering the small genomic size of aviadenoviruses (~ 46 kb) and the presence of tremendous amounts of bacterial DNA, which are much larger than viral DNA, the proportion of CrAdV-1 reads in sample 346 was considerably high

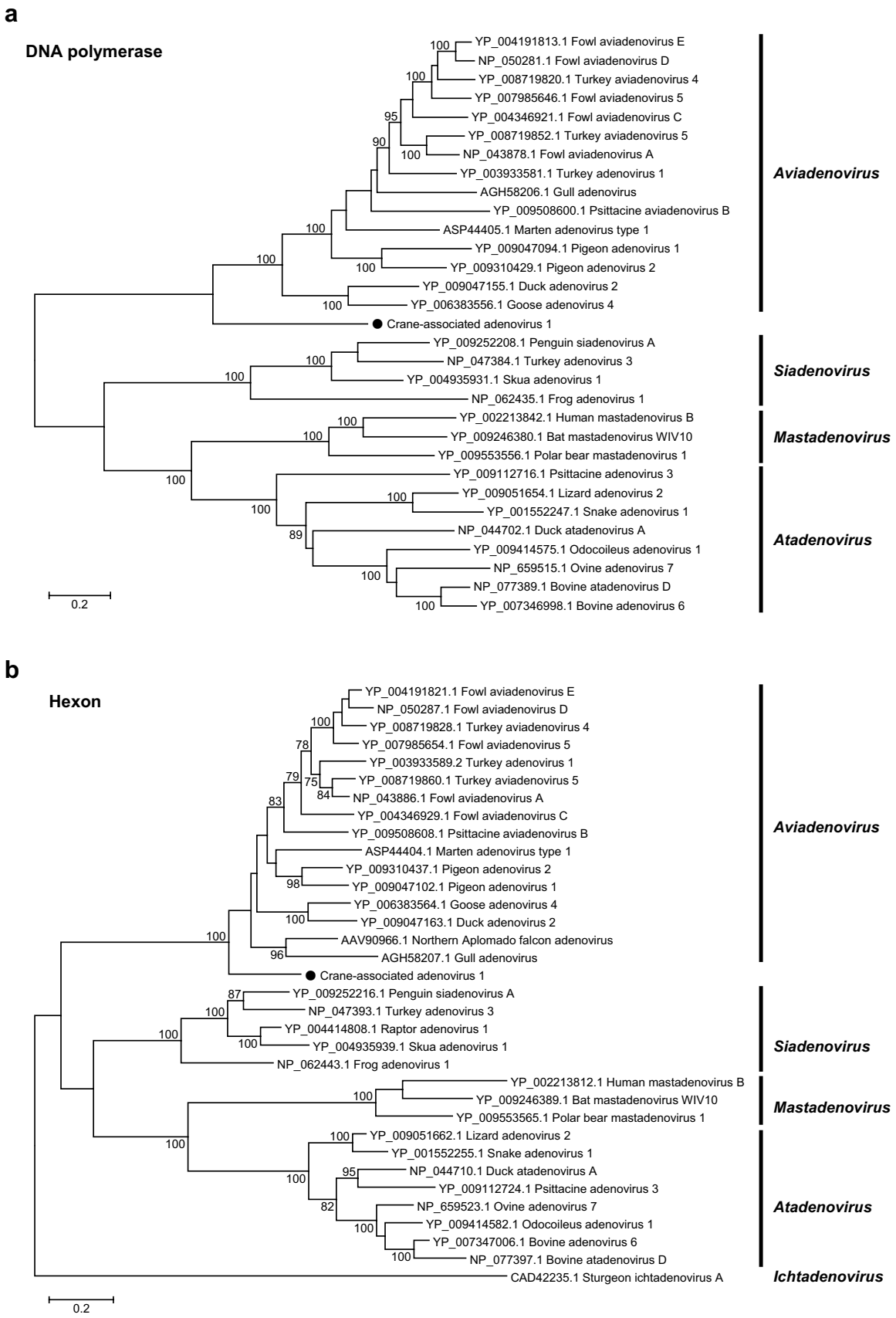
a
DNA polymerase
YP_004191813.1 Fowl aviadenovirus E
NP_050281.1 Fowl aviadenovirus D
YP_008719820.1 Turkey aviadenovirus 4
YP_007985646.1 Fowl aviadenovirus 5
YP_004346921.1 Fowl aviadenovirus C
YP_008719852.1 Turkey aviadenovirus 5
NP_043878.1 Fowl aviadenovirus A
YP_003933581.1 Turkey adenovirus 1
AGH58206.1 Gull adenovirus
YP_009508600.1 Psittacine aviadenovirus B
ASP44405.1 Marten adenovirus type 1
YP_009047094.1 Pigeon adenovirus 1
YP_009310429.1 Pigeon adenovirus 2
YP_009047155.1 Duck adenovirus 2
YP_006383556.1 Goose adenovirus 4
Crane-associated adenovirus 1
YP_009252208.1 Penguin siadenovirus A
NP_047384.1 Turkey adenovirus 3
YP_004935931.1 Skua adenovirus 1
NP_062435.1 Frog adenovirus 1
YP_002213842.1 Human mastadenovirus B
YP_009246380.1 Bat mastadenovirus WIV10
YP_009553556.1 Polar bear mastadenovirus 1
YP_009112716.1 Psittacine adenovirus 3
YP_009051654.1 Lizard adenovirus 2
YP_001552247.1 Snake adenovirus 1
NP_044702.1 Duck atadenovirus A
YP_009414575.1 Odocoileus adenovirus 1
NP_659515.1 Ovine adenovirus 7
NP_077389.1 Bovine atadenovirus D
YP_007346998.1 Bovine adenovirus 6
0.2
Aviadenovirus
Siadenovirus
Mastadenovirus
Atadenovirus
b
Hexon
YP_004191821.1 Fowl aviadenovirus E
NP_050287.1 Fowl aviadenovirus D
YP_008719828.1 Turkey aviadenovirus 4
YP_007985654.1 Fowl aviadenovirus 5
YP_003933589.2 Turkey adenovirus 1
YP_008719860.1 Turkey aviadenovirus 5
NP_043886.1 Fowl aviadenovirus A
YP_004346929.1 Fowl aviadenovirus C
YP_009508608.1 Psittacine aviadenovirus B
ASP44404.1 Marten adenovirus type 1
YP_009310437.1 Pigeon adenovirus 2
YP_009047102.1 Pigeon adenovirus 1
YP_006383564.1 Goose adenovirus 4
YP_009047163.1 Duck adenovirus 2
AAV90966.1 Northern Aplomado falcon adenovirus
AGH58207.1 Gull adenovirus
Crane-associated adenovirus 1
YP_009252216.1 Penguin siadenovirus A
NP_047393.1 Turkey adenovirus 3
YP_004414808.1 Raptor adenovirus 1
YP_004935939.1 Skua adenovirus 1
NP_062443.1 Frog adenovirus 1
YP_002213812.1 Human mastadenovirus B
YP_009246389.1 Bat mastadenovirus WIV10
YP_009553565.1 Polar bear mastadenovirus 1
YP_009051662.1 Lizard adenovirus 2
YP_001552255.1 Snake adenovirus 1
NP_044710.1 Duck atadenovirus A
YP_009112724.1 Psittacine adenovirus 3
NP_659523.1 Ovine adenovirus 7
YP_009414582.1 Odocoileus adenovirus 1
YP_007347006.1 Bovine adenovirus 6
NP_077397.1 Bovine atadenovirus D
CAD42235.1 Sturgeon ichtadenovirus A
0.2
Aviadenovirus
Siadenovirus
Mastadenovirus
Atadenovirus
Ichtadenovirus

Table 4 Pairwise distance between crane adenovirus 1 and other aviadenoviruses

Virus (species) name	Accession number	aa pairwise distance
Psittacine aviadenovirus B	YP 009508600.1	0.468
Pigeon adenovirus 1	YP 009047094.1	0.468
Pigeon adenovirus 2	YP 009310429.1	0.468
Turkey adenovirus 1	YP 003933581.1	0.468
Marten adenovirus type 1	ASP44405.1	0.466
Turkey aviadenovirus 5	YP 008719852.1	0.465
Goose adenovirus 4	YP 006383556.1	0.464
Fowl aviadenovirus D	NP 050281.1	0.458
Turkey aviadenovirus 4	YP 008719820.1	0.456
Fowl aviadenovirus A	NP 043878.1	0.456
Fowl aviadenovirus 5	YP 007985646.1	0.455
Duck adenovirus 2	YP 009047155.1	0.455
Fowl aviadenovirus E	YP 004191813.1	0.454
Gull adenovirus	AGH58206.1	0.454
Fowl aviadenovirus C	YP 004346921.1	0.452

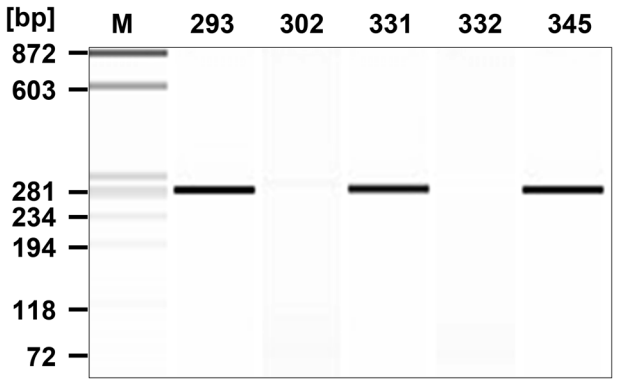

Fig. 4 Detection of CrAdV-1 DNA in fecal samples. A nested-PCR analysis was performed using primers targeting the CrAdV-1 DNA polymerase gene. The PCR products were analyzed using the MultiNA DNA 1000 microchip electrophoresis system, and the pseudo gel image is shown. Sample numbers are provided above the image. *M* DNA ladder

(2016 reads were mapped to the CrAdV-1 contig, which accounts for approximately 0.0054% of the total reads; Table 2). Additionally, previous studies have shown no evidence that *G. monacha* eats small birds at the sampling site [28–30]. Finally, CrAdV-1 DNA is also detected in conjunctival swabs of *G. monacha* (Ozawa et al., unpublished data). These data suggest that CrAdV-1 infected *G. monacha*. However, further studies, using other samples such as those from several organs, are needed to confirm the hosts of CrAdV-1.

Assuming that CrAdV-1 infects *G. monacha*, our analyses also suggest that CrAdV-1 is circulating in the crane population. Although the sample size was small, our analyses revealed that at least four of the six crane fecal samples contained CrAdV-1 DNA (Fig. 4; Table 2). These data suggest that multiple individuals were infected with CrAdV-1. Further molecular epidemiological studies are needed to understand the biology of CrAdV-1.

The discovery of CrAdV-1 might be also important from the perspective of protecting threatened species. In this study, we detected CrAdV-1 in cranes, probably *G. monacha*, which is listed as a vulnerable species in the IUCN Red List (https://www.iucnredlist.org). Additionally, the Izumi plane is an overwintering site for several other species of cranes, including *G. vipio*, which is also an IUCN vulnerable species. Some adenoviruses, such as fowl adenoviruses, are known to be pathogenic [31]. Although it remains unclear whether CrAdV-1 causes disease, these baseline data may be important for the protection of cranes. Further studies are needed to address this issue.

Taken together, results from this study contribute to a deeper understanding of the diversity and evolution of viruses, and may contribute to the conservation of threatened species. We focused only on eukaryotic DNA viruses and not on others such as bacteriophages and RNA viruses. Therefore, further analyses may reveal additional novel viruses in cranes.

Acknowledgements We thank Dr. Keiko Takemoto (Kyoto University) for her kind help with the bioinformatic analyses. We also thank Naomi Fujimoto and Mikiko Tanaka for technical support. The supercomputing resource was provided by the Human Genome Center (University of Tokyo). This study was supported by the Hakubi Project at Kyoto University (MH); a Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Science, Sports, and Technology (MEXT) of Japan, Grant Numbers 17H05821 (MH), 19H04833 (MH),17H05823 (SN), 19H04843 (SN), 16H06429 (TW), 16K21723 (TW), and 16H06434 (TW); the Cooperative Research Program (Joint Usage/Research Center) of the Institute for Frontier Life and Medial Sciences at Kyoto University (SN); and a Grant for contracted research activity related to crane conservation by the City of Izumi, Japan (MO). This research was commissioned by the Kagoshima Crane Conservation Committee.

Author contributions MO, TM, KT, YK, TW, and MH designed the study. YM, YT, MO, TW, and MH performed the researches. KK, SN, TI and MH analyzed the data. YM, SN, TW, and MH wrote the paper.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Research involving human participants and/or animals This article does not contain any studies with human participants or animals, performed by any of the authors.

Informed consent Informed consent concerns are not applicable.

References

1. Geoghegan JL, Holmes EC (2017) Predicting virus emergence amid evolutionary noise. Open Biol 7(10):170189. https://doi.org/10.1098/rsob.170189
2. Shi M, Zhang YZ, Holmes EC (2018) Meta-transcriptomics and the evolutionary biology of RNA viruses. Virus Res 243:83–90. https://doi.org/10.1016/j.virusres.2017.10.016
3. Bodewes R, Ruiz-Gonzalez A, Schapendonk CM, van den Brand JM, Osterhaus AD, Smits SL (2014) Viral metagenomic analysis of feces of wild small carnivores. Virol J 11:89. https://doi.org/10.1186/1743-422X-11-89
4. Honma H, Suyama Y, Nakai Y (2011) Detection of parasitizing coccidia and determination of host crane species, sex and genotype by faecal DNA analysis. Mol Ecol Resour 11(6):1033–1044. https://doi.org/10.1111/j.1755-0998.2011.03048.x
5. Okuya K, Kanazawa N, Kanda T, Kuwahara M, Matsuu A, Horie M, Masatani T, Toda S, Ozawa M (2017) Genetic characterization of an avian H4N6 influenza virus isolated from the Izumi plain, Japan. Microbiol Immunol. https://doi.org/10.1111/1348-0421.12545
6. Nakagawa H, Okuya K, Kawabata T, Matsuu A, Takase K, Kuwahara M, Toda S, Ozawa M (2018) Genetic characterization of low-pathogenic avian influenza viruses isolated on the Izumi plain in Japan: possible association of dynamic movements of wild birds with AIV evolution. Arch Virol 163(4):911–923. https://doi.org/10.1007/s00705-017-3698-1
7. Ozawa M, Matsuu A, Tokorozaki K, Horie M, Masatani T, Nakagawa H, Okuya K, Kawabata T, Toda S (2015) Genetic diversity of highly pathogenic H5N8 avian influenza viruses at a single overwintering site of migratory birds in Japan, 2014/15. Euro Surveill 20(20):21132
8. Ozawa M, Matsuu A, Khalil AM, Nishi N, Tokorozaki K, Masatani T, Horie M, Okuya K, Ueno K, Kuwahara M, Toda S (2019) Phylogenetic variations of highly pathogenic H5N6 avian influenza viruses isolated from wild birds in the Izumi plain, Japan, during the 2016–2017 winter season. Transbound Emerg Dis 66(2):797–806. https://doi.org/10.1111/tbed.13087
9. Ratnasingham S, Hebert PD (2007) The barcode of life data system. Mol Ecol Notes 7(3):355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x
10. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinform 10:421. https://doi.org/10.1186/1471-2105-10-421
11. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34(17):i884–i890. https://doi.org/10.1093/bioinformatics/bty560
12. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. Genome Res 27(5):824–834. https://doi.org/10.1101/gr.213959.116
13. Shen W, Le S, Li Y, Hu F (2016) SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS ONE 11(10):e0163962. https://doi.org/10.1371/journal.pone.0163962
14. Steinegger M, Soding J (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol 35(11):1026–1028. https://doi.org/10.1038/nbt.3988
15. 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(5):455–477. https://doi.org/10.1089/cmb.2012.0021
16. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923
17. Barnett DW, Garrison EK, Quinlan AR, Stromberg MP, Marth GT (2011) BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics 27(12):1691–1692. https://doi.org/10.1093/bioinformatics/btr174
18. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772–780. https://doi.org/10.1093/molbev/mst010
19. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096
20. Marek A, Nolte V, Schachner A, Berger E, Schlotterer C, Hess M (2012) Two fiber genes of nearly equal lengths are a common and distinctive feature of Fowl adenovirus C members. Vet Microbiol 156(3–4):411–417. https://doi.org/10.1016/j.vetmic.2011.11.003
21. Harrach B, Benkö M, Both G, Brown M, Davison A, Echavarría M, Hess M, Jones M, Kajon A, Lehmkuhl H (2011) Family adenoviridae. Virus taxonomy: classification and nomenclature of viruses. In: Ninth report of the international committee on taxonomy of viruses (2011). Elsevier, San Diego, CA, pp 125–141
22. Tan B, Yang XL, Ge XY, Peng C, Liu HZ, Zhang YZ, Zhang LB, Shi ZL (2017) Novel bat adenoviruses with low G+C content shed new light on the evolution of adenoviruses. J Gen Virol 98(4):739–748. https://doi.org/10.1099/jgv.0.000739
23. Ojkic D, Nagy E (2000) The complete nucleotide sequence of fowl adenovirus type 8. J Gen Virol 81(Pt 7):1833–1837. https://doi.org/10.1099/0022-1317-81-7-1833
24. Corredor JC, Garceac A, Krell PJ, Nagy E (2008) Sequence comparison of the right end of fowl adenovirus genomes. Virus Genes 36(2):331–344. https://doi.org/10.1007/s11262-007-0194-9
25. Griffin BD, Nagy E (2011) Coding potential and transcript analysis of fowl adenovirus 4: insight into upstream ORFs as common sequence features in adenoviral transcripts. J Gen Virol 92(Pt 6):1260–1272. https://doi.org/10.1099/vir.0.030064-0
26. Marek A, Kosiol C, Harrach B, Kajan GL, Schlotterer C, Hess M (2013) The first whole genome sequence of a Fowl adenovirus B strain enables interspecies comparisons within the genus *Aviadenovirus*. Vet Microbiol 166(1–2):250–256. https://doi.org/10.1016/j.vetmic.2013.05.017
27. Ojkic D, Nagy E (2001) The long repeat region is dispensable for fowl adenovirus replication in vitro. Virology 283(2):197–206. https://doi.org/10.1006/viro.2000.0890
28. Hagiwara S (1988) Wild plants for winter foods in *Grus monacha* & *G. vipio* in Izumi, Japan (studies of the Cranes in Izumi, Kagoshima, Japan. 13). Misc Rep Natl Park Nat Stud 19:83–97 **(in Japanese)**
29. Board of Education in Kagoshima Prefecture (1995) Research report on the long-term crane protective measures investigation studies, 1994 fiscal year **(in Japanese)**
30. Board of Education in Kagoshima Prefecture (1996) Research report on the long-term crane protective measures investigation studies, 1995 fiscal year **(in Japanese)**
31. Li PH, Zheng PP, Zhang TF, Wen GY, Shao HB, Luo QP (2017) Fowl adenovirus serotype 4: epidemiology, pathogenesis, diagnostic detection, and vaccine strategies. Poult Sci 96(8):2630–2640. https://doi.org/10.3382/ps/pex087

Affiliations

Yahiro Mukai[1,2] · Yuriko Tomita[3] · Kirill Kryukov[4] · So Nakagawa[4] · Makoto Ozawa[5,6,7] · Tsutomu Matsui[8] · Keizo Tomonaga[1,2,9] · Tadashi Imanishi[4] · Yoshihiro Kawaoka[10,11] · Tokiko Watanabe[3] · Masayuki Horie[1,12]

1 Laboratory of RNA Viruses, Department of Virus Research, Institute for Frontier Life and Medical Sciences, Kyoto, Japan

2 Department of Mammalian Regulatory Network, Graduate School of Biostudies, Kyoto University, Kyoto, Japan

3 Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo, Japan

4 Department of Molecular Life Science, Tokai University School of Medicine, Tokyo, Japan

5 Joint Faculty of Veterinary Medicine, Laboratory of Animal Hygiene, Kagoshima University, Kagoshima, Japan

6 Transboundary Animal Diseases Research Center, Joint Faculty of Veterinary Medicine, Kagoshima University, Kagoshima, Japan

7 United Graduate School of Veterinary Science, Yamaguchi University, Yamaguchi, Japan

8 Kagoshima Crane Conservation Committee, Izumi, Kagoshima, Japan

9 Department of Molecular Virology, Graduate School of Medicine, Kyoto University, Kyoto, Japan

10 Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, Madison, WI, USA

11 Department of Special Pathogens, International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo, Japan

12 Hakubi Center for Advanced Research, Kyoto University, Kyoto 606-8507, Japan