Pathogenicity, genomic analysis and structure of abalone asfa-like virus: evidence for classification in the family Asfaviridae

JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE

Matsuyama *et al.*, *Journal of General Virology* 2023;104:001875

DOI 10.1099/jgv.0.001875

MICROBIOLOGY SOCIETY

OPEN ACCESS

Pathogenicity, genomic analysis and structure of abalone asfa-like virus: evidence for classification in the family *Asfarviridae*

Tomomasa Matsuyama*, Ikunari Kiryu, Tohru Mekata†, Tomokazu Takano, Kousuke Umeda and Yuta Matsuura

Abstract

This paper presents the rationale for classifying abalone asfa-like virus (AbALV) in the family *Asfarviridae* based on analyses of the host, whole genome and electron microscopic observations. AbALV caused >80% cumulative mortality in an experimentally infected mollusc, *Haliotis madaka*. The AbALV genome was found to be linear, approximately 281 kb in length, with a G+C content of 31.32%. Of the 309 predicted ORFs, 48 of the top hits with African swine fever virus (ASFV) genes in homology analysis were found to be in the central region of the genome. Synteny in the central region of the genome was conserved with ASFV. Similar to ASFV, paralogous genes were present at both ends of the genome. The pairwise average amino acid identity (AAI) between the AbALV and ASFV genomes was 33.97%, within the range of intra-family AAI values for Nucleocytoviricota. Electron microscopy analysis of the gills revealed ~200 nm icosahedral virus particles in the cytoplasm of epithelial cells, and the size and morphology resembled ASFV. In addition to swine, ASFV also infects ticks, which are protostomes like abalone. The overall genome structure and virion morphology of AbALV and ASFV are similar, and both viruses infect protostomes, suggesting that AbALV is a new member of the family *Asfarviridae*.

INTRODUCTION

African swine fever (ASF) virus (ASFV) is a large, enveloped virus classified in the phylum Nucleocytoviricota and the only known DNA arbovirus [1, 2]. ASFV exhibits high infectivity and lethality in swine. ASF was first diagnosed in the Republic of Kenya in 1907 [3] and later spread from Africa to Europe, Asia and Oceanian countries [4]. According to a recent World Organization for Animal Health (WOAH) situation report, ASF cases have been confirmed in 46 countries since 2021, causing more than 1.1 million animal fatalities [5]. Currently, there is no effective vaccine, making ASF one of the most feared infectious diseases of livestock.

Mature ASFV virions are 175–215 nm in diameter and have a linear, dsDNA genome 170–190 kb in length, which varies among strains and contains multigene families in terminals [1]. Although several viruses with characteristics similar to ASFV have been reported, ASFV remains the only species in the family *Asfarviridae* in the ICTV classification [6]. Three virus species isolated using protozoan amoebae as hosts, including Faustovirus [7], Kaumoebavirus [8] and Pacmanvirus [9], harbour core genes homologous to those of ASFV. However, the genomes of these viruses are more than twice the size of the ASFV genome, and the taxonomic groups of hosts differ from those of ASFV. Unlike ASFV, which has a linear genome, Faustovirus [7] and Kaumoebavirus [8] have a circular genome. Due to these distinct features, there is some debate as to whether these three viruses should be classified in the family *Asfarviridae* or comprise putative new families, with some researchers classifying Faustovirus, Kaumoebavirus and Pacmanvirus as extended *Asfarviridae* [10]. In this paper, we describe these three viruses as members of an extended family *Asfarviridae*, although they are still listed as unclassified viruses in the ICTV taxonomy [6].

Received 08 June 2023; Accepted 18 July 2023; Published 02 August 2023

Author affiliations: [1]Pathology Division, Aquaculture Research Department, Fisheries Technology Institute, Japan Fisheries Research and Education Agency, Minami-Ise, Mie, Japan.

***Correspondence:** Tomomasa Matsuyama, matsuyama_tomomasa55@fra.go.jp

Keywords: abalone asfa-like virus; *Asfarviridae*; African swine fever virus; *Haliotis*; genome; electron microscopy.

Abbreviations: AAI, average amino acid identity; AbALV, abalone asfa-like virus; ASF, African swine fever; ASFV, African swine fever virus; MCP, major capsid protein; nr, non-redundant; Po1B, type B DNA polymerase; TEM, transmission electron microscopy; WOAH, World Organization for Animal Health.

†Present address: Department of Veterinary Medicine, Faculty of Veterinary Medicine, Okayama University of Science, Imabari, Ehime, Japan.

The complete genome sequence of AbALV was submitted to GenBank and assigned accession number LC637659.

Impact Statement

Genomic and morphological characteristics indicate that abalone asfa-like virus (AbALV) should be classified in the family *Asfarviridae*. Although African swine fever virus (ASFV) is currently considered the sole member of the family *Asfarviridae*, the family appears to be more diverse than previously thought. Characterization of AbALV will provide new perspectives for advancing the understanding of ASFV, which has a complex infection mechanism and an unclear evolutionary origin.

The sequence of the type B DNA polymerase (PolB) gene of the *Heterocapsa circularisquama* DNA virus, which infects marine dinoflagellates, is highly similar to the PolB gene sequence of ASFV [11]. Additionally, published genomic data of the gastropod *Elysia marginata* [12] include a scaffold containing multiple coding sequences homologous to the gene of the ASFV genome [13, 14]. Furthermore, metagenomic analyses have identified sequences homologous to the ASFV in marine [15, 16] and freshwater [17] sources. ASFV-like integrated elements have also been found in the genomes of ticks [18] and fungi [14], suggesting that the ancestors of these organisms were infected with ASFV-like viruses. ASFV and related viruses appear to be more diverse than previously thought.

We previously reported the approximately 155-kb partial genome sequence of a viral pathogen of abalone amyotrophia, which causes mass mortality in juvenile *Haliotis* spp. [19, 20]. The partial genome sequence was conserved in synteny with the genome of ASFV, and we proposed abalone asfa-like virus (AbALV) as a tentative name for the causative virus [19]. Phylogenetic analysis of its core gene sets indicated that AbALV is most closely related to ASFV among currently known viruses [13, 14, 16, 19]. Here, we discuss the rationale for including AbALV in the family *Asfarviridae* based on analyses of the host and AbALV whole genome sequence as well as electron microscopic observations of the virus.

METHODS

Animals

Details regarding the abalone used in this study are shown in Table 1. For infection experiments, healthy 12-month-old *Haliotis gigantea* and *Haliotis madaka* were used as recipients. The source of infection was a 5-year-old *Haliotis discus discus* that was still alive 165 days after initiation of the experimental infection conducted in our previous study [20]. For electron microscopic analysis, 6-month-old healthy *H. discus discus* were used as recipients, and *H. gigantea*, which survived and became 18 months old 182 days after the start of the above infection experiments, was used as the source of infection. Abalone were reared in 65 litre tanks filled with 56 lires of running (approximately 300 ml min^{-1}), UV-irradiated seawater maintained at 18–20 °C under the natural photoperiod in the laboratory. The abalone were fed dried kelp (*Saccharina japonica*) once each week during the experimental period. The infection status of abalone used as recipients and donors was determined prior to the start of the infection test. DNA was extracted from the muscles (including shell muscle and foot muscle) of eight randomly sampled abalone prior to the start of the study and tested for AbALV using quantitative PCR (qPCR), as described below.

Table 1. Details of the *Haliotis* abalone used in this study

					Body length (mm)	
Experiments	**Species scientific name**	**Starting disease status**	**Use in experiments**	**Age**	**Range**	**Mean**
Infection experiment						
	H. discus discus	Experimentally infected*	Infectious source	5 years	49–61	55.8±1.3
	H. gigantea	Healthy	Recipient	12 months	18–24	21.1±1.9
	H. madaka†	Healthy	Recipient	12 months	22–29	25.3±1.7
Electron microsopy						
	H. gigantea	Experimentally infected‡	Infectious source	18 months	31–35	32.9±0.6
	H. discus discus	Healthy	Recipient	6 months	7–11	9.7±0.2

*Survivors from a previously reported experimental infection study [20]. A total of 165 days had passed since infection.
†A DNA sample with the highest AbALV copy number per unit of extracted total DNA was used for AbALV genomic analysis.
‡Survivors from the above infection experiment. A total of 182 days had passed since infection.

qPCR

Based on the AbALV major capsid protein (MCP) gene sequence, primers and fluorescent probes were designed using primer3 (v. 0.4.0) [21, 22]. The sequences of the primers and probe were as follows: forward primer, CCCATCCAACACTCATTCTC; reverse primer, AGCCGAACATCTTCATTAAACC; fluorescent probe, 5′FAM-CGGCACTAAATGGTCCACAAACACCAA-3′BHQ. DNA was extracted from approximately 50 mg of tissue using an Agencourt DNAdvance kit (GE Healthcare) and eluted with 200 µl of nuclease-free water. The reaction mixture was prepared using a KAPA probe force qPCR kit (Roche) with 1 µl of the eluate as a template, according to the kit instructions. A LightCycler 96 (Roche) was used for PCR assays, with 40 cycles of 98 °C for 2 min followed by 95 °C for 10 s and 60 °C for 30 s. The copy number of the MCP gene sequence in the sample was calculated by creating a calibration curve using a dilution series of plasmid containing the MCP gene [23]. The DNA concentration in the eluate was calculated from the absorbance at 260 nm using Nivo (PerkinElmer), and the copy number of the MCP gene in 1 µg of genomic DNA was then calculated.

Infection experiment

Twelve-month-old *H. gigantea* and *H. madaka* were divided into three observation groups, each consisting of 50 abalone and a sampling group consisting of 500 abalone. Ten infected 5-year-old *H. gigantea* were reared upstream of the two observation groups and the sampling group, and the effluent of the rearing tank was supplied to the groups. Another observation group was directly supplied with UV-sterilized seawater as a negative control. Each week, ten abalone were taken from the sampling tanks, and the AbALV genome copy number in the muscle (including shell muscle and foot muscle) and gill was determined by qPCR. The sample with the highest copy number per extracted microgram of DNA was used for AbALV genome analysis. Five abalone were sampled for immunohistological analysis on days 14, 28 and 42 of infection. The whole abalone was fixed in Davidson's fixative [24], and paraffin sections were prepared using a standard method. Immunostaining using mouse antiserum prepared against the recombinant MCP of AbALV and determination of the number of positive cells per area of microscopic field of view were performed using methods described in our previous paper [20].

Genome sequencing and assembly

Both long-read sequencing (nanopore) and short-read sequencing (DNBSEQ sequencer) were performed using a sample with the highest virus copy number per unit of DNA obtained from the infection test described above. The libraries for nanopore sequencing were prepared using a rapid sequencing kit (Oxford Nanopore Technologies), and sequencing was carried out using an R9.4 flow cell and MinION instrument. Base calling was performed using MinKNOW v19.12.5 software in FAST mode, and the draft genome was assembled using the CLC Genomics Workbench v21.0.3 *de novo* assembly software (Qiagen). For short-read sequencing, libraries were prepared using an MGI Easy FS DNA library prep set (MGI Tech), and 150 bp paired-end sequencing using the DNBSEQ platform from MGI Tech was outsourced to the Beijing Genomics Institute. The final full-length genome sequence of AbALV was constructed by short-read polishing of the draft long-read genome assembly using CLC Genomics Workbench v21.0.3.

To determine whether the AbALV genome is circular or linear, two primer sets (Table 2) were designed from both ends of the full-length sequence outward. For the positive control for PCR assays, two primer sets were designed to amplify two locations of the genome. Using 1 µl of the DNA solution used for the genome analysis as a template, PCR was performed in

Table 2. Primer sets used for genome structure estimation

Use*	Primer†	Sequence	Predicted product size (bp)
inv	928 R	CCGATCAATATACTGACTTACCACCTCTAACAAACGG	–
	281249 F	GAAGATTATTTTGAAATAGCAATTAGAAGTAAACCGG	
inv	1316 R	CCAATTTACCTATACCGATTTACCACCTTTAACAAATGG	–
	281187 F	CCGCTACATCCCGACCTTTCCAACATTG	
posi	22084 F	GCAATTGTGTTACACACATTGTAAGTTTCAAATTGG	14225
	36309 R	CATCCGGAGCATTTACACAAACTTGTGTAATCAAAGG	
posi	221277 F	GGTTTACCAGATGTCGGCACTATAAAAGAAGTCC	10770
	232047 R	GGCCATGAAATTTTGTTAATATTTTCGGGTAACACATCACGG	

*inv: inverse PCR to infer genome structure; posi: positive control for PCR.
†Number in the primer name indicates the genomic position of the 5′-end of the primer sequence.
F and R indicate forward and reverse primer, respectively.

a 50 µl reaction using a KOD-Fx-neo DNA polymerase kit (TOYOBO), and the desired products were confirmed by agarose gel electrophoresis. The PCR mixture was prepared according to the manufacturer's instructions. Reaction conditions were as follows: 98 °C for 30 s, followed by 98 °C for 10 s and 72 °C for 10 min (three cycles), 98 °C for 10 s and 70 °C for 10 min (three cycles), 98 °C for 10 s and 68 °C for 10 min (three cycles), 98 °C for 10 s, 65 °C for 30 s and 68 °C for 10 min (three cycles), 98 °C for 10 s, 60 °C for 30 s and 68 °C for 10 min (28 cycles), and finally holding at 68 °C for 10 min.

Genome annotation

ORFs were predicted using GeneMarkS (v4.32) [25]. The predicted ORFs were subjected to homology searches against the non-redundant (nr) and virus databases (5 May 2023) using BLASTp. Homology was considered significant if the E-value was $<1\times10^{-3}$. Paralogous genes were identified through an all-against-all protein sequence comparison of AbALV ORFs in BLASTp with identified ORF combinations in which one pair had coverage >60% and E-value $<1\times10^{-6}$ as paralogues. A genome map was generated using drawGeneArrows3 (http://www.ige.tohoku.ac.jp/joho/labhome/tool.html). The results of the homology analysis and paralogue search are colour-coded on the genome map. tRNA genes were identified using tRNAscan-SE (2.0) [26, 27] and ARAGORN software [28]. The pairwise average amino acid identity (AAI) between the genomes of AbALV and related viruses was calculated using the AAI calculator available online from the Kostas lab (http://enve-omics.ce.gatech.edu/aai/) [29].

Electron microscopy

Gills were used for transmission electron microscopy (TEM) observation because immunopositive cells were observed at the highest frequency in this organ. An experimental infection was performed to obtain gill tissue with a high density of AbALV-infected cells. The effluent from a tank in which 40 infected *H. gigantea* were kept was supplied to 50 healthy, 6-month-old *H. discus discus*. Four weeks after the supply of effluent, ten *H. discus discus* were killed, and the gills were collected and divided into two parts. One tissue piece was fixed in Davidson's fixative [24] and immunostained to identify samples with the highest frequency of AbALV-infected cells, as described above. The other tissue piece was fixed for electron microscopy in Karnovsky fixative [30] and then stored at 4 °C in 0.1 M phosphate buffer until use. For electron microscopy, a sample from an abalone with the highest density of AbALV-infected cells was post-fixed in 1% osmium tetroxide, embedded in Epon 812 (TAAB) and used for TEM observation. Ultra-thin section preparation and electron microscopy were contracted to Tokai Electron Microscopy.

Statistical analyses

Statistical analyses were performed using the Statcel3 software package (OMS). The frequency of immunopositive cells among three tissues were compared using Kruskal–Wallis tests with Bonferroni correction. Cumulative mortality was analysed using Fisher's exact test. Differences in virus load between muscle and gill were analysed using Student's *t*-test. A *P*-value of <0.05 was considered significant.

RESULTS AND DISCUSSION

Infection experiment

H. madaka was more susceptible to AbALV infection than *H. gigantea*, as evidenced by mortality and viral replication. In both species, mortality was observed only in the infected group; the cumulative mortality in the duplicated infected groups of *H. madaka* was 92 and 80%, and that of *H. gigantea* was 18 and 10%, respectively (Fig. 1a, b). Virus levels were higher in *H. madaka* than in *H. gigantea* throughout the measurement period (Fig. 1c, d). In *H. madaka*, virus levels peaked in the gills after 35 days of infection and in the muscle after 56 days of infection. The trend of increased virus levels in muscle secondary to the gill was reported in a study that analysed AbALV-infected cells of 7-month-old *H. gigantea* by immunostaining [20]. The reason for the change in infected tissue is unknown. In *H. gigantea*, which has a low mortality rate, the amount of virus detected was also low, and in contrast to *H. madaka*, the amount of virus was higher in muscle than gill throughout the study period. Immunostaining also showed a higher number of positive cells in *H. madaka* than in *H. gigantea*. Positive cells were observed in gill epithelial cells, and they were also seen in muscle and hypobranchial glands (Figs 1e, f and 2). A few positive cells were also found in the kidneys (Fig. 2). There were no statistically significant differences in the frequency of positive cells in *H. gigantea* among the muscles, gills and hypobranchial glands. Due to the lack of information about abalone cells, target cells could not be identified from the immunostaining images. ASFV replicates in cell types of the mononuclear–phagocytic system, including fixed-tissue macrophages [6]. Abalone haemocytes, like mammalian macrophages, have phagocytosis, reactive oxygen species-producing capacity and intracellular degradation of engulfed foreign material [31]. However, based on the form and distribution of cells that exhibited positive immunostaining, AbALV's target cells do not appear to be haemocytes. Identification of the target cells of the virus will be important in elucidating the gateway of entry of the virus and in understanding how it causes death and other pathological conditions.

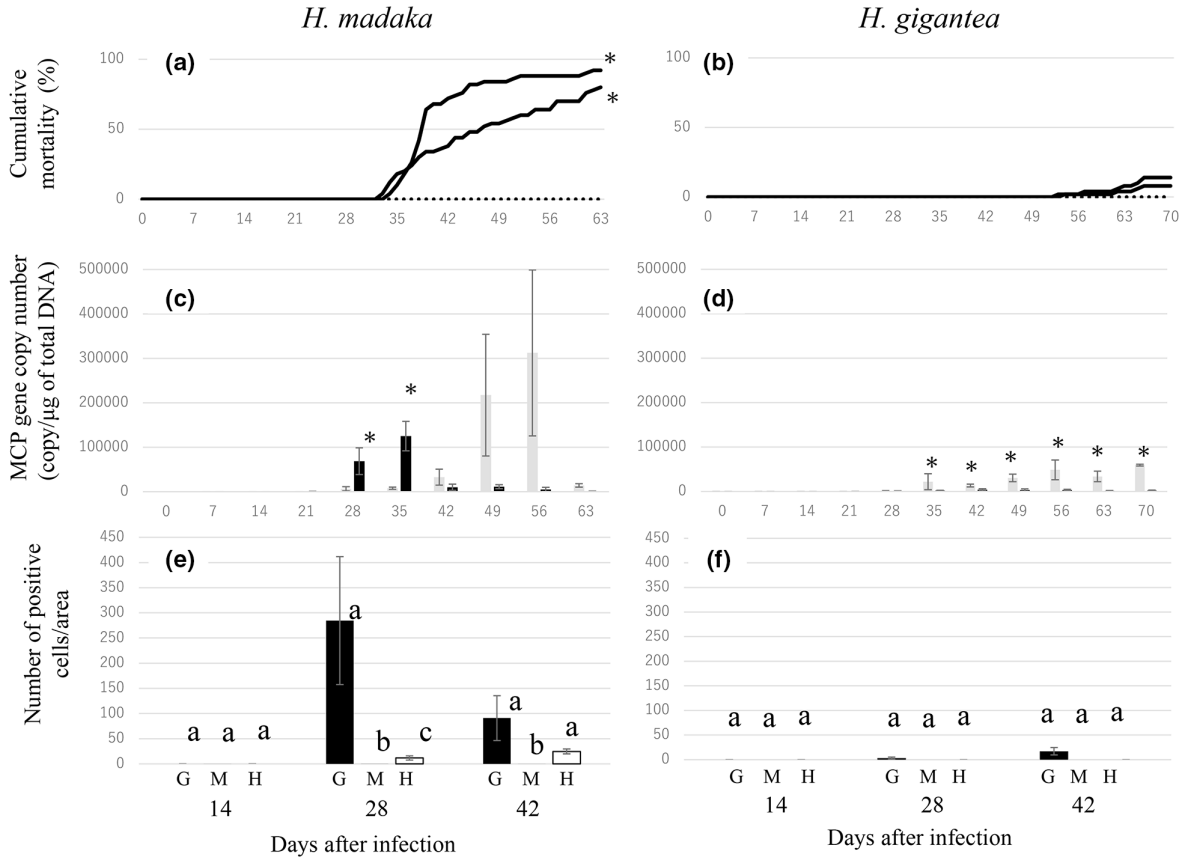

Fig. 1. Kinetics of cumulative mortality, copy number of the AbALV MCP gene and number of immunopositive cells in experimentally infected 12-month-old *H. madaka* and *H. gigantea*. (a, b) Bold and dashed lines indicate cumulative mortality in the challenge and control groups, respectively. The two bold lines show results of duplicate tests. Asterisks indicate significant differences compared to the control group. (c, d) Copy number of the AbALV MCP gene in muscles (grey bars) and gills (black bars). Results are expressed as copy number per microgram of total DNA. Asterisks indicate significant differences between the two tissues. (e, f) Number of MCP-immunopositive cells in gills (G, black bars), muscle (M, grey bars) and hypobranchial glands (H, white bars). Different letters indicate significant differences among tissues. Bars within a group with different letters indicate significant differences among the observed tissues. (a, c, e) *H. madaka*, (b, d, f) *H. gigantea*.

Although ASFV causes high mortality regardless of pig age [32, 33], susceptibility to AbALV appears to decrease as abalone grow. Although the 12-month-old *H. gigantea* used in this study had low mortality, in a previous study, 7-month-old *H. gigantea* showed significant AbALV proliferation, with a cumulative mortality rate approaching 50% [20]. These data suggest that *H. gigantea* becomes less susceptible to AbALV between 7 and 12 months of age. A previous study showed that three species of abalone (*H. gigantea*, *H. discus hannai* and *H. discus discus*) are susceptible to AbALV, but even in these species, adult shellfish have low mortality rates when infected [20]. Co-habitation of juvenile abalone with apparently healthy adult abalone collected in the field has been reported to cause the disease in juveniles [34, 35], indicating that wild adult abalones are carriers of AbALV. ASFV is not pathogenic to its natural hosts, warthogs (*Phacochoerus africanus*) [36], bush pigs (*Potamochoerus* spp.) [37] and ticks (*Ornithodoros moubata*) [38], and these animals serve as reservoirs. Although AbALV infection has not been studied in animals other than abalone, the virus may be maintained in nature with adult abalone as carriers.

Overall structure of the AbALV genome

Both long- and short-read sequencing were performed for a DNA sample extracted from the muscle of an *H. madaka* on day 56 of infection in the infection experiment described above. In this analysis, 57.3 kb of the left terminal and 68.8 kb of the right terminal portion of the genome were newly sequenced. A comparison of the overall genome structure between AbALV and related virus species is shown in Table 3. The size of the AbALV genome was 281 224 bp, with 309 predicted ORFs. The genome size and the number of ORFs of AbALV were between those of ASFV (170–194 kb, 150–167 ORFs [39]) and extended *Asfarviridae* (351–466 kb, 457–465 ORFs [8, 9, 40]). AbALV had the lowest GC content, at 31.32%. The genome was considered linear, as PCR with two primer sets designed outward at both ends of the genome yielded no amplified products (data not shown). Unlike Pacmanvirus [9], the AbALV genome did not encode tRNA.

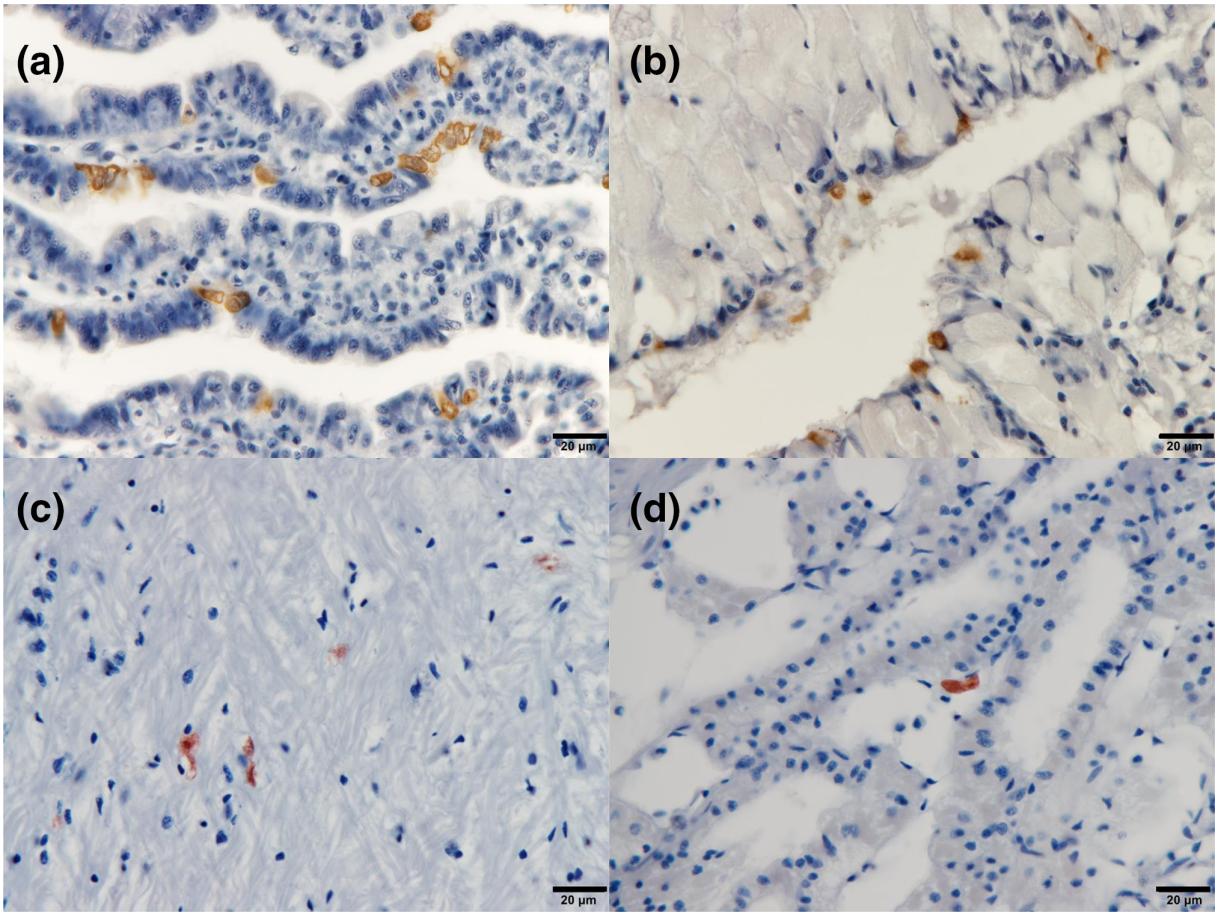

Fig. 2. Immunostaining images of *H. madaka* experimentally infected with AbALV. (a) Gills, (b) hypobranchial glands, (c) muscle, (d) kidney. Bars, 20 µm.

The pairwise AAI between the genomes of AbALV and ASFV strain Georgia 2007/1 was 33.97%, within the range of intra-family AAI values (26–100%) in the Nucleocytoviricota [41] (Table 4). The three extended *Asfarviridae* viruses also had AAI values with ASFV within the range of the above intra-family AAI, but AbALV had the highest value.

Organization of the AbALV genome

The overall structure of the AbALV genome was determined based on BLASTp best hits against the virus database and a search for paralogous genes (Fig. 3). The AbALV genome shared some features with the ASFV genome. The central region of the genome encodes 48 ORFs that were top hits with ASFV genes, 32 of which showed homology with ASFV core genes [42]. As in the genome

Table 3. Comparison of the characteristics of AbALV and related viruses

			Extended *Asfarviridae*		
	AbALV	**ASFV**	**Faustovirus**	**Kaumoebavirus**	**Pacmanvirus**
Genome length (kbp)	281	170–194	456–491	351–363	395–419
Proteins	309	150–167	467–520	465–507	445–465
GC content (%)	31.32	appox. 39%	36.22–39.59	43.06–43.70	33.20–33.62
Genome structure	Linear	Linear	Circular	Circular	Linear
No. of tRNAs	0	0	0	0	1
Virion size (nm)	200	175–215	200	250	175

Table 4. Pairwise average amino acid identity (AAI, %) distances among the genomes of AbALV and related viruses

	AbALV	ASFV Georgia 2007	Faustovirus E12	Kaumoebavirus KLCC10	Pacmanvirus A23
AbALV	100.00				
ASFV Georgia 2007	33.97	100.00			
Faustovirus E12	31.86	29.71	100.00		
Kaumoebavirus KLCC10	32.12	31.05	31.64	100.00	
Pacmanvirus A23	30.43	31.45	34.42	31.61	100.00

of ASFV [39], paralogous genes (referred to as multigene family types 1–7) were distributed in the newly sequenced regions (i.e. the left terminal 68.5 kb and right terminal 79 kb regions). Paralogues of the same types are often located adjacent to each other in the same direction, suggesting that they are multigene families. Type 1, which had the highest copy number, and types 2 and 3, which had the next highest copy numbers, were distributed on both ends, whereas the other four types (types 4–7) were distributed only on one side. None of the paralogous genes exhibited homology to any known virus sequence.

All of the ORFs that showed homology to ASFV and extended *Asfarviridae* genes were contained in a partial genome sequence of approximately 155 kb (corresponding to positions 57283–212447 of the full-length sequence) reported in a previous paper [19]. Previous studies have reported the partial genome sequence and results of synteny analyses with the ASFV genome [13, 19], phylogenetic analyses using multiple genes [13, 16, 19, 43] and comparisons of virion protein genes with related viruses [13]. Because no differences were observed between the full-length and the partial genome in these analyses in the present study, the results are not presented in this paper.

Twenty-one ORFs showed homology to viruses other than ASFV: our to Faustovirus, two to Pacmanvirus, seven to other Nucleocytoviricota and eight to viruses other than Nucleocytoviricota. Six of the eight ORFs that showed homology to viruses other than Nucleocytoviricota were top hits to bacteriophage genes. Namely, the tail fibre protein of *Podoviridae* and Caudovirales (ORFs

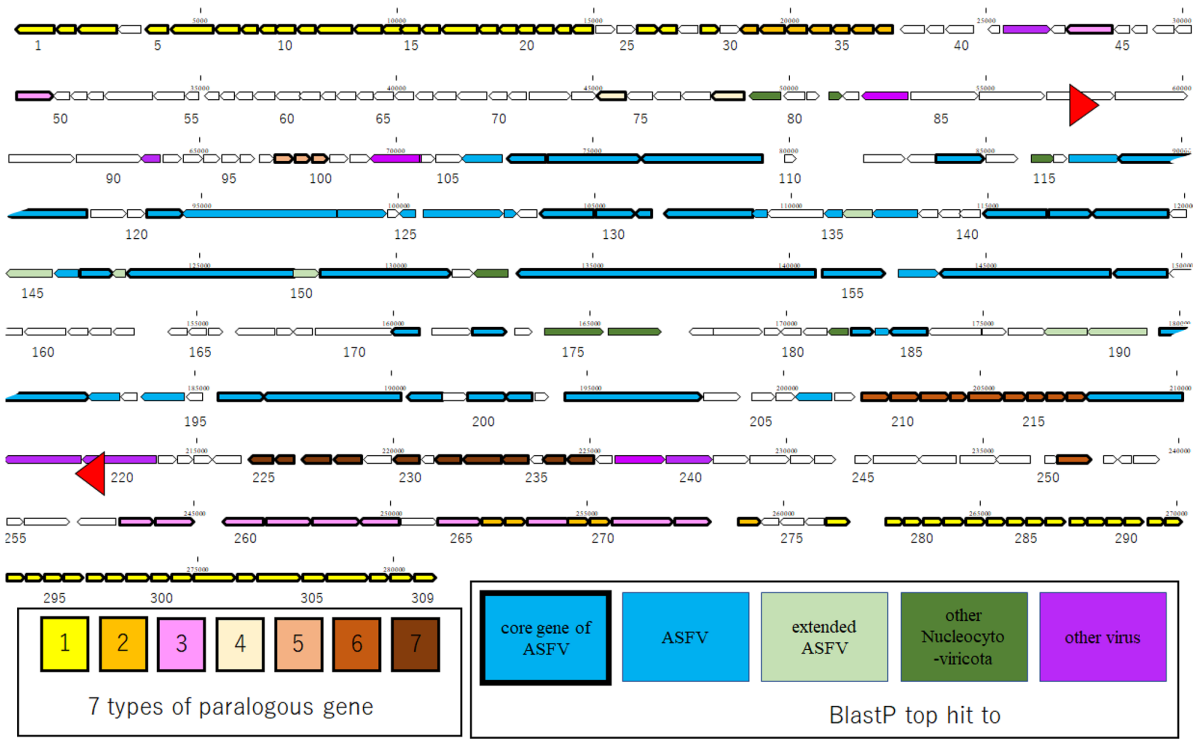

Fig. 3. Structure of the AbALV genome. Box arrows indicate ORFs, with the ORF number listed below. The area between the red arrowheads is the region reported in our previous paper (accession no: LC506465 [19]). Seven types of paralogous genes are indicated by bold arrows, as explained in the box at the lower left. Results of BLASTp searching are colour-coded, as explained in the box at the lower right.

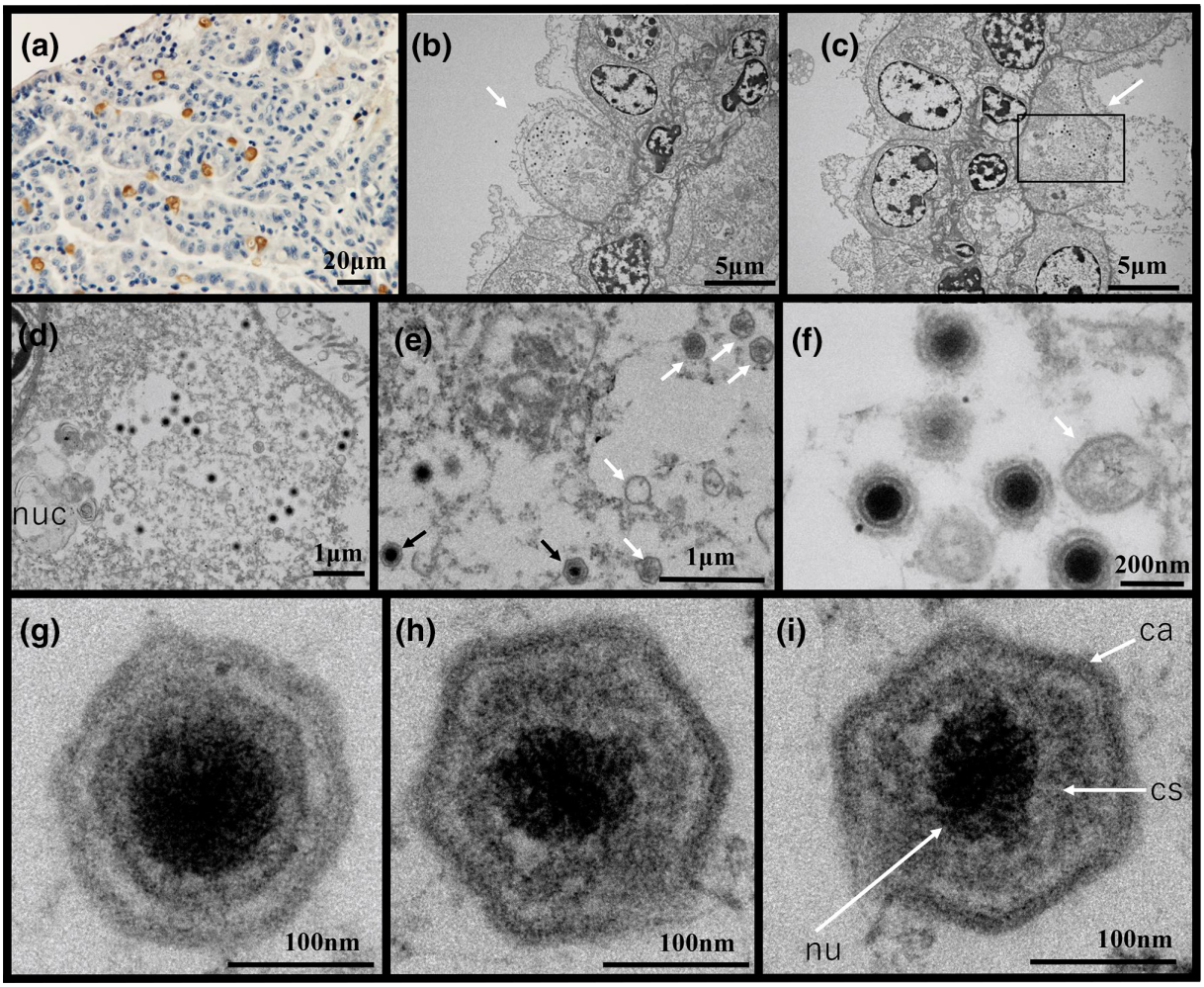

Fig. 4. Immunostaining and TEM observation. (a) Gill of an abalone analysed by TEM showing many positive cells (brown) by immunostaining with AbALV-MCP mouse antiserum. (b, c) Outline of infected cells (white arrows). (d) Enlarged view of the frame in c (nuc: cell nucleus). (e) White and black arrows indicate immature and mature virions, respectively. (f) White arrow indicates a particle that appears to be in the process of assembly. (g–i) Enlarged view of mature virions (ca: capsid, cs: core-shell, nu: nucleoid).

42, 219, 220), ATP-dependent DNA helicase (ORF 239), the DNA ligase of *Siphoviridae* (ORF 240), and the NADAR domain of Caudovirales-containing protein (ORF 091). The other two ORFs exhibited homology with vinnexin-1 in *Polydnaviridae* (ORF 084) and inhibition of apoptosis protein IAP-3 homologue in *Baculoviridae* (ORF 103).

Based on the BLASTp best hits in a search against the nr database, 34 ORFs showed the best hit to non-viral sequences. Of these, there were 18 top hits to sea slug (*E. marginata*), eight to bacteria, two to bivalves, two to protozoa, two to fungi, one to insects and one to plants. Twenty-eight of the 34 ORFs showed homology also to viral sequences, while the remaining six ORFs exhibited no homology to any known virus. Of the six ORFs that did not show hits to viruses, four (ORFs 105, 172, 193 and 195) showed top hits to *E. marginata*, and two (ORFs 271 and 272) showed top hits to bacterial genes. In the *E. marginata* genomic analysis, a considerable number of coding sequences homologous to ASFV and AbALV genes were found in one scaffold (accession no. BMAT01005152.1) that formed independently of the host genome scaffolds [13, 14]. It is likely that an unknown virus closely related to AbALV and ASFV had infected or was present in the *E. marginata* used for the genome analysis. Like abalone, *E. marginata* is a gastropod. The presence of viruses related to *Asfarviridae* in these two gastropod species is interesting, because it suggests that gastropods may be hosts for a variety of viruses in the family *Asfarviridae*.

Electron microscopy

Gill tissue of an abalone with the highest number of infected cells as determined by immunostaining (Fig. 4a) was observed by TEM. Viral particles morphologically resembling ASFV [6] were observed (Fig. 4b–i). All of the cells in which viral particles were observed showed a dome-like shape with wedge-shaped gaps between adjacent cells (Fig. 4b, d). This morphology was

also observed in cells showing positive immunostaining (Figs 2a and 4a). All virions observed were located in the cytoplasm (Fig. 4b–d), as is characteristic of ASFV [6]. Nuclei of infected cells exhibited reduced electron density and were located in the corners of the cells (Fig. 4c). Immature particles with low electron density in the centre of the virion (Fig. 4e) and virions in the process of assembly, which were not closed, were also observed (Fig. 4f). Nucleoid, core-shell and capsid structures were also observed in virions (Fig. 4g–i). In contrast to the extended *Asfarviridae*, in which the virus factory is filled with virions [7–9], no clear viral factory was observed in AbALV, and the particles were scattered in the cytoplasm. No budding images, extracellular virions, or particles within endosomes and lysosomes were observed. The viral particles presented an icosahedral form approximately 200 nm in size (Fig. 4g–i), similar to ASFV (175–215 nm [6]) (Table 3).

CONCLUSION

Our study supports the proposal that AbALV belongs to the family *Asfarviridae* due to the many similar characteristics of these viruses. For example, the pairwise AAI with an ASFV strain genome was 33.97%, within the range of intra-family AAI values for Nucleocytoviricota. Phylogenetic analyses using multiple genes have shown that AbALV and ASFV are the most closely related among the known viruses [13, 16, 19, 43]. The central region of the AbALV genome contains a considerable number of genes homologous to those in the ASFV genome, with shared synteny [13, 19]. As with ASFV [39], both ends of the AbALV genome contain paralogous gene repeats. The morphology and size of the viral particles also resemble those of ASFV. Furthermore, both AbALV and ASFV infect protostome hosts. Abalone asfa-like virus is inappropriate as a species name because the virus is now named using the binomial nomenclature [44]. We will propose a classification of this virus to the ICTV. Further understanding of AbALV will not only aid efforts to counter the threat to abalone aquaculture but also enhance understanding of ASFV.

Funding information
This work was supported by JSPS KAKENHI Grant Numbers JP 18H02282 and 22H02440.

Acknowledgements
We thank Kiyoshi Isowa of Mie Prefecture Sea Farming Center for providing healthy abalone.

Author contributions
T.M., T.M., T.K. and K.M. analysed sequence data; I.K. carried out the histological study; T.M., T.K. and Y.M. conducted experimental infection study and sampling; T.M. wrote the manuscript. All authors assisted in the research, analysis and formation of the manuscript and agreed to the final version.

Conflicts of interest
The authors declare no conflicts of interest.

Ethical statement
Abalone handling, husbandry and sampling were conducted based on the policy of the Institutional Animal Care and Use Committee of the Fisheries Research and Education Agency under approval of the committee (IACUC-NRIA nos. 20005 and 22004).

References

1. Tulman E, Delhon GA, Ku B, Rock DL. African swine fever virus. In: *Lesser Known Large dsDNA Viruses*. 2009. pp. 43–87.
2. Gaudreault NN, Madden DW, Wilson WC, Trujillo JD, Richt JA. African swine fever virus: an emerging DNA Arbovirus. *Front Vet Sci* 2020;7:215.
3. Eustace Montgomery R. On a form of swine fever occurring in British East Africa (Kenya Colony). *J Comp Pathol Ther* 1921;34:159–191.
4. Penrith ML. Current status of African swine fever. *CABI Agric Biosci* 2020;1:11.
5. WOAH. World Organisation for Animal Health, African Swine Fever (ASF)—Situation Report 30; 2023 [accessed 30 March 2023].
6. Alonso C, Borca M, Dixon L, Revilla Y, Rodriguez F, *et al*. ICTV virus taxonomy profile: *Asfarviridae*. *J Gen Virol* 2018;99:613–614.
7. Reteno DG, Benamar S, Khalil JB, Andreani J, Armstrong N, *et al*. Faustovirus, an asfarvirus-related new lineage of giant viruses infecting amoebae. *J Virol* 2015;89:6585–6594.
8. Bajrai LH, Benamar S, Azhar EI, Robert C, Levasseur A, *et al*. Kaumoebavirus, a new virus that clusters with *Faustoviruses* and *Asfarviridae*. *Viruses* 2016;8:278.
9. Andreani J, Khalil JYB, Sevvana M, Benamar S, Di Pinto F, *et al*. Pacmanvirus, a new giant icosahedral virus at the crossroads between *Asfarviridae* and *Faustoviruses*. *J Virol* 2017;91:e00212–00217.
10. Geballa-Koukoulas K, Andreani J, La Scola B, Blanc G. The Kaumoebavirus LCC10 genome reveals a unique gene strand bias among "Extended Asfarviridae" *Viruses* 2021;13:148.
11. Ogata H, Toyoda K, Tomaru Y, Nakayama N, Shirai Y, *et al*. Remarkable sequence similarity between the dinoflagellate-infecting marine girus and the terrestrial pathogen African swine fever virus. *Virol J* 2009;6:1–8.
12. Maeda T, Takahashi S, Yoshida T, Shimamura S, Takaki Y, *et al*. Chloroplast acquisition without the gene transfer in kleptoplastic sea slugs, *Plakobranchus ocellatus*. *Elife* 2021;10:e60176.
13. Hannat S, La Scola B, Andreani J, Aherfi S. Asfarviruses and closely related giant viruses. *Viruses* 2023;15:1015.
14. Zhao H, Hikida H, Okazaki Y, Ogata H. A 1.5 Mb continuous endogenous viral region in the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Microbiology* 2017;2023:2023. DOI: 10.1101/2023.04.17.537115.
15. Monier A, Claverie J-M, Ogata H. Taxonomic distribution of large DNA viruses in the sea. *Genome Biol* 2008;9:1–15.
16. Karki S, Moniruzzaman M, Aylward FO. Comparative genomics and environmental distribution of large dsDNA viruses in the family *Asfarviridae Front Microbiol* 2021;12:657471.
17. Wan X-F, Barnett JL, Cunningham F, Chen S, Yang G, *et al*. Detection of African swine fever virus-like sequences in ponds in the Mississippi Delta through metagenomic sequencing. *Virus Genes* 2013;46:441–446.

18. Forth JH, Forth LF, Lycett S, Bell-Sakyi L, Keil GM, *et al*. Identification of African swine fever virus-like elements in the soft tick genome provides insights into the virus' evolution. *BMC Biol* 2020;18:1–18.
19. Matsuyama T, Takano T, Nishiki I, Fujiwara A, Kiryu I, *et al*. A novel Asfarvirus-like virus identified as a potential cause of mass mortality of abalone. *Sci Rep* 2020;10:1–12.
20. Matsuyama T, Kiryu I, Inada M, Takano T, Matsuura Y, *et al*. Susceptibility of four abalone species, *Haliotis gigantea*, *Haliotis discus discus*, *Haliotis discus hannai* and *Haliotis diversicolor*, to abalone asfa-like virus. *Viruses* 2021;13:2315.
21. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, *et al*. Primer3--new capabilities and interfaces. *Nucleic Acids Res* 2012;40:e115– .
22. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 2007;23:1289–1291.
23. Matsuyama T, Kiryu I, Inada M, Nakayasu C. Verification of sample collection sites for PCR-based detection assay for abalone asfa-like virus. *FISH Pathol* 2021;56:18–21.
24. Shaw BL, Battle HI. The gross and microscopic anatomy of the digestive tract of the oyster *Crassostrea virginica* (Gmelin). *Can J Zool* 1957;35:325–347.
25. Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 2001;29:2607–2618.
26. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;25:955–964.
27. Lowe TM, Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 2016;44:W54–7.
28. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004;32:11–16.
29. Rodriguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. *Microbe Magazine* 2014;9:111–118.
30. Graham RC, Karnovsky MJ. The histochemical demonstration of monoamine oxidase activity by coupled peroxidatic oxidation. *J Histochem Cytochem* 1965;13:604–605.
31. Donaghy L, Hong H-K, Lambert C, Park H-S, Shim WJ, *et al*. First characterisation of the populations and immune-related activities of hemocytes from two edible gastropod species, the disk abalone, Haliotis discus discus and the spiny top shell, Turbo cornutus. *Fish Shellfish Immunol* 2010;28:87–97.
32. Blome S, Gabriel C, Dietze K, Breithaupt A, Beer M. High virulence of African swine fever virus caucasus isolate in European wild boars of all ages. *Emerg Infect Dis* 2012;18:708.
33. Penrith M-L, Vosloo W, Jori F, Bastos ADS. African swine fever virus eradication in Africa. *Virus Res* 2013;173:228–246.
34. Nakatsugawa T, Okabe M, Muroga K. Horizontal transmission of amyotrophia in Japanese black abalone. *Fish Pathol* 2000;35:11–14.
35. Okada K, Nishimura M, Kawamura T. Prevention of amyotrophia in mass production of 0-year-old abalone, *Haliotis discus*, by quarantine. *Aquac Sci* 2000;48:657–663.
36. Thomson GR, Gainaru MD, Van Dellen AF. Experimental infection of warthos (*Phacochoerus aethiopicus*) with African swine fever virus. *Onderstepoort J Vet Res* 1980;47:19–22.
37. Oura CA, Powell PP, Parkhouse RM. Detection of African swine fever virus in infected pig tissues by immunocytochemistry and in sity hybridisation. *J Virol Methods* 1998;72:205–217.
38. Kleiboeker SB, Scoles GA. Pathogenesis of African swine fever virus in Ornithodoros ticks. *Anim Health Res Rev* 2001;2:121–128.
39. Dixon LK, Chapman DAG, Netherton CL, Upton C. African swine fever virus replication and genomics. *Virus Res* 2013;173:3–14.
40. Benamar S, Reteno DGI, Bandaly V, Labas N, Raoult D, *et al*. Faustoviruses: comparative genomics of new Megavirales family members. *Front Microbiol* 2016;7:3.
41. Moniruzzaman M, Martinez-Gutierrez CA, Weinheimer AR, Aylward FO. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nat Commun* 2020;11:1710.
42. Wang L, Luo Y, Zhao Y, Gao GF, Bi Y, *et al*. Comparative genomic analysis reveals an 'open' pan-genome of African swine fever virus. *Transbound Emerg Dis* 2020;67:1553–1562.
43. Geballa-Koukoulas K, Abdi S, La Scola B, Blanc G, Andreani J. Pacmanvirus S19, the second pacmanvirus isolated from sewage waters in Oran, Algeria. *Microbiol Resour Announc* 2021;10:e00693-00621.
44. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Adriaenssens EM, *et al*. Recent changes to virus taxonomy ratified by the International Committee on Taxonomy of Viruses (2022). *Arch Virol* 2022;167:2429–2440.