Virology 452-453 (2014) 95–116

Contents lists available at ScienceDirect

Virology

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

The complete genome sequence of the *Alphaentomopoxvirus Anomala cuprea entomopoxvirus*, including its terminal hairpin loop sequences, suggests a potentially unique mode of apoptosis inhibition and mode of DNA replication

Wataru Mitsuhashi *, Kazuhisa Miyamoto, Sanae Wada

National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8634, Japan

ARTICLE INFO

Article history:
Received 1 October 2013
Returned to author for revisions
28 October 2013
Accepted 26 December 2013
Available online 31 January 2014

Keywords:
Entomopoxvirus
Alphaentomopoxvirus
Genome
Terminal hairpin loop
Apoptosis inhibition
DNA replication

ABSTRACT

Complete genome sequence of *Anomala cuprea entomopoxvirus*, which belongs to the genus *Alphaentomopoxvirus*, including its terminal hairpin loop sequences, is reported. This is the first genome sequence of *Alphaentomopoxvirus* reported, and hairpin loops in entomopoxviruses have not previously been sequenced. The genome is 245,717 bp, which is smaller than had previously been estimated for *Alphaentomopoxvirus*. The inverted terminal repeats are quite long, and experimental results suggest that one genome molecule has one type of hairpin at one end and another type at the other end. The genome contains unexpected ORFs, e.g., that for the ubiquitin-conjugating enzyme E2 of eukaryotes. The BIR and RING domains found in a single ORF for an inhibitor of apoptosis in baculoviruses and entomopoxviruses occurred in two different, widely separated ORFs. Furthermore, an ORF in the genome contains a serpin domain that was previously found in vertebrate poxviruses for apoptosis inhibition but not in insect viruses.

Introduction

The family *Poxviridae* contains two subfamilies: the *Chordopoxvirinae*, whose members infect vertebrates, and the *Entomopoxvirinae*, which infect insects. The *Entomopoxvirinae* is subdivided into the following three genera: *Alphaentomopoxvirus* of Coleoptera, *Betaentomopoxvirus* of Lepidoptera and Orthoptera, and *Gammaentomopoxvirus* of Diptera. The virions of entomopoxviruses (EVs (or EPVs)) are ovoid or brick-shaped (220–265 nm × 270–470 nm) and contain a large, linear, double-stranded DNA of 225–380 kbp (Becker and Moyer, 2007; Skinner et al., 2011) and thus their genomes are generally larger than those of *Chordopoxviruses* (ChPVs) (Becker and Moyer, 2007; Skinner et al., 2011).

To date, most of the viruses of the poxviridae, whose genomes have been fully sequenced are members of *Chordopoxvirinae*; the genome of at least one member of each genus of ChPVs has been sequenced (Hautaniemi et al., 2010; Perera et al., 2010). In contrast, complete genomic sequences have been obtained for only a few EVs, an orthopteran EV (*Melanoplus sanguinipes* EV(MSEV)), whose genus has been unassigned and likely should be established as a new genus, and five lepidopteran EVs (*Amsacta moorei* EV (AMEV), *Adoxophyes honmai* EV, *Choristoneura biennis* EV (CBEV), *Choristoneura rosaceana* EV (CREV), and *Mythimna separata* EV (MySEV)), whose genus is *Betaentomopoxvirus* (Afonso et al., 1999; Bawden et al., 2000; Thézé et al., 2013). Therefore genomic sequences of members of *Alphaentomopoxvirus* and *Gammaentomopoxvirus* had not been determined previously. Several features have hampered efforts to obtain complete genomic sequences of EVs. First, long stretches of AT-rich sequences in EVs have been problematic (Perera et al., 2010). Second, it is not easy to mass-rear a sufficient number of some scarab species in order to produce enough virions and to extract sufficiently large amounts of genome DNA from *Alphaentomopoxvirus* species because of difficulty in dissolving their inclusion bodies, namely, spheroids (Langridge and Roberts, 1977; Mitsuhashi et al., 1997).

Furthermore, little has been reported regarding the sequences external to the perfectly base-paired region of poxviruses, that is, the hairpin loop region sequences, because of the difficulties associated with the cloning or PCR amplification of such regions (Baroudy et al., 1982). Because hairpin loop regions are important during the process of DNA replication in the ChPV, vaccinia virus (VACV) (Smith, 2007; Moss, 2001), the elucidation of the sequences and structures of these regions should lead to understanding of the process of DNA replication in EVs and improve our understanding of this process in ChPVs.

New information regarding the entire genomes of *Alphaentomopoxvirus* and *Gammaentomopoxvirus* members will contribute

* Corresponding author. Tel.: +81 29 838 6081; fax: +81 29 838 6028.
E-mail address: mitsuhas@affrc.go.jp (W. Mitsuhashi).

0042-6822/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.12.036

to a better understanding of EVs themselves, which are less well-studied than are ChPVs, and of the relationships between EVs, ChPVs, and other viruses, and therefore of the evolution of poxviruses. In addition, this new knowledge likely will provide insights into the features and mechanisms of horizontal gene transfer between EVs and organisms, including their host insects. EV genome analyses may yield important clues about interactions between insect viruses and their hosts, including control of host responses. In particular, compared with lepidopteran insects and their parasites, coleopteran insects generally grow slowly and *Alphaentomopoxvirus* viruses replicate slowly in their coleopteran hosts. These differences suggest the presence of as-yet unknown *Alphaentomopoxvirus*-specific genetic factors that allow the viruses to adapt or regulate their hosts.

Spheroids are candidate reagents for insect pest control (King et al., 1998). In addition, spindles – another type of inclusion body composed of the EV-encoded protein fusolin and formed in host cells – may improve the efficacy of insecticides containing insect viruses or the bacterium *Bacillus thuringiensis* (Bt) because they enhance infection by insect viruses and the insecticidal activity of Bt spores and their toxin (Mitsuhashi et al., 1998a, 2000; 2014; Mitsuhashi and Miyamoto, 2003; Wijonarko and Hukuhara, 1998; Furuta et al., 2001). Therefore, elucidating the genome of *Alphaentomopoxvirus* species will also positively contribute to these basic and applied issues.

Here we report the full genome sequence, including the terminal hairpin loop structures, of *Anomala cuprea entomopoxvirus* (ACEV), which is the first complete genome sequence of *Alphaentomopoxvirus* reported and the first description of the hairpin loop sequences in *Entomopoxvirinae* members. We also discuss the implications of the role of the hairpin loop with a region adjacent to it in DNA replication, and compare the genome with other poxviruses. Finally, we describe several noteworthy ORFs, including those potentially involved in a unique apoptosis-inhibition system.

Results and discussion

General features of the genomes

We first obtained a 225-kb contig from which most of the right inverted terminal repeat (ITR) region was missing (see Section 4). Then we have obtained the full length of the right ITR region and have confirmed the sequence of the left ITR region by "additional sequencing and assembly" described in Section 4. On the basis of results from several types of experiment described in Section 4, we have determined the full genome sequence. The left-most nucleotide of the double-stranded region was designated base 1 (nt1). The ACEV genome consists of a double-stranded region, whose size is 245,717 bp and a hairpin loop region of 88 nucleotides adjacent to each terminus of the double-stranded region. The G+C content in the double stranded region is 20.0%, which is within the range of the other EV genomes (17.8–21%) (Afonso et al., 1999; Bawden et al., 2000; Thézé et al., 2013). The genome size of *Alphaentomopoxvirus* viruses has been reported to be about 260–370 kbp according to results obtained by electron microscopy and sedimentation rates (Becker and Moyer, 2007; Skinner et al., 2011; Langridge and Roberts, 1977). However, our results indicate that the genome of the *Alphaentomopoxvirus* ACEV is 246 kbp and therefore smaller than had been reported. It is unclear whether this ACEV genome size is atypical among *Alphaentomopoxvirus*.

ACEV contains 263 putative functional ORFs (Fig. 1; Table 1). ChPV ORFs near the left and right termini of the genome tend to be transcribed towards their respective ends (Upton et al., 1988), but AMEV does not follow this convention (Bawden et al., 2000), and ACEV and MSEV appear to be intermediate between the two models (Fig. 1; Table 1; Afonso et al., 1999). On the other hand, whereas most ChPV genes in the central region were transcribed toward the center of the genome (Lefkowitz et al., 2006); neither ACEV, MSEV, nor AMEV showed any discernible pattern in this regard (Fig. 1; Table 1; Afonso et al., 1999; Bawden et al., 2000). The two trends in the orientation of gene transcription in ChPV are assumed to reduce collisions between transcription complexes (Perera et al., 2010).

The ACEV genome included the 49 genes (core genes) conserved among poxviruses (Upton et al., 2003). Mammalian ChPVs generally share a co-linear arrangement of these core genes (Skinner et al., 2011; Bawden et al., 2000: Goebel et al., 1990; Massung et al., 1994; Senkevich et al., 1997; Afonso et al., 2000). However their order was not preserved among four EV genomes (Fig. 2), although gene order may have been weakly conserved between ACEV and MSEV (Fig. 2). The gene order of *Betaentomopoxvirus* has already been reported to be non-co-linear with that of MSEV (Perera et al., 2010; Bawden et al., 2000; Thézé et al., 2013). These reports combined with our present results suggest that the order of these core genes is not conserved between different EV genera.

ITRs

Each ACEV ITR is 22, 978 bp, which is quite long for poxviruses. The ITRs of poxviruses range from $<$ 100 bp to 16 kb (Perera et al., 2010); for example, those of MSEV and AMEV are 7 kb and 9 kb, respectively (Afonso et al., 1999; Bawden et al., 2000). Exceptions to the range in poxviral ITR length are spontaneous mutants of the cow poxvirus, whose ITRs are 21–50 kb (Pickup et al., 1984) and that of 23.8 kb in CREV (Thézé et al., 2013). The ACEV ITR sequence at one terminal, once it is converted comprehensively and reversibly, is perfectly identical to that of the other ITR. Each ACEV ITR contains a series of tandemly repeated sequences, like other poxvirus ITRs. In the very terminus of each ITR, we found a putative concatemer resolution motif (Fig. 3), which is essential for the resolution of concatemeric DNA molecules during ChPV DNA replication (Smith, 2007). This motif spans nucleotides (nts) 117 through 135 downstream of each terminal nucleotide of the ACEV two-DNA strand region, consistent with a report that the motifs have been detected within the 150-bp terminal-most regions of the two DNA strands of ChPV members (Afonso et al., 2006). This is the first detection of a putative concatemer resolution motif among EVs.

Hairpin loop structures

By "Subsequent PCRs" (see Section 4) for amplifying respective hairpin loop type sequences and the sequencing of these PCR products, we could obtain two types of sequence at each genome end (Fig. 4). Therefore it was found that one type of sequence that had not been detected by the sequencing of nested PCR product from a 37 kb restriction fragment (see Section 4) is also present at the end. The two forms (referred to as types A and B) were inverted and complementary to each other, as is the case for two ChPV species, VACV (strain WR) and the Shope fibroma virus (SFV), and for African swine fever (ASF) virus (family Asfarviridae) (Baroudy et al., 1982, DeLange et al., 1986; González et al., 1986). Poxvirus terminal hairpin loop sequences have been reported only in the two species mentioned above. The sizes of the hairpin loops in ACEV, VACV (strain WR), VACV (strain Copenhagen) (Goebel et al., 1990), SFV, and ASF virus are 88, 104, 101, 64, and 37 bps, respectively. All of these sequences are highly AT rich but did not appear to share homology.

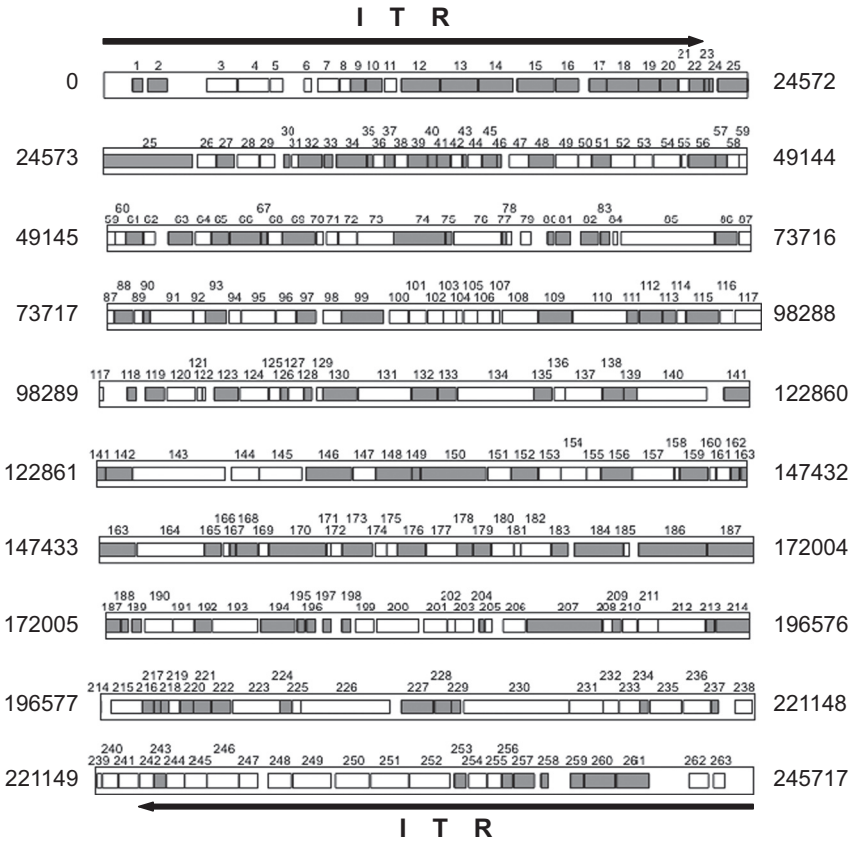

Fig. 1. Linear map of the ACEV genome. ORFs are numbered consecutively from left to right based on the position of the left terminus. ORFs transcribed to the right are shown as open boxes and those transcribed to the left as filled boxes. Nucleotide (base) numbers are listed to the left and right of the diagrams representing a part (24,572 bases) of the genome. Left-most nucleotide of the double-stranded region was designated nucleotide number 1. The figure was illustrated by using PlasDraw (GENETYX Corporation, Tokyo, Japan). ITR: inverted terminal repeat.

The present study revealed the presence of hairpin loop structures and concatemer resolution motifs very near the hairpin loops in an EV, and these "devices" are essential for DNA replication in VACV (Smith, 2007; Moss, 2007). Therefore DNA replication in ACEV is considered to be similar to that in VACV.

Are the hairpin sequences at the two ends of a molecule different or the same?

As mentioned above, two types of hairpin loop sequence that are in a relationship as an inverted repeat (IR), were detected at each terminus of the ACEV genomes. It had not been clearly elucidated whether hairpin loop sequences of two termini of one molecule in poxviruses are in a relationship as IR or whether they are identical because previous analyses of hairpin loops in each viral species or strain were performed by using a DNA sample of each end of a genome that had been separated from the rest of the genome (Baroudy et al., 1982; Moss, 2001, 2007; DeLange et al., 1986). Goebel et al. (1990) reported that they found only one type of hairpin loop sequence in VACV (strain Copenhagen), which was at both ends of the DNA double-stranded region. A model for VACV DNA replication has proposed an IR relationship between the hairpin loop at one genome end and that at the other end (Moss, 2001, 2007; Culyba et al., 2009).

In the electrophoresis of the "Subsequent PCR" products, fluorescence of the type B band derived from the left-telomeric fragment (see Section 4) induced by UV radiation was weaker than that from the right-telomeric fragment (data not shown). Furthermore, in the sequence analysis of the *Escherichia coli* clones harboring the nested PCR products, there was a significant difference in the type A to type B clone ratio between the left-terminal hairpin loop structure and right-terminal hairpin loop structure: the numbers of type A sequence and type B sequence clones were 20 and 0 at the left hairpin loops, respectively, and those at the right hairpin loops were 14 and 6, respectively (chi-square test, $p < 0.01$). These results suggest that the number of type A hairpin loops at the left end of the genome sample was greater than that at the right end and also that the number of type B hairpin loops at the left end of the genome sample was less than that at the right end; therefore, it is suggested that there is one type of ACEV genome molecules, each of which possesses two hairpin loops whose sequences are in an IR relationship with each other. The present experimental results suggest that both terminal hairpin sequences differ in one molecule of the poxvirus genome. Our present results support the proposal (Moss, 2001, 2007; Culyba et al., 2009) for an IR relationship between the two hairpin loops in one molecule of the VACV genome in the model for VACV DNA replication.

Structure/morphogenesis

ORFs sorted according to functional categories are described from this subsection "Structure/morphogenesis" to the subsection "Other functions"; ORFs that can be identified based on their similarities of sequences with known proteins whose functions remain unknown in viruses were regarded as functionally-unknown ORFs and excluded from the subsections. Some ORFs were assigned by our judgment mainly based on the function and structure of their homologs but include ORFs whose categorizations may not be absolute.

Table 1
Predicted ORFs of the ACEV genome.

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
ACV001	1498–1070	143	YP_008004312.1, CBEV N1R/p28-like protein	2.00E-15	42	○	MTG-like gene family protein
ACV002	2418–1687	244	YP_008003947.1, AHEV N1R/p28-like protein	7.00E-47	60	○	MTG-like gene family protein
ACV003	3915–5096	394	YP_008004161.1, CBEV N1R/p28-like protein	7.00E-92	46	○	N1R/p28-like protein
ACV004	5143–6291	383	YP_008004641.1, CREV N1R/p28-like protein	3.00E-94	49	○	N1R/p28-like protein
ACV005	6338–6820	161	YP_008004161.1, CBEV N1R/p28-like protein	6.00E-22	41	○	N1R/p28-like protein
ACV006	7664–7933	90	YP_008004472.1, CREV N1R/p28-like protein	2.00E-12	50	○	N1R/p28-like protein
ACV007	8180–8956	259				×	hypothetical protein
ACV008	9005–9391	129	YP_008004089.1, CBEV N1R/p28-like protein	0.002	33	○	hypothetical protein
ACV009	9943–9413	177				×	hypothetical protein
ACV010	10644–9976	223	XP_001662055.1, *Aedes aegypti* matrix metalloproteinase	4.00E-06	36	×	matrix metalloproteinase
ACV011	10743–11183	147	YP_008004481.1, CREV N1R/p28-like protein	7.00E-16	38	○	N1R/p28-like protein
ACV012	12813–11359	485	YP_008003639.1, MySEV N1R/p28-like protein	2.00E-115	46	○	MTG-like gene family protein
ACV013	14277–12850	476	YP_008003639.1, MySEV N1R/p28-like protein	2.00E-121	49	○	MTG motif gene family protein
ACV014	15603–14326	426	YP_008003639.1, MySEV N1R/p28-like protein	4.00E-116	48	○	MTG-like gene family protein
ACV015	17201–15768	478	YP_008003639.1, MySEV N1R/p28-like protein	6.00E-118	48	○	MTG-like gene family protein
ACV016	18127–17264	288	YP_008003944.1, AHEV N1R/p28-like protein	7.00E-70	49	○	MTG-like gene family protein
ACV017	19216–18521	232	YP_008003781.1, unknown similar to AMEV240 (MySEV)	8.00E-40	38	○	putative exonuclease RNase T and DNA polymerase III
ACV018	20424–19237	396	XP_004364154.1, *Capsaspora owczarzaki* ATCC 30864predicted protein	4.00E-11	30	○	leucine-rich repeat gene family protein
ACV019	21212–20421	264	NP_065057.1, AMEV hypothetical protein AMVITR05b	0.02	26	○	hypothetical protein
ACV020	21900–21232	223	XP_004017602.1, PREDICTED: *Ovis aries* LOW QUALITY PROTEIN: matrix metalloproteinase-17	2.00E-04	25	○	hypothetical protein
ACV021	21930–22373	148				×	hypothetical protein
ACV022	22916–22371	182	ADR00582.1, *Gossypium hirsutum* somatic embryogenesis receptor-like kinase 1 protein	0.05	31	×	hypothetical protein
ACV023	23113–22943	57				×	hypothetical protein
ACV024	23275–23126	50				×	hypothetical protein
ACV025	27957–23440	1506	ENN70211.1, *Dendroctonus ponderosae* hypothetical protein YQE_12997, partial	0	31	○	putative ATP-binding cassette transporter
ACV026	28152–28913	254	XP_002047944.1, *Drosophila virilis* GJ11638	6.00E-41	35	×	phosphatidic acid phosphatase type 2
ACV027	29581–28916	222				×	hypothetical protein
ACV028	29703–30509	269	YP_008003624.1, MySEV ser/thr kinase (Cop-B1R)	2.00E-18	31	○	protein kinase
ACV029	30556–31101	182	YP_008004440.1, unknown similar to AMEV022 (CREV)	1.00E-13	40	○	hypothetical protein

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
ACV030	31673–31458	72				×	hypothetical protein
ACV031	31762–32040	93				×	hypothetical protein
ACV032	32920–32039	294	NP_048265.1, MSEV ALI motif gene family protein(MSV194)	3.00E-18	33	○	ALI motif gene family protein
ACV033	33337–33002	112				×	hypothetical protein
ACV034	34596–33478	373	BAA25629.1, ACEV fusolin	0	100	○	fusolin
ACV035	34885–34640	82				×	hypothetical protein
ACV036	34914–35348	145				×	hypothetical protein
ACV037	35698–35321	126	NP_064865.1, AMEV hypothetical protein AMV083	6.00E-25	59	○	putative membrane protein
ACV038	35703–36215	171				×	hypothetical protein
ACV039	36933–36199	245	YP_008004337.1, CBEV uracil-DNA glycosylase, DNA polymerase processivity factor	1.00E-68	49	○	putative uracil DNA glycosylase UNG
ACV040	37299–36964	112	YP_008004356.1, CBEV IMV membrane protein entry/ fusion complex component (Cop-A21L)	3.00E-46	63	○	IMV membrane protein involved in fusion and entry
ACV041	37830–37303	176	WP_022119592.1, *Firmicutes bacterium* CAG:56 na/Pi-cotransporter family protein	0.004	27	×	hypothetical protein
ACV042	37895–38305	137	YP_008004034.1, AHEV putative late 16kDa membrane protein (Cop-J5L)	7.00E-50	57	○	membrane protein
ACV043	38461–38306	52				×	hypothetical protein
ACV044	38508–39047	180	YP_008004030.1, AHEV RNA polymerase RPO18	1.00E-39	46	○	putative RNA polymerase subunit
ACV045	39616–39044	191	YP_008004332.1, unknown similar to AMEV226 (CBEV)	6.00E-08	43	○	hypothetical protein
ACV046	39604–39771	56				×	hypothetical protein
ACV047	40070–40849	260	YP_008004339.1, unknown similar to AMEV233 (CBEV)	2.00E-19	46	○	formyltetrahydrofolate synthetase
ACV048	41792–40845	316	WP_002698739.1, *Microscilla marina* deoxyuridine 5′-triphosphate nucleotidohydrolase family	6.00E-32	32	○	putative dUTPase
ACV049	41854–42690	279	NP_064875.1, AMEV putative mRNA capping enzyme small subunit (AMV093)	1.00E-79	53	○	putative mRNA capping enzyme small subunit
ACV050	42721–43212	164				×	hypothetical protein
ACV051	43967–43215	251				×	hypothetical protein
ACV052	43993–44856	288	YP_008003630.1, unknown similar to AMEV090 (MySEV)	6.00E-50	46	○	hypothetical protein
ACV053	44884–45540	219	YP_717593.1, *Clanis bilineata* nucleopolyhedrovirus Ld138-like protein	7.00E-11	32	×	nicotinamide riboside kinase 1
ACV054	45585–46577	331	YP_008003634.1, unknown similar to AMEV096 (MySEV)	9.00E-109	53	○	hypothetical protein
ACV055	46650–46853	68				×	hypothetical protein
ACV056	47946–46948	333	YP_008004471.1, CREV entry-fusion complex component, myristylprotein	2.00E-123	56	○	poxvirus myristoyl protein
ACV057	48382–47966	139	YP_008004634.1, CREV Ca2+ BP	2.00E-36	51	○	putative calcium binding protein
ACV058	48398–48847	150	NP_048167.1, MSEV hypothetical protein MSV096	2.00E-04	33	○	hypothetical protein
ACV059	48847–49452	202		6.00E-18	49	○	

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
			YP_008003931.1, unknown similar to AMEV102 (AHEV)				hypothetical protein
ACV060	49448–49837	130				×	hypothetical protein
ACV061	50517–49840	226	YP_008004037.1, unknown similar to AMEV235 (AHEV)	2.00E-69	58	○	metallophosphoesterase domain-containing protein
ACV062	50575–50988	138				×	hypothetical protein
ACV063	52396–51482	305	YP_008003825.1, AHEV N1R/p28-like protein	3.00E-57	42	○	KilA-N domain-containing protein
ACV064	52503–53093	197	YP_008004177.1, CBEV hypothetical protein CHBEV_107	1.00E-31	50	○	hypothetical protein
ACV065	53799–53119	227				×	hypothetical protein
ACV066	55001–53850	384	NP_048161.1, MSEV putative Molluscum contagiosum virus MC121L (vaccinia A16L) homolog (MSV090)	9.00E-179	68	○	putative myristylated membrane protein
ACV067	55237–55013	75				×	hypothetical protein
ACV068	55274–55813	180	NP_048164.1, MSEV putative vaccinia E10R homolog (MSV093)	2.00E-35	65	○	sulfhydryl oxidase
ACV069	57090–55816	425	Q05894.1, *Heliothis armigera entomopoxvirus* fusolin	5.00E-43	48	○	fusolin-like protein
ACV070	57163–57399	79				×	hypothetical protein
ACV071	57522–57947	142	NP_048203.1, MSEV putative AMEV G4R homolog (vaccinia A28L) (MSV132)	8.00E-64	69	○	membrane protein
ACV072	57989–58660	224				×	hypothetical protein
ACV073	58694–60103	470	ERP50118.1, *Populus trichocarpa* hypothetical protein POPTR_0017s00560g	0.003	25	×	hypothetical protein
ACV074	62037–60100	646	YP_008003739.1, MySEV ATPase, NPH1	0	65	○	nucleoside triphosphate phosphohydrolase I
ACV075	62288–62052	79				×	hypothetical protein
ACV076	62360–64180	607	NP_048126.1,MSEV hypothetical protein MSV055	2.00E-08	36	○	translation elongation factor eEF-3 like
ACV077	64388–64224	55				×	hypothetical protein
ACV078	64393–64575	61				×	hypothetical protein
ACV079	64930–65298	123	XP_001640022.1, *Nematostella vectensis* predicted protein	3.00E-09	45	×	hypothetical protein
ACV080	66170–65925	82				×	hypothetical protein
ACV081	66830–66267	188	ESO86625.1, *Lottia gigantea* hypothetical protein LOTGIDRAFT_166892	0.001	34	×	hypothetical protein
ACV082	67892–67224	223	YP_008004407.1, unknown similar to AcMNPV orf7 (CREV)	4.00E-35	45	○	hypothetical protein
ACV083	68302–67973	110				×	hypothetical protein
ACV084	68437–68601	55				×	hypothetical protein
ACV085	68733–72302	1190	NP_064848.1, AMEV DNA-directed RNA polymerase subunit (AMV066)	0	64	○	DNA-directed RNA polymerase subunit RPO132
ACV086	73201–72341	287				×	serpin

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
ACV087	73274–73960	229	XP_001007673.1, *Tetrahymena thermophila* protein kinase domain containing protein	6.00E-12	33	×	protein kinase domain containing protein
ACV088	74681–73977	235	YP_008003742.1, unknown similar to AMEV200 (MySEV)	3.00E-22	38	○	hypothetical protein
ACV089	74727–75050	108				×	hypothetical protein
ACV90	75310–75065	82	YP_008003740.1, unknown similar to AMEV198 (MySEV)	3.00E-20	57	○	hypothetical protein
ACV91	75340–76932	531	YP_008003741.1, MySEV NAD-dependent DNA ligase	4.00E-164	51	○	putative NAD+ dependent DNA ligase
ACV92	76952–77392	147	ADJ67808.1, *Hypophthalmichthys molitrix* copper/zinc superoxide dismutase	4.00E-46	53	○	superoxide dismutase
ACV93	78180–77401	260	NP_048258.1, MSEV putative late transcription factor VLTF-2 homolog (vaccinia A1L) (MSV187)	3.00E-57	44	○	putative late transcription factor VLTF-2
ACV94	78236–78733	166	YP_008003602.1, unknown similar to AMEV062 (MySEV)	2.00E-35	45	○	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
ACV95	78753–80024	424	NP_064996.1, AMEV hypothetical protein AMV214	1.00E-88	44	○	hypothetical protein
ACV96	80067–80813	249	YP_008004487.1, CREV ss/dsDNA binding protein VP8 (Cop-L4R)	1.00E-54	47	○	ssDNA/dsDNA binding protein
ACV97	81554–80820	245	NP_064999.1, AMEV putative myristylated membrane protein (AMV217)	8.00E-113	71	○	putative myristylated membrane protein
ACV98	81831–82502	224				×	hypothetical protein
ACV99	84092–82518	525	YP_008003758.1, unknown similar to AMEV216 (MySEV)	9.00E-51	31	○	hypothetical protein
ACV100	84312–85016	235	YP_008004580.1, CREV RNA polymerase RPO19	1.00E-51	46	○	DNA-dependent RNA polymerase subunit rpo19
ACV101	85059–85688	210	YP_007317337.1, *Procambarus fallax* NADH dehydrogenase subunit 5 (mitochondrion)	0.039	33	×	cation transport ATPase
ACV102	85718–86332	205	YP_008003717.1, unknown similar to AMEV164 (MySEV)	1.00E-59	53	○	tryptophan-rich sensory protein
ACV103	86350–86826	159	YP_008003995.1, AHEV holliday junction resolvase	7.00E-44	50	○	holliday junction resolvase
ACV104	86837–87061	75	YP_008003714.1, MySEV viral membrane associated, early morphogenesis protein (Cop-A9L)	1.00E-25	64	○	virion membrane protein
ACV105	87110–87679	190				×	hypothetical protein
ACV106	87630–88184	185	YP_008003998.1, unknown similar to AMEV159 (AHEV)	9.00E-06	39	○	hypothetical protein
ACV107	88204–88470	89				×	hypothetical protein
ACV108	88562–89914	451	YP_008003979.1, AHEV virion core cysteine protease	3.00E-168	56	○	virion core cysteine protease
ACV109	91180–89906	425	YP_008003732.1, MySEV FEN1-like nuclease (Cop-G5R)	3.00E-57	36	○	FEN1-like nuclease
ACV110	91213–93228	672	NP_064956.1, AMEV putative early transcription factor small subunit (AMV174)	0	69	○	putative early transcription factor small subunit
ACV111	93654–93235	140	YP_008004654.1, unknown similar to AMEV247 (CREV)	4.00E-32	50	○	hypothetical protein

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
ACV112	94564–93710	285	YP_008004048.1, AHEV IMV heparin binding surface protein	3.00E-94	53	○	putative glycosyl transferase
ACV113	95074–94574	167	YP_008003916.1, unknown similar to AMEV098 (AHEV)	8.00E-21	42	○	hypothetical protein
ACV114	95114–95452	113	YP_008003729.1, unknown similar to AMEV172 (MySEV)	3.00E-23	50	○	hypothetical protein
ACV115	96689–95460	410	XP_003743508.1, PREDICTED: *Metaseiulus occidentalis* serine/threonine-protein kinase VRK1-like	2.00E-31	28	○	putative ser/thr protein kinase
ACV116	96736–97248	171	ENN77264.1, *Dendroctonus ponderosae* hypothetical protein YQE_06093, partial	7.00E-53	50	×	translation initiation factor eIF-4E
ACV117	97305–98429	375	XP_637909.1, *Dictyostelium discoideum* AX4 hypothetical protein DDB_G0286039	2.00E-36	32	○	inhibitor of apoptosis protein
ACV118	99675–99334	114				×	hypothetical protein
ACV119	100725–100018	236	XP_002399419.1, *Ixodes scapularis* putative leucine-rich repeat (LRR) protein	3.00E-08	28	×	leucine-rich repeat protein
ACV120	100825–101892	356	EFN66173.1, *Camponotus floridanus* G-protein coupled receptor Mth2	4.00E-49	33	○	G-protein coupled receptor Mth2-like
ACV121	101957–102148	64				×	hypothetical protein
ACV122	102157–102318	54	BAF79968.1, *Closterium ehrenbergii* receptor-like kinase	0.007	48	×	serin/threonine kinase like
ACV123	103556–102648	303	XP_005802524.1, PREDICTED: *Xiphophorus maculatus* tenascin-like	1.00E-12	29	×	tenascin
ACV124	103618–104670	351				×	hypothetical protein
ACV125	104716–105120	135				×	hypothetical protein
ACV126	105414–105136	93	YP_008003728.1, unknown similar to AMEV171 (MySEV)	3.00E-32	59	○	hypothetical protein
ACV127	105443–106021	193	YP_008004006.1, unknown similar to AMEV145 (AHEV)	5.00E-34	42	○	hypothetical protein
ACV128	106324–106019	102	NP_064950.1, AMEV hypothetical protein AMV168	6.00E-14	41	○	hypothetical protein
ACV129	106507–106731	75				×	hypothetical protein
ACV130	108005–106734	424	YP_008004286.1, unknown similar to AMEV173 (CBEV)	2.00E-57	39	○	hypothetical protein
ACV131	108063–110057	665	YP_008003703.1, P4b precursor (MySEV)	0	47	○	putative core protein
ACV132	111028–110060	323	NP_064922.1, AMEV hypothetical protein AMV140	1.00E-15	25	○	hypothetical protein
ACV133	111786–111064	241	YP_008003711.1, unknown similar to AMEV157 (MySEV)	4.00E-50	47	○	hypothetical protein
ACV134	111848–114739	964	YP_008004269.1, unknown similar to AMEV156 (CBEV)	3.00E-11	27	○	prominin
ACV135	115429–114707	241	YP_008004003.1, AHEV ATPase/DNA packaging protein	2.00E-75	54	○	putative ATP/GTP-binding protein
ACV136	115469–115879	137	YP_008004002.1, unknown similar to AMEV151(AHEV)	0.046	35	○	hypothetical protein
ACV137	115894–117282	463	YP_008004001.1, AHEV essential ser/thr kinase morph (Cop-F10L)	2.00E-133	50	○	putative serine/threonine protein kinase
ACV138	118076–117285	264	ENH62146.1, *Fusarium oxysporum* f. sp. cubense race 1 Protein roadkill	6.00E-06	29	×	BTB/POZ domain-containing protein
ACV139	118576–118091	162	NP_048139.1, MSEV hypothetical protein MSV068	1.00E-29	42	○	hypothetical protein
ACV140	118585–121257	891		0	51	○	

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
			NP_064917.1, AMEV putative mRNA capping enzyme large subunit (AMV135)				putative mRNA capping enzyme large subunit
ACV141	123184–121898	429	YP_008003832.1, unknown similar to AMEV033 (AHEV)	2.00E-20	93	○	hypothetical protein
ACV142	124182–123202	327	YP_008004011.1, AHEV viral membrane formation (Cop-A11R)	1.00E-44	39	○	virion membrane formation
ACV143	124207–127716	1170	YP_008004010.1, AHEV P4a precursor	2.00E-84	33	○	precursor p4a of core protein 4a
ACV144	127960–128997	346	NP_048251.1, MSEV putative vaccinia L3L homolog (MSV180)	5.00E-89	46	○	early gene transcription related protein
ACV145	129020–130609	530	NP_064826.1, AMEV hypothetical protein AMV044	2.00E-23	27	○	hypothetical protein
ACV146	132476–130776	567	YP_008004155.1, CBEV poly (A) polymerase catalytic subunit VP55	0	56	○	putative poly (A) polymerase large subunit
ACV147	132543–133391	283	NP_048221.1, MSEV putative NTP pyrophosphohydrolase mutT motif homolog (vaccinia D10R) (MSV150)	3.00E-58	47	○	putative NTP pyrophosphohydrolase mutT motif
ACV148	134781–133390	464	YP_008003599.1, MySEV DNA helicase, transcript release factor	1.00E-126	48	○	DNA helicase, transcriptional elongation
ACV149	135112–134804	103				×	hypothetical protein
ACV150	137582–135141	814	YP_008003596.1, MySEV RAP94 RNA pol assoc protein	0	49	○	putative RNA polymerase associated transcriptional specificity factor
ACV151	137647–138522	292	YP_008004476.1, unknown similar to AMEV053 (CREV)	3.00E-09	34	○	hypothetical protein
ACV152	139526–138525	334	YP_008003875.1, AHEV DNA Topoisomerase type I	4.00E-126	61	○	DNA topoisomerase type I
ACV153	139572–140396	275	NP_048132.1, MSEV putative LINE reverse transcriptase (MSV061)	6.00E-07	29	○	reverse transcriptase
ACV154	140389–141405	339	NP_048220.1, MSEV putative RPO35 homolog (vaccinia A29L) (MSV149)	1.00E-57	39	○	DNA-dependent RNA polymerase RPO35
ACV155	141412–141939	176	NP_064822.1, AMEV hypothetical protein AMV040	6.00E-17	37	○	hypothetical protein
ACV156	143116–141947	390	NP_048121.1, MSEV hypothetical protein MSV050	1.00E-14	30	○	hypothetical protein
ACV157	143133–144683	517				×	hypothetical protein
ACV158	144725–144901	59				×	hypothetical protein
ACV159	145971–144910	354	YP_008003700.1, unknown similar to AMEV141(MySEV)	2.00E-69	44	○	hypothetical protein
ACV160	146031–146261	77	CDI81783.1, *Eimeria praecox* ubiquitin / ribosomal protein CEP52 fusion protein, putative	5.00E-40	87	○	ubiquitin/ ribosomal protein
ACV161	146302–146802	167	XP_001021849.1, *Tetrahymena thermophila* ubiquitin-conjugating enzyme family protein	4.00E-23	42	×	ubiquitin-conjugating enzyme E2
ACV162	147134–146829	102	XP_001816367.1, PREDICTED: similar to *Tribolium castaneum* GA19017-PA	6.00E-14	49	×	hypothetical protein
ACV163	148783–147200	528	YP_008003730.1, MySEV virion protein (Cop-E6R)	1.00E-144	50	○	core protein
ACV164	148841–151369	843	YP_008004214.1, CBEV NTPase, DNA primase	0	47	○	putative NTPase
ACV165	152009–151377	211	NP_048159.1, MSEV hypothetical protein MSV088	3.00E-34	44	○	hypothetical protein
ACV166	152081–152323	81				×	peptidoglycan binding domain-

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
							containing protein
ACV167	152539–152330	70	YP_008003619.1, MySEV putative thioredoxin	1.00E-15	46	○	putative thioredoxin
ACV168	153368–152565	268	XP_005998662.1, PREDICTED: *Latimeria chalumnae* matrix metalloproteinase-20-like	7.00E-25	39	×	zinc-dependent metalloprotease
ACV169	153417–153791	125	YP_008003897.1, unknown similar to AMEV080 (AHEV)	2.00E-10	40	○	hypothetical protein
ACV170	155926–153794	711	YP_008003621.1, MySEV RNA helicase, DExH-NPH-II domain	0	52	○	RNA helicase NPH-II
ACV171	155974–156123	50				×	hypothetical protein
ACV172	156136–156573	146	YP_008003932.1, unknown similar to AMEV101 (AHEV)	7.00E-31	45	○	capcid protein VP1-like
ACV173	157686–156547	380				×	hypothetical protein
ACV174	157798–158232	145	YP_008003651.1, unknown similar to AMEV120 (MySEV)	7.00E-18	38	○	ATP-dependent DNA ligase
ACV175	158228–158665	146				×	hypothetical protein
ACV176	159683–158622	354	NP_064901.1, AMEV putative protein phosphatase 2C (AMV119)	3.00E-108	51	○	putative protein phosphatase 2C
ACV177	159721–160848	376	YP_008004421.1, CREV tryptophan repeat gene family	2.00E-10	28	○	tryptophan repeat gene family protein
ACV178	161463–160861	201	YP_008004553.1, CREV entry-fusion complex essential component (Cop-H2R)	7.00E-93	66	○	putative viral membrane protein
ACV179	162126–161473	218	AGB75902.1, Vaccinia virus hypothetical protein	8.00E-06	33	○	nucleopolyhedrovirus p26 protein
ACV180	162174–162992	273	NP_048112.1, MSEV putative poly(A) polymerase small subunit PAP-S homolog (vaccinia J3R) (MSV041)	3.00E-19	31	○	putative poly (A) polymerase small subunit
ACV181	163023–163250	76				×	hypothetical protein
ACV182	163274–164407	378	NP_001071929.1, *Ciona intestinalis* zinc finger protein	3.00E-04	45	○	RING zinc finger-containing protein
ACV183	165276–164410	289	NP_064915.1, AMEV triacylglycerol lipase (AMV133)	2.00E-94	55	○	lipase
ACV184	167106–165280	609	YP_008004054.1, AHEV metalloprotease (Cop-G1L)	1.00E-51	33	○	putative vaccinia G1L metaloprotease
ACV185	167128–167334	69	YP_008004347.1, unknown similar to AMEV241 (CBEV)	0.044	36	○	hypothetical protein
ACV186	170245–167690	852	YP_008004202.1, CBEV NTPase/helicase	0	42	○	putative NTPase/helicase
ACV187	172559–170277	761	YP_008004229.1, CBEV VETF-L early transcription factor large	0	61	○	putative early transcription factor large subunit
ACV188	172845–172567	93				×	hypothetical protein
ACV189	173380–172973	136	YP_008003909.1, unknown similar to AMEV088 (AHEV)	9.00E-25	44	○	hypothetical protein
ACV190	173508–174554	349	NP_064873.1, AMEV DNA-directed RNA polymerase beta chain (AMV091)	3.00E-72	49	○	DNA-directed RNA polymerase beta chain
ACV191	174576–175373	266	YP_008004541.1, unknown similar to AMEV121 (CREV)	1.00E-53	43	○	hypothetical protein
ACV192	176046–175381	222	NP_048136.1, MSEV putative late transcription factor VLTF-3 homolog (vaccinia A2L) (MSV065)	6.00E-81	58	○	late transcription factor VLTF-3
ACV193	176101–177798	566		0	66	○	

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
			YP_008003693.1, MySEV trimeric virion coat protein rifampicin res				rifampicin resistance protein
ACV194	179252–177915	446	NP_064881.1, AMEV hypothetical protein AMV099	4.00E-47	34	○	hypothetical protein
ACV195	179609–179343	89				×	hypothetical protein
ACV196	180027–179686	114	XP_002020894.1, *Drosophila persimilis* GL16255	5.00E-06	40	×	PDGF- and VEGF-related factor 2
ACV197	180604–180299	102	NP_064959.1, AMEV putative antirepressor (AMV177)	3.00E-11	47	○	N1R/p28-like protein
ACV198	181358–181014	115	YP_008004159.1, CBEV N1R/p28-like protein	7.00E-14	63	○	N1R/p28-like protein
ACV199	181524–182261	246				×	hypothetical protein
ACV200	182383–183948	522	WP008647708.1, *Bacteroides* sp. 3_1_23 conserved hypothetical protein	0.021	25	×	hypothetical protein
ACV201	184146–185060	305	YP_008003953.1, AHEV N1R/p28-like protein	3.00E-45	41	○	KilA-N domain-containing protein
ACV202	185091–185369	93	AAB96622.1, *Heliothis armigera* EV 17K ORF	5.00E-09	50	○	KilA-N domain-containing protein
ACV203	185369–186049	227	YP_008003520.1, MySEV N1R/p28-like protein	8.00E-48	52	○	KilA-N domain-containing protein
ACV204	186444–186274	57				×	hypothetical protein
ACV205	186489–186758	90	XP_001604408.1, PREDICTED: *Nasonia vitripennis* hypothetical protein LOC100120806	0.005	46	○	Peritrophin-A domain containing protein
ACV206	187171–188058	296	NP_065001.1, AMEV putative N-myristoyl transferase (AMV219)	1.00E-32	38	○	putative N-myristoyl transferase
ACV207	190938–188113	942	BAA33399.1, ACEV spheroidin	0	100	○	spheroidin
ACV208	191034–191360	109				×	hypothetical protein
ACV209	191695–191369	109				×	hypothetical protein
ACV210	191759–192295	179	NP_064798.1, AMEV thymidine kinase (AMV016)	2.00E-32	48	○	thymidine kinase
ACV211	192337–193098	254	XP_001306343.1, *Trichomonas vaginalis* G3 hypothetical protein	1.00E-07	35	○	leucine rich repeat gene family protein
ACV212	193102–194931	610	NP_064992.1, AMEV putative DNA polymerase beta/AP endonuclease (AMV210)	0	55	○	putative DNA polymerase beta/AP endonuclease
ACV213	195261–194932	110				×	hypothetical protein
ACV214	196594–195314	427	YP_008003832.1, unknown similar to AMEV033 (AHEV)	4.00E-21	27	○	hypothetical protein
ACV215	196962–198131	390	YP_008004118.1, unknown similar to AMEV020 (CBEV)	1.00E-146	67	○	hypothetical protein
ACV216	198575–198150	142	NP_048208.1, MSEV hypothetical protein MSV137	3.00E-19	42	○	hypothetical protein
ACV217	198824–198582	81				×	hypothetical protein
ACV218	199547–198837	237	YP_008003785.1, MySEV S-S bond formation pathway protein substrate (Cop-F9L)	3.00E-50	46	○	S-S bond formation pathway protein
ACV219	199144–199563	140	NP_065026.1, AMEV hypothetical protein AMV244	3.00E-09	53	○	hypothetical protein
ACV220	200053–199553	167	NP_955165.1, Canarypox virus CNPV142 N1R/p28-like protein	6.00E-05	36	○	KilA-N domain-containing protein
ACV221	200725–200066	220	XP_003546035.1, PREDICTED: *Glycine max* probable LRR	0.005	25	×	leucine rich repeat

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
			receptor-like serine/ threonine-protein kinase At2g24230-like				containing protein
ACV222	201460–200738	241	XP_001621475.1, *Nematostella vectensis* hypothetical protein NEMVEDRAFT_v1g144768	3.00E-06	33	×	hypothetical protein
ACV223	201526–203310	595	YP_008004461.1, unknown similar to AMEV045 (CREV)	4.00E-14	30	○	putative glycoprotein B
ACV224	203756–203307	150	YP_008003886.1, unknown similar to AMEV072 (AHEV)	2.00E-22	46	○	gamma-glutamyl cyclotransferase like
ACV225	203774–204109	112	NP_048120.1, MSEV hypothetical protein MSV049	1.00E-20	51	○	hypothetical protein
ACV226	204105–207446	1114	YP_008004452.1, CREV DNA polymerase	0	42	○	DNA polymerase
ACV227	209055–207868	396	EFN86490.1, *Harpegnathos saltator* collagen-like protein 5	2.00E-08	23	×	hypothetical protein
ACV228	209743–209120	208	YP_008003863.1, AHEV NLPc/ P60 superfamily protein (Cop-G6R)	8.00E-58	60	○	in vivo virulence related protein
ACV229	210133–209750	128				×	hypothetical protein
ACV230	210213–214166	1318	YP_008004022.1, AHEV RNA polymerase RPO147	0	62	○	putative RNA polymerase largest subunit
ACV231	214190–215467	426				×	hypothetical protein
ACV232	215460–216026	189				×	hypothetical protein
ACV233	216058–216831	258				×	hypothetical protein
ACV234	217149–216829	107				×	hypothetical protein
ACV235	217212–218393	394	YP_007630399.1, *Edwardsiella tarda* C07–087 putative eliminase	2.00E-25	30	×	right handed beta helix region containing protein
ACV236	218465–219520	352	WP_009347858.1, *Alloprevotella rava* hypothetical protein	6.00E-17	38	○	cell wall surface anchor family protein
ACV237	219788–219504	95				×	hypothetical protein
ACV238	220407–221048	214				×	hypothetical protein
ACV239	221211–221375	55				×	hypothetical protein
ACV240	221413–221991	193	XP_004965107.1, PREDICTED: *Setaria italica* metalloendoproteinase 1-like	3.00E-05	28	×	zinc-dependent metalloprotease
ACV241	222019–222756	246	XP_003746806.1, PREDICTED: *Metaseiulus occidentalis* 72 kDa type IV collagenase-like	1.00E-08	35	×	zinc-dependent metalloprotease
ACV242	222802–223347	182	ADR00582.1, *Gossypium hirsutum* somatic embryogenesis receptor-like kinase 1 protein	0.05	31	×	hypothetical protein
ACV243	223788–223345	148				×	hypothetical protein
ACV244	223818–224486	223	XP_004017602.1, PREDICTED: *Ovis aries* LOW QUALITY PROTEIN: matrix metalloproteinase-17	2.00E-04	25	×	hypothetical protein
ACV245	224506–225297	264	NP_065057.1, AMEV hypothetical protein AMVITR05b	0.02	26	○	hypothetical protein
ACV246	225294–226481	396	XP_004364154.1, *Capsaspora owczarzaki* ATCC 30864predicted protein	4.00E-11	30	○	leucine rich repeat gene family protein
ACV247	226502–227197	232		8.00E-40	38	○	

Table 1 (*continued*)

ORF	Position[a]	aa[b]	Best match[c]	*E* value[c]	Identity[c]	ORF(s) similar to that of ACEV in other known EVs (presence, ○; absence, ×)	Product, function, or structure
			YP_008003781.1, unknown similar to AMEV240 (MySEV)				putative exonuclease RNase T and DNA polymerase III
ACV248	227591–228454	288	YP_008003944.1, AHEV N1R/p28-like protein	7.00E-70	49	○	MTG-like gene family protein
ACV249	228517–229950	478	YP_008003639.1, MySEV N1R/p28-like protein	6.00E-118	48	○	MTG-like gene family protein
ACV250	230115–231392	426	YP_008003639.1, MySEV N1R/p28-like protein	4.00E-116	48	○	MTG-like gene family protein
ACV251	231441–232868	476	YP_008003639.1, MySEV N1R/p28-like protein	2.00E-121	49	○	MTG motif gene family protein
ACV252	232905–234359	485	YP_008003639.1, MySEV N1R/p28-like protein	2.00E-115	46	○	MTG-like gene family protein
ACV253	234975–234535	147	YP_008004481.1, CREV N1R/p28-like protein	7.00E-16	38	○	N1R/p28-like protein
ACV254	235074–235742	223	XP_001662055.1, *Aedes aegypti* matrix metalloproteinase	4.00E-06	36	×	matrix metalloproteinase
ACV255	235775–236305	177				×	hypothetical protein
ACV256	236713–236327	129	YP_008004089.1, CBEV N1R/p28-like protein	0.002	33	○	hypothetical protein
ACV257	237538–236762	259				×	hypothetical protein
ACV258	238054–237785	90	YP_008004472.1, CREV N1R/p28-like protein	2.00E-12	50	○	N1R/p28-like protein
ACV259	239380–238898	161	YP_008004161.1, CBEV N1R/p28-like protein	6.00E-22	41	○	N1R/p28-like protein
ACV260	240575–239427	383	YP_008004641.1, CREV N1R/p28-like protein	3.00E-94	49	○	N1R/p28-like protein
ACV261	241803–240622	394	YP_008004161.1, CBEV N1R/p28-like protein	7.00E-92	46	○	N1R/p28-like protein
ACV262	243300–244031	244	YP_008003947.1, AHEV N1R/p28-like protein	7.00E-47	60	○	MTG-like gene family protein
ACV263	244220–244648	143	YP_008004312.1, CBEV N1R/p28-like protein	2.00E-15	42	○	MTG-like gene family protein

AMEV, *Amsacta moorei entomopoxvirus*. CBEV, *Choristoneura biennis entomopoxvirus*. AHEV, *Adoxophyes honmai entomopoxvirus*. CREV, *Choristoneura rosaceana entomopoxvirus*. MySEV, *Mythimna separata entomopoxvirus*. MSEV, *Melanoplus sanguinipes entomopoxvirus*.

[a] The left-most nucleotide of the double-stranded region was designated as base no. 1. Stop codon was not included.

[b] Amino acids.

[c] The top hit in blastp in NCBI, whose *E* value was less than 0.05, was listed. PSI-BLAST analysis was performed for top hit whose *E* value was less than 0.05, if there were no hits in blastp.

ChPV homologs in ACEV in the category Structure/morphogenesis are shown in Table 2. The others in this category are thioredoxin (ACV167 (ORF167)) and spheroidin (ACV207; Mitsuhashi et al., 1998b). AMEV and MSEV contain homologs of all ACEV ORFs in this category.

ACV184 is a homolog of VACV G1L, which is a virion core protein that is considered to be a metalloprotease. In addition to ACV184, at least four more ACEV ORFs (ACV010, ACV168, ACV240, and ACV241) were detected as metalloproteases, but their functions are unclear. Therefore, ACEV has more metalloprotease ORFs than do MSEV and AMEV (3 and 1, respectively) (Afonso et al., 1999; Bawden et al., 2000). The four metalloprotease ORFs of unknown function in ACEV showed no similarity to other EV ORFs. As has been suggested for ORFs in MSEV (MSV176 and MSV179), at least ACV010, ACV240 and ACV241 may be extracellular metalloproteases because they all have a His-Glu-2X-His domain and putative signal peptide and show similarities to matrixins (Afonso et al., 1999; Jongeneel et al., 1989; Rawlings and Barrett, 1995). None of the metalloproteases in ACEV showed similarities to the baculovirus metalloprotease, enhancin; AMVITR10 in AMEV is the only EV ORF to show any, albeit low, similarity to enhancin (Bawden et al., 2000).

Transcription/RNA modification

In addition to the ChPV homologs present (Table 2), ACEV contains an ORF for NTPase/helicase (ACV186).

Among the ChPV homologs in ACEV (Table 2) is a homolog of BTB/POZ domain-containing protein (ACV138), which does not occur in other EVs. This domain is present near the N-terminus of proteins that contain the Kelch domain in ChPVs, and is known to be involved in protein–protein interactions such as regulation of transcription in eukaryotes (Bardwell and Treisman, 1994). Homologs of all the other ChPV homologs in ACEV in this category are present in AMEV and MSEV. It is interesting that a homolog of VLTF-1, which is essential for VACV replication and is conserved in ChPVs, was not detected in ACEV, AMEV, or MSEV. However, it seems that the transcriptional process is well conserved among

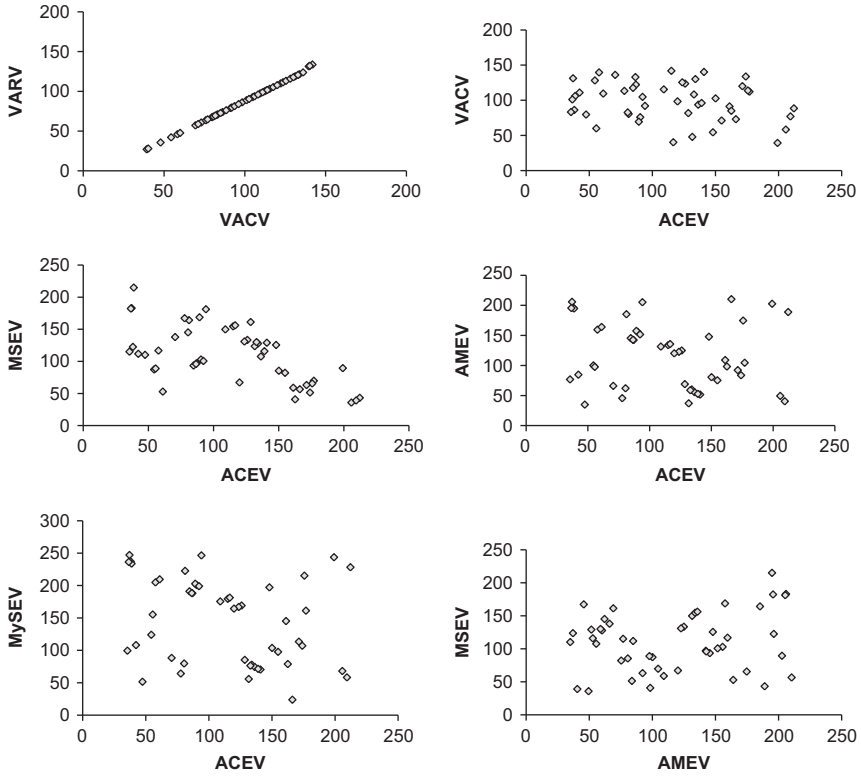

Fig. 2. Gene parity plots between several poxviruses. Forty-nine conserved poxvirus genes (ORFs) (Upton et al., 2003) are plotted. The horizontal and vertical axes represent the relative position of each ORF along the genome in kb. AMEV, *Amsacta moorei entomopoxvirus*; MSEV, *Melanoplus sanguinipes entomopoxvirus*; MySEV, *Mythimna separata entomopoxvirus*; VACV, vaccinia virus Copenhagen; VARV, variola virus. GenBank accession numbers: AMEV, NC_002520; MSEV, NC_001993; MySEV, HF679134; VACV, M35027 (Goebel et al., 1990); VARV, NC_001611 (Shchelkunov et al., 1994, 1996; Shchelkunov and Totmenin, 1995).

ACEV	ATTTATTATGA-TAAAAAAA
Rabbit fibroma virus	ATTTATAACCCTAGAAAAAA
Myxoma virus	ATTTATAGCTCTTAAAAAAA
Yaba-like disease virus	ATTTATAACTGGAAAAAAAA
Sheeppox virus	ATTTATAGG-CTTAAAAAAA
Fowlpox virus	ATTTATATA--GTAAAAAAA
Swinepox virus	ATTTATAAACGGTAAAAAAA
Vaccinia virus	ATTTAGTGTCTAGAAAAAAA
Crocodilepox virus	ATTTATAGCCTGTAAAAAAA
	*****.. ..******

Fig. 3. Alignment of the putative ACEV concatemer resolution motif with those of other poxviruses. The motif is underlined. The analysis was performed by using GENETYX MAC ver.15 (GENETYX Corporation, Tokyo, Japan) and the resultant alignment was slightly modified manually. Asterisks indicate identity among the nine nucleotides, and dots indicate nucleotides that are conserved more than 50% among the nine nucleotides. Genbank accession numbers: Shope (Rabbit) fibroma virus, AF170722 (Willer et al., 1999); myxoma virus, AF170726 (Cameron et al., 1999); Yaba-like disease virus, AY386371 (Brunetti et al., 2003); sheeppox virus, M28823 (Gershon and Black, 1989); fowlpox virus, AJ581527 (Laidlaw and Skinner, 2004); swinepox virus, AF410153 (Afonso et al., 2002); vaccinia virus, AY243312 (Baroudy et al., 1982); crocodilepox virus; DQ356948 (Afonso et al., 2006).

Type A

5'-TTTAT**ATTATGGATCAAATGTTTATAAAGCACAATTGCGAGAGTAACTTAATG GACGAGTGCATTTACGTCAAATTAATGTTGATCCATACGA**ATAAA-3'

Type B

5'-TTTAT**TCGTATGGATCAACATTAATTTGACGTAAATGCACTCGTCCATTAAGT TACTCTCGCAATTGTGCTTTATAAACATTTGATCCATAAT**ATAAA-3'

Fig. 4. Sequences of hairpin loop regions of the ACEV genome. Two types of sequence (type A, type B) that consist of inverted repeat sequences each other were obtained from each genome terminus. Bold text indicates a hairpin loop sequence, and the 5-nucleotide sequences adjacent to the hairpin loop are in normal font.

ChPVs and ACEV, as is the case for MSEV (Afonso et al., 1999; this study), given that ACEV contained no ORFs for transcription/RNA modification that were not detected in ChPVs.

DNA replication/repair

In addition to the ChPV homologs present (Table 2), ACEV contains ORFs for exonuclease RNase T and DNA polymerase III (ACV017), NAD^+ dependent DNA ligase (ACV91), reverse transcriptase (ACV153), ATP-dependent DNA ligase (ACV174) and DNA polymerase beta/AP endonuclease (ACV212). AMEV has homologs of all of the ACEV ORFs in this category, but MSEV is missing those of dUTPase (ACV048) and thymidine kinase (ACV210). The presence of the genes for both NAD^+ dependent DNA ligase and ATP-dependent DNA ligase was a distinguishing feature of EVs including ACEV; almost all viral species, including both ChPVs and insect viruses (except for some kind of iridescent virus), and eukaryotic organisms encode only ATP-dependent DNA ligases (Afonso et al., 1999; Lasko et al., 1990; Lindahl and Barnes, 1992). This pattern suggests that the NAD^+ dependent DNA ligase gene in EVs has a prokaryotic origin (Afonso et al., 1999). The thymidine kinase gene is involved in nucleotide metabolism, and is present in other *Betaentomopoxvirus* species (e.g., CBEV, *Choristoneura fumiferana* EV, AMEV, and MySEV) (Gruidl et al., 1992; Lytvyn et al., 1992; Thézé et al., 2013) but not in MSEV (Afonso et al., 1999). This finding suggests that DNA replication of *Alphaentomopoxvirus* and *Betaentomopoxvirus* members is rather different from that of EVs infecting grasshoppers, including MSEV. MSEV replication may be heavily dependent on host cell nucleotide biosynthesis (Afonso et al., 1999).

Other functions

In addition to various ChPV homologs (Table 2), ACEV carries ORFs similar to fusolin (ACV034), formyltetrahydrofolate synthetase (ACV047), nicotinamide riboside kinase 1 (ACV053), calcium binding protein (ACV057), metallophosphoesterase domain-containing protein

Table 2
Chordopoxvirus (ChPV) homologs in ACEV.

Function category	ACEV ORF	Length (aa)[a]	VACV ORF[b]	Length (aa)	Gene name and/or function (or structure)
Structure/morphogenesis	ACV037	126	L5R	114	membrane protein
	ACV040	112	A21L	117	membrane protein involved in fusion and entry
	ACV042	137	J5L	133	membrane protein
	ACV056	333	G9R	340	poxvirus myristoyl protein
	ACV066	384	A16L	378	myristylated membrane protein
	ACV071	142	A28L	146	membrane protein
	ACV096	249	L4R	251	ssDNA/dsDNA binding protein
	ACV097	245	L1R	250	myristylated membrane protein
	ACV104	75	A9L	99	virion membrane protein (MP)
	ACV108	451	I7L	423	virion core cysteine protease
	ACV109	425	G5R	434	FEN1-like nuclease
	ACV112	285	H3L	324	glycosyl transferase
	ACV131	665	A3L	644	core protein
	ACV135	241	A32L	300	ATP/GTP-binding protein
	ACV137	463	F10L	439	serine/threonine protein kinase
	ACV142	327	A11R	318	virion membrane formation
	ACV143	1170	A10L	891	precursor p4a of core protein
	ACV163	528	E6R	567	core protein
	ACV178	201	H2R	189	viral membrane protein
	ACV184	609	G1L	591	vaccinia G1L metalloprotease
	ACV193	566	D13L	551	rifampicin resistance protein
	ACV218	237	F9L	212	S–S bond formation pathway protein
Transcription/RNA modification	ACV044	180	D7R	161	RNA polymerase subunit
	ACV049	279	D12L	287	mRNA capping enzyme small subunit
	ACV074	646	D11L	631	nucleoside triphosphate phosphohydrolase I
	ACV085	1190	A24R	1164	DNA-directed RNA polymerase subunit RPO132
	ACV093	260	A1L	150	late transcription factor VLTF-2
	ACV100	235	A5R	164	DNA-dependent RNA polymerase subunit rpo19
	ACV110	672	D6R	637	early transcription factor small subunit
	ACV138	264	List174	564	BTB/POZ domain-containing protein
	ACV140	891	D1R	844	mRNA capping enzyme large subunit
	ACV144	346	L3L	350	early gene transcription related protein
	ACV146	567	E1L	479	poly(A) polymerase large subunit
	ACV148	484	A18R	493	DNA helicase, transcriptional elongation
	ACV150	814	H4L	795	RNA polymerase associated transcriptional specificity factor
	ACV152	334	H6R	314	DNA topoisomerase type I
	ACV154	339	A29L	305	DNA-dependent RNA polymerase RPO35
	ACV170	711	I8R	676	RNA helicase NPH-II
	ACV180	273	J3R	333	poly(A) polymerase small subunit
	ACV187	761	A7L	710	early transcription factor large subunit
	ACV190	349	A23R	382	DNA-directed RNA polymerase beta chain
	ACV192	222	A2L	224	late transcription factor VLTF-3
	ACV230	1318	J6R	1286	RNA polymerase largest subunit
DNA replication/repair	ACV028	269	B1R	300	ser/thr protein kinase
	ACV039	245	D4R	218	uracil DNA glycosylase UNG
	ACV048	316	F2L	147	dUTPase
	ACV103	159	A22R	176	holliday junction resolvase
	ACV115	410	B1R	300	ser/thr protein kinase
	ACV147	283	D10R	248	NTP pyrophosphohydrolase mutT motif
	ACV164	843	D5R	785	NTPase
	ACV210	179	J2R	177	thymidine kinase
	ACV226	1114	E9L	1006	DNA polymerase
Other functions	ACV068	180	E10R	95	sulfhydryl oxidase
	ACV086	287	DUKE-205	222	serpin
	ACV092	147	A45R	125	superoxide dismutase
	ACV160	77	CNPV096[c]	85	ubiquitin/ribosomal protein
	ACV183	289	M5L[d]	288	lipase
	ACV196	114	Vegf-e[e]	132	vascular endothelial growth factor
	ACV228	208	G6R	165	in vivo virulence related protein
Unknown	ACV003	394	CNPV169[c]	332	
	ACV004	383	CNPV227[c]	359	
	ACV063	305	CNPV165	346	
	ACV179	218	B2R	219	
	ACV201	305	CNPV227[c]	359	
	ACV203	227	CNPV168[c]	358	
	ACV220	167	CNPV142[c]	321	

[a] Amino acids.
[b] Vaccinia virus homolog. Where no VACV homolog exists, other ChPV homolog is shown.
[c] ORF of canary poxvirus.
[d] ORF of cowpox virus.
[e] OFR of orf virus.

(ACV061), fusolin-like protein (ACV069, described in detail in another subsection below), translation elongation factor eEF-3 like (ACV076), protein kinase domain containing protein (ACV087), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (ACV94), cation transport ATPase (ACV101), translation initiation factor elF-4E (ACV116), inhibitor of apoptosis protein (ACV117, described in detail in another subsection below), G-protein coupled receptor Mth2-like rotein (ACV120), serin/threonine kinase like protein (ACV122), tenascin (ACV123), capsid protein VP1-like (ACV172), putative protein phosphatase 2C (ACV176), and Peritrophin-A domain containing protein (ACV205).

Comparison of ORF content to other EVs

Among the 241 ORFs in ACEV in which each duplicate copy gene in the ITRs is counted as one ORF, 152 ORFs showed similarity to those in other EVs (Table 1), corresponding to a rate of 63.1%. In detail, for AMEV and MSEV, 137 ORFs were similar between ACEV and AMEV, and 128 were similar between ACEV and MSEV. This finding implies that ACEV is taxonomically more closely related to AMEV than MSEV. This assumption is supported by the results of a phylogenetic analysis of EVs based on spheroidin (Perera et al., 2010; Hernández-Crespo et al., 2000).

Among the 22 ORFs in the ACEV ITR region, the rate of detection of similarities to other EVs was 77.3%. Therefore, whether the ORF content in ITR regions has higher diversity than does that in the central region is unclear; this result does not appear to be consistent with the previously reported trend that the majority of novel genes are generally found within poxviral ITRs or toward the genomic termini (Bawden et al., 2000). However, all 49 core genes in poxviruses are contained in the central region bounded by ITRs in ACEV, consistent with the other poxviruses (Perera et al., 2010).

On the other hand, some ORFs that are present in other EV (s) but are missing in ACEV include dual specificity phosphatase (AMV078), CPD photolyase (AMV025 and MSV235), protein tyrosine phosphatase (AMV246) and FALPE (AMV032). ChPVs harbor three of these four proteins; only FALPE (Alaoui-Ismaili and Richardson, 1996; Van Oers and Vlak, 1997) which forms cytoplasmic fibrils in infected cells, which are associated with spheroids, is missing from ChPVs.

EV gene families

Eight EV gene families have been reported; leucine-rich repeat, alanine–leucine–isoleucine (ALI) motif subgroup 1, ALI motif subgroup 2, tryptophan repeat, methionine–threonine–glycine (MTG), AMV176, KilA-N domain-containing proteins, and serine–cysteine–glycine (SCG) (Perera et al., 2010). None of these genes has been functionally characterized, although KilA-N domain-containing proteins in DNA viruses may play a role in transcription and/or DNA replication (Perera et al., 2010). Among the eight families, ALI motif family subgroup 1, AMV176 family, and SCG family were not found in ACEV (Table 1). As for other EVs such as AMEV and MSEV, the SCG family is missing from AMEV, and MSEV lacks the AMV176 and KilA-N domain-containing protein families. ORFs belonging to the leucine-rich repeat family are ACV018, ACV019, and ACV211, and those that belong to ALI motif family subgroup 2 are ACV003, ACV004, ACV005, ACV006, ACV011, ACV032, ACV197, and ACV198. ACV177 is a member of the tryptophan repeat family. ORFs that belong to the MTG family are ACV001, ACV002, ACV012, ACV013, ACV014, ACV015, and ACV016. ORFs that belong to the KilA-N domain-containing protein family are ACV063, ACV201, ACV202, ACV203, and ACV220. ACEV contains more ALI motif family subgroup 2 and MTG family ORFs (8 and 7, respectively) than do AMEV (4 and 3, respectively) and MSEV (3 and 4, respectively).

ChPV homologs within the ACEV genome

ACEV ORFs that show similarity to those in ChPVs are shown in Table 2. Among the ORFs similar between ACEV and ChPVs, BTB/POZ domain-containing protein (ACV138) and serpin (ACV086; discussed in detail in another subsection below) are not present in other EVs.

ACV160 corresponds to a ubiquitin/ribosomal protein that is one of the 49 core poxviral proteins. This protein not only plays a role in ubiquitin-mediated protein degradation in eukaryotes but is involved in numerous non-proteolytic functions, including viral budding (Reilly and Guarino, 1996; Hochstrasser, 2009). ACV160 ubiquitin showed an amino acid (aa) identity of up to 89% to a number of eukaryotic ubiquitins, and these high similarities are consistent with those for AMEV and MSEV ubiquitins. However, baculovirus ubiquitins showed somewhat lower identities (approximately 75%) to those in eukaryotes (Reilly and Guarino, 1996).

Inhibitors of apoptosis

Apoptosis is an important immune response against viral infections in host cells. The ACEV genome does not harbor *inhibitor of apoptosis* (*iap*) gene(s) that are typical in structure. ORFs in the IAP family are characterized by the presence of an inhibitor of apoptosis (IAP) repeat (BIR) domain, and many IAPs also contain a C-terminal RING domain (Salvesen and Duckett, 2002; Taylor and Barry, 2006). However, the ACEV genome does contain one ORF (ACV117) that contains a C-terminal BIR domain but not RING and another ORF (ACV182) that contains RING at its C-terminal but not BIR (Fig. 5); we confirmed the locations of these ORFs in the genome sequence by verifying the corresponding reads and reads around the ORFs of the plasmids of a template shotgun library, which had been generated by the sequencer (see Section 4). These two ORFs are distant from each other (Fig. 1); ACV117 is located from 97.3 to 98.4 kb downstream of the left terminus of the two DNA strand region of the genome, whereas ACV182 is located from 163.3 to 164.4 kb downstream of the left terminus. ACV117 also contains a DUF4419 motif (aa 22 through aa 290) and showed highest similarity to a protein in *Dictyostelium discoideum* (Dictyosteliida: Dictyosteliaceae) (blastp *E* value, 1e-36). However, ACV182 may not be involved in the apoptosis inhibition but in other functions because RING has been shown to be a domain for a wide range of functions. Furthermore, P35 in baculovirus and its AMEV homolog, P33, represent one type of apoptosis inhibition protein (Means et al., 2007), but this ORF was not present in ACEV.

Interestingly, the ACEV genome has an ORF containing a serpin domain (ACV086), which was detected by using blastp, PSI-BLAST and InterProScan programs (Fig. 5). Serpin-domain-containing genes are absent from the AMEV, MSEV, and other insect viruses

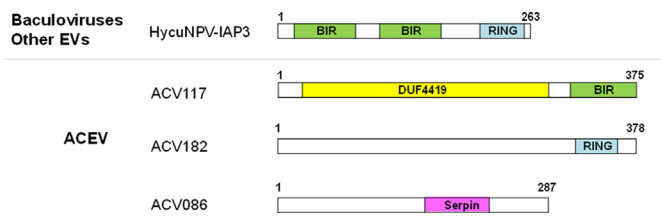

Fig. 5. Basic structure of three potential apoptosis inhibition-related ORFs in the ACEV genome. The names of ORFs are listed to the left of the diagrams. HycuNPV-IAP3, an inhibitor of apoptosis (Hycu-IAP3) in *Hyphantria cunea* nucleopolyhedrovirus (Ikeda et al., 2004); BIR, baculoviral inhibition of apoptosis protein repeat domain; RING, really interesting new gene domain; DUF4419, domain of unknown function (DUF4419); Serpin, the domain of serine protease inhibitor.

(Becker and Moyer, 2007; Perera et al., 2010). Serpins inhibit apoptosis and actually function in ChPVs (Taylor and Barry, 2006; Turner and Moyer, 1998). In analyses by using blastp and PSI-BLAST, ACV086 did not show similarity with ChPV serpin sequences. On the other hand, searches using DELTA-BLAST yielded good *E*-values (maximum, 4e-18) between ACV086 and both mammalian and ChPV serpin sequences. These combined results suggest that the ACEV ORF is not closely related to those of the mammalian and ChPV serpins.

The results of the present analysis indicate that ACEV does not harbor an IAP gene typical of those in other insect viruses because the two IAP-characteristic domains in ACEV lie on two different ORFs and that ACEV harbors a serpin gene found in ChPVs. Therefore the system of apoptosis inhibition in ACEV may be different from that in other insect viruses and ChPVs. If the above mentioned difference in status of IAP-associated domains from that in other insect viruses lead to differences in function from that in other insect viruses, the ORF for serpin in ACEV may compensate for the gap.

Another immune-response-related protein, superoxide dismutase (SOD), is functional in AMEV and might act to overcome innate immune responses in the insect gut (Becker et al., 2004). A homologous ORF was found in ACEV (ACV92) but not in MSEV. Furthermore, SOD of VACV and myxoma virus is non-functional.

Ubiquitin-conjugating enzyme

We identified an ORF (ACV161) similar to ubiquitin-conjugating enzyme E2 that was located adjacent to ACEV ubiquitin (ACV160) (Figs. 1 and 6). Ubiquitin-conjugating enzymes have not been found in viruses, except for ASF virus and some giant viruses such as *Megavirus chilensis* and *Acanthamoeba polyphaga mimivirus* (Rodriguez et al., 1992; Arslan et al., 2011; Legendre et al., 2011). In addition, we found by analysis using blastp and PSI-BLAST that ubiquitin-conjugating enzyme E2 in a moumouvirus and other species of megaviruses is included in GenBank. The ubiquitin-conjugating enzyme E2 is necessary for ubiquitin-mediated protein degradation in eukaryotes (Taylor and Barry, 2006), which is one of the most essential functions in organisms, and a similar system was found in a bacterium (Pearce et al., 2008; Striebel et al., 2009). These findings suggest a virus-host-related function, even if the ORF is not involved in ubiquitin-mediated protein degradation in ACEV.

ATP-binding cassette protein

ACV025 showed high similarity to ATP-binding cassette proteins (ABC transporter). The most similar protein was an ATP-binding cassette protein in the beetle, *Dendroctonus ponderosae* (blastp *E* value, 0; 31% aa identity from aa 5 through 1,504 of ACV025), and other proteins showing high similarity were ATP-binding cassette proteins of insects. Other than AMEV (Bawden et al., 2000), ACEV is the sole virus in which an ATP-binding cassette protein has been documented, although this protein should be present in all organisms. Interestingly, the ACEV ATP-binding cassette protein showed relatively low similarity to that (AMV130) in AMEV; proteins indicated by blastp with default parameters as being similar to ACV025 did not include the ATP-binding cassette protein of AMEV. These findings suggest that the ancestor of the EVs obtained the gene from insect(s), perhaps its host. ACV025 is the largest ORF (1506 aa) in ACEV, as is true for AMEV (1384 aa). The ACEV putative protein consists of TMS–ABC–TMS–ABC domains, and this arrangement is the same as that of the AMEV protein. ATP-binding cassette proteins function as membrane translocators, but some act as ion channels, ion channel regulators, and sensing proteins (Bawden et al., 2000). It is of interest whether the EV ATP-binding cassette protein is functional, and if so, what its function is.

Fusolin-like ORF

The ACEV genome harbored two fusolin-group genes, one of which was known previously (ACV034; Mitsuhashi et al., 1997; Takemoto et al., 2008); the novel ORF (ACV069) showed 32% identity at the aa level in the region of aa 14 to 243 to ACV034 by blastp analysis (Fig. 7a). In addition, blastp analysis of the novel ORF revealed a chitin-binding domain 3, which has been conserved in fusolin-group genes. It has not been reported that insect viruses have multiple fusolin-GP37 group genes per genome (Perera et al., 2010; Arif, 1995). Analysis of the hydrophilicity and hydrophobicity of ACV069 revealed a distributed pattern of hydropathy plots that are similar to those of the known fusolins; ACV069 contains a short N-terminal highly hydrophobic region and a relatively long C-terminal hydrophilic region (Fig. 7b). These structural features of ACV069 imply that it is a signal peptide and disrupts the host peritrophic membrane as does the known fusolin, ACV034 (Mitsuhashi et al., 2007; Mitsuhashi and Miyamoto, 2003).

```
ACV161       ---MDYKQRIMKEISNLQNNKLENIIINIGDTINIIH--FILIGYKDTVFENGYYYCKLF  55
Tetrahymena  MKFSDTVNRINKEYQRLQKNPVENVL-AIPDPKNMFQWHFCIYGLVDCPFEGGIYHGILS  59
                 *   ** **   ** *  **    * *  *     *   *  *  ** * *   *

ACV161       LN-KYPVTAPDIMMITPNGLFKPNTKLCIDGLTSHHNETWAITTKLDKILIAFQSFMNDT  114
Tetrahymena  LPPEYPMKPPSIKILTPNGRFKEGTNICT-SFTNYHPESWQLTWNIEKMLIAMISFMND-  117
             *   **   * *   **** **  *  *    *  * * *  *    * ***  *****

ACV161       IEEEFIGKIHTTITEKKILSKSSIKNNL-ENTEFVKTFKDTDIYRKMIETYNKF------  167
Tetrahymena  -NDPSAGVVQTSESEKRRLAKKSIIWNIKNDEEFVRLFKPYYKQLNIDPSLFTDPQKLKE  176
                   *   *   **  * * **  *     ***  **

ACV161       ---------------------------------------------   167
Tetrahymena  YEEQMQFQQSTTDHEKKVRNFEKVLFFGASIFLVMCSYLYMKSLK   221
```

Fig. 6. Alignment of the amino acid sequence of the putative ubiquitin-conjugating enzyme E2 of ACEV with that of *Tetrahymena thermophila* (Ciliophora). The putative ubiquitin-conjugating enzyme E2 (ACV161) of ACEV was compared with that of a species in *Tetrahymena*, which was the most similar to ACV161 according to blastp. GENETYX MAC ver.15 (GENETYX Corporation, Tokyo, Japan) was used to perform the analysis. Asterisks indicate identity between the two amino acids. GenBank accession number: *Tetrahymena thermophila*, XP_001021849.1

A

```
ACV069          -----------------MKRVIGLYMLINNVFCGNYISFPIPRQLKCNNIENRMCEKSYN    43
ACEV fusolin    -----MLKLYILLILLGYVYGHGYVTFPIARQRRCNVQGGFWWPPEGTNIPDPMCRAAYQ    55
HAEV fusolin    MNKFYYICIYIN-ILYVCVSGHGYMTFPIARQRRCSVRGGQWWPPNGDGITDTMCRAAYQ    59
                          ..  ..   . ..*. .......... . .. .... . .* . **...*.

ACV069          YIYDKIYDSNGNEHIAHDVANKMFQSYDSYMKYAGPKYYSLTNIKNNIIPNNLCSA----    99
ACEV fusolin    YVFNKVLSEGGSTSQAASAAQYMFQQDNEYAALAGPNFRDICWIKEQVVPDYLCAAGADT   115
HAEV fusolin    NVYNKVLNQYNDPQEAATAAQYMFQQDNEYAALAGPDYTNLCNLQQNVVPNNLCAAGADD   119
                ....*..   .    *. .*..***....*...*** .  ..... ...*..**.*...

ACV069          ---NTY-DKSGMDLIGDWKPNIL------DDSGYLEIALCSDYIDDYSYIEVYITHSNFN   149
ACEV fusolin    WRIRPFGDKTGMDIVGSWPPTVIPLENNFVNTIPIELEFCPTAIHEPSYFEVYVTTPEFN   175
HAEV fusolin    WDVVPFGDKSGMDLPGNWVPTVIPLDSNHQSSVALELEFCPTAVHDPSYYEVYITNSGFN   179
                .   ...**.***. * * *.....  .   .  .*...*.......** ***.* . **

ACV069          VEKHKIEWNILYLIYNNSVSLSANNNNYNCKND-YIYKFIIDVPYRNSQFLLYIRHQLAN   208
ACEV fusolin    VYRDKVTWPLLELVFNSTVPLVNRRADSLCTANARVYRMIVPVPYRQTQFVIYVRWQRID   235
HAEV fusolin    VHTDNVVWGNLELIFNDTVPLRPKSSTSTCNANPNVYRFTVSIPVRPAQFVLYVRWQRID   239
                *  ... *  *.*..* .*.*      . * ..  .*.... .*.*  **..*.*.*...

ACV069          SNGLGTYNCIDMIFKTHEHNCCKHDCCKHDCKHDCCKHDCCKHDCKHDCCKHDCKHDCCK   268
ACEV fusolin    PVGEGFYNCVDAVFANRPGPDPEDMIPPP-PAEEDCDYT-DSQGNRYCP-F-STFSNSYY   291
HAEV fusolin    PVGEGFYNCVDMAFDYAAGPSEEDVIYPDYEAPGQNAYTCHANRNKYGGNYENTIDEDKY   299
                ..*.*.***.*. *    ..  .. . .   .  . ...    ...     .    .

ACV069          HDCKHDCCKHDCKHDCCKHDCCKHDCKHDCCKHDCKHDCCKHDCKHDCCKHDCKHDCCKH   328
ACEV fusolin    SNREHYNHKHDYNRYSNY----YDHNTYTYNYGKYNK-NKYSDTTYNNRAWEIAYAGYTE   346
HAEV fusolin    QAQLDESIKSRYDKYSRHKGGKFGQKQCNGNKHHYNKYTKYYNQNYKNNKNY--------   351
                    .   *...  ..             ... ...  .. .  . . .

ACV069          DCKHDCCKHDCCKHDCCKHDCCKHDCCKHDCCKHDCKH   366
ACEV fusolin    DHTGLNRQTECDGISRFCVSTVRNRVY-----------   373
HAEV fusolin    ----------------------------------------   351
                    .         .
```

B

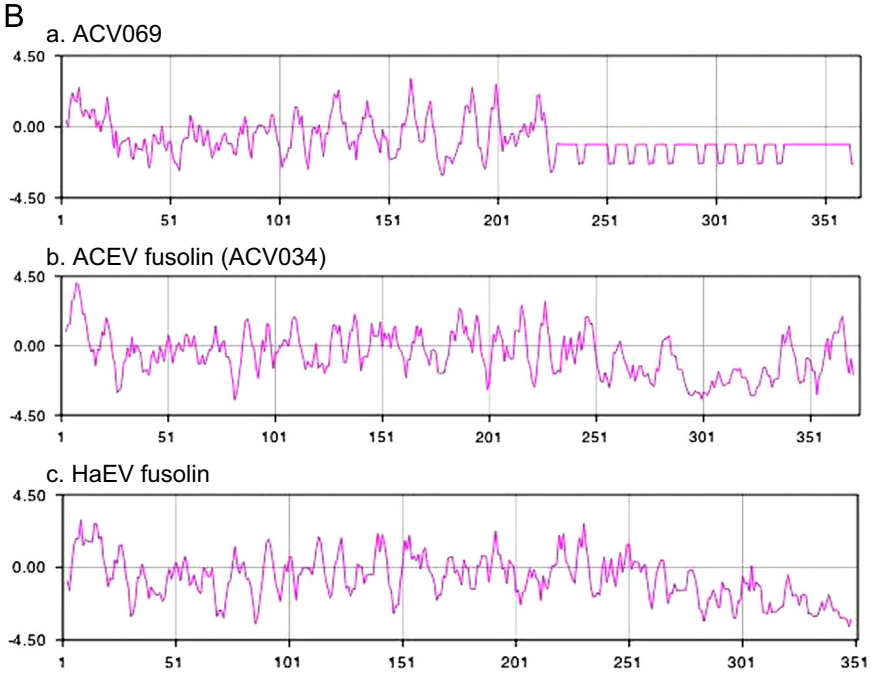

Fig. 7. Analyses of an ORF for a putative fusolin-like protein in the ACEV genome. (A) Alignment of the amino acid sequence of ACV069 with that of ACV034 (known ACEV fusolin) and *Heliothis armigera* EV fusolin. GENETYX MAC ver.15 (GENETYX Corporation, Tokyo, Japan) was used to perform the analysis. Asterisks indicate identity among the three amino acids, and dots indicate amino acids that are conserved more than 50% among the three amino acids. Five conserved regions among fusolin/GP37 that have been identified by Vialard et al. (1990) are underlined. HAEV fusolin, *Heliothis armigera* EV fusolin (Dall et al., 1993; GenBank accession number, L08077). (B) Comparison of hydropathy plots among ACV069, ACV034 and L08077. The analysis was performed by using Genetyx-Mac ver.15 (GENETYX Corporation, Tokyo) with default parameters; the Kyte and Doolittle algorithm (1982) was used with an average of five values.

Nucleotide sequence accession numbers

The GenBank accession numbers of sequences are as follows: ACEV genome (DNA double-stranded region), AP013055; left-terminal hairpin loops, AB787183 and AB787184; and right-terminal hairpin loops, AB787185 and AB787186.

Conclusion

To discuss EVs at the whole-genome level and obtain a better understanding of EV evolution requires the complete sequence of at least one species in both *Alphaentomopoxvirus* and *Gammaentomopoxvirus*. Furthermore, because of the unavailability of these complete genomic sequences, proteome and gene-function analyses of EVs have lagged behind those of ChPVs and baculoviruses. The results of the present study thus have merit in regard to both of these issues. Our present study shows genetically a rather distant relationship between ACEV and the orthopteran EV MSEV and confirms the necessity of creating a new genus for orthopteran EVs. The first elucidation of the terminus at each end of the genome provides a fundamental insight regarding the mode of DNA replication in EVs. The results of the present study suggest various features in EV-host interactions, particularly specific features in EV-scarab beetle host interactions, for example, apoptosis-associated aspects and the presence of a fusolin-like gene. The specificities in EV-scarab beetle host interactions may derive from the circumstances around the larval habitat underground in addition to coleopteran specificity. Future studies to obtain a better understanding of EV-insect host interactions are expected on the basis of the ACEV genome information, and these results will increase our understanding of the coevolution of EVs and their hosts.

Compared with baculoviruses, EVs have been less well-studied regarding their use as pest control agents but they offer some merits in this application. Baculoviruses do not infect all pest species; for example, almost all beetle pest species are unsusceptible to baculoviruses, but a number of beetle species are hosts of EVs. Therefore, EVs are potential control agents for pests that are resistant to baculoviruses. In addition, fusolins of EVs are potential co-agents for viral insecticides and *B. thuringiensis* formulations because of these proteins' ability to enhance viral infectivity and insecticidal activity of *B. thuringiensis*. A better understanding of EV-host interactions and EVs themselves at the molecular level will facilitate the design and development of highly efficient biocontrol strategies and technologies.

Materials and methods

Virus and DNA preparation

The ACEV isolate (strain CV6M) used in the present study was obtained through an in vivo cloning process (Smith and Crook, 1988) as previously described (Mitsuhashi et al., 2014) and consequently is a clone or nearly clonal.

The strain was propagated in *A. cuprea* larvae in our laboratory, and the DNA was extracted from *A. cuprea* cadavers according to the methods of Mitsuhashi et al. (1997, 2007).

DNA cloning and sequencing

After ACEV DNA was randomly fragmented by using Hydro-Shear (GeneMachines, San Carlos, CA), the fragments were blunt-ended by using a Mighty Cloning Kit (TakaRa Bio, Shiga, Japan). Fragments of approximately 2–2.5 kb were gel-purified by using a MinElute Gel Extraction Kit (QIAGEN, Tokyo, Japan) and cloned into the dephosphorylated *Hin*c II site of pUC118 (TakaRa Bio). Recombinant plasmids were transformed into *E. coli* DH10B cells by electroporation (Gene Pulser, BIORAD, Hercules, CA).

Plasmids were purified from transformed cells by using an alkaline-SDS method and were sequenced using sequencing primers M13-47 and R-8, an ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA), and the 3730 × l DNA Analyzer (Applied Biosystems).

Sequence assembly

Chromatograms were base-called by using the software, Phred ver. 0.020425.c (CodonCode, Dedham, MA), and vector-associated and low-quality-value sequences in reads were masked by using the software PFP ver. 2.6.2 (Paracel, Pasadena, CA). The resultant sequences were assembled into contigs by using the software CAP4 ver. 2.6.2 (Paracel).

Additional sequencing and assembly

A 225-kb contig was obtained from the above-mentioned procedure, and this contig-sequence represented on average a twelve-fold redundancy at each base position. The sequences of five regions (approximate lengths: 300, 600, 300, 3300 and 600 bps) for which the depth of reads was lower than that of the other regions were confirmed by direct-sequencing of PCR products obtained by using (as template) plasmids in the shotgun library.

Early in the present study, we found that two 2370-bp regions (nucleotide (nt) 20,499 to nt 22,868 and nt 222,630 to the terminus (nt 224,999)) in the 225-kb contig are IRs of each other. We therefore speculated that a region from the left terminus (nt 1) to nt 22,868 is an inverted terminal repeat (ITR) and that most of its right counterpart (right terminal ITR) was missing from the contig sequence. It was supposed that the lack of most of the putative right ITR was due to the software's misassembly of right ITR reads into the assembly for the 20.5-kb left ITR region. In reads of the 20.5-kb left ITR, no significant hetero-sequence was observed at each nucleotide position.

We therefore used the following method to obtain the possible missing right ITR sequence and to confirm the left ITR sequence. Virions purified from the spheroids according to Mitsuhashi et al. (1997) were embedded in 0.75% agarose (SeaPlaque GTG, TakaRa Bio) gel in molds. Each of the resulting plugs was twice treated with TE buffer (pH 8.0) and then with proteinase K (final concentration, 50 μg/ml) and 0.1% SDS overnight at 50 °C. Plugs then were treated and extracted DNA was digested with *Sac*I, according to Sambrook and Russell (2001). The plugs were loaded into wells of a gel of pulsed-field gel electrophoresis and electrophoresis was performed by using the CHEF Mapper System (BIO-RAD, Alfred Nobel Drive, Hercules, CA) according to the manufacturer's manual. The separated DNA fragments (left-most fragment of the genome, 30 kb; right-most fragment, 37 kb) were purified from the gel by using agarase (β-Agarase, TakaRa Bio) and then were treated with S1-nuclease to eliminate hairpin structures. Shotgun library of each terminal fragment was then constructed, inserts within plasmids were sequenced, and contigs were formed, as described above. The assembly of the reads from the shotgun library of the *Sac*I digested left terminal fragment yielded fourteen contigs, and gaps (each shorter than 200 bp) were filled by direct sequencing of PCR products generated by using the 30-kb left terminal fragment as a template. Assembly of the reads from the right fragment yielded a continuous 37-kb sequence that corresponded to the expected 37-kb sequence containing the entire right ITR.

Sequencing of regions external to the double-stranded region of the genome

To sequence the regions linked to the termini of the two DNA strands, we performed nested PCR. Briefly, the first PCR was performed by using a TaKaRa LA PCR Kit (TakaRa Bio) and a single primer (sequence; 5′-TTGATTAATCCATGTAAACCCGT-3′) that corresponded to a region just beyond a hairpin loop in the DNA double-stranded region of the 30-kb left or 37-kb right terminal fragment obtained according to the above-mentioned method with the exception of the omission of treatment with S1-nuclease. The second PCR was performed by using a single primer (5′-TGATTAATCCATGTAAACCCGTA-3′) and the same PCR kit as for the first PCR. The PCR products obtained (ca. 280 bp) were cloned into pT7 Blue vector (Novagen, Madison, Wisconsin), and the plasmids were introduced into *E. coli* HST08 by electroporation. Purified plasmids of twenty *E. coli* clones for each terminus of the genome were sequenced by using the primer M13-47 as mentioned above. This process yielded two types of sequence from the right-most (37 kb) fragment as a PCR template, which correspond to the two types of hairpin loops, and only one type of sequence (that is, the same as one of those from the right-most fragment) from the left-most fragment (30 kb) as a template.

We therefore performed subsequent PCRs (referred to as "Subsequent PCRs") to confirm the complete sequences of the hairpin loop regions. Based on the two types of sequence obtained from the nested PCRs, we synthesized four primers (5′-GTGACATAAATGCACTCGTCCA-3′; 5′-GTCAAATTAATGTTGATCCATACG-3′; 5′-CTGAAGCACAATTGCGAGAGT-3′; and 5′-TGTGCTTTATAAACATTTGATCCA-3′), and then carried out PCR by using each of the four primers together with a primer (5′-(C)TGAA(or C)AA(or C)TGTAATGTAATGTATTAAC-3′) that corresponded to the sequence of the double-stranded region and an extremely high-fidelity PCR kit (PrimeSTAR Max DNA Polymerase, TaKaRa Bio). The PCR products obtained were cloned to pUC118 vector, and the resultant plasmids were introduced into *E. coli* HST08 by electroporation. Purified plasmids of 10 *E. coli* clones from each of the four kinds of PCR product were sequenced by using primers M13-47 and RV-M as mentioned above.

Confirmation of genome organization

The organization of the ACEV genome determined in the present study was checked in several ways. First, we checked on the restriction site maps (physical maps) of the ACEV genome that were constructed based on the final consensus sequence. Using ACEV genomic DNA as a template, we performed PCRs that were designed to amplify each "putative" region on the final consensus sequence, which consisted of a *Hin*dIII, *Bam*HI or *Pst*I site and the regions just upstream and downstream of the target restriction site and directly sequenced the PCR products. We also compared the sizes of the fragments generated by the complete digestion of ACEV genomic DNA with *Hin*dIII, *Bam*HI or *Pst*I followed by 0.4% agarose gel electrophoresis with those of putative restriction fragments. The results were that each sequence of the PCR products was identical to that of each corresponding region on the final consensus sequence and that the sizes of the electrophoresed restriction fragments were compatible with those of the putative corresponding "fragments" based on the final consensus sequence (data not shown).

Second, we performed pulsed-field gel electrophoresis of the whole genomic DNA and fragments digested with the enzyme *Kpn*I, *Pme*I, *Sac*I, *Sac*II or *Sma*I and compared their observed sizes, taking into account the particular migration characteristics of AT-rich DNA fragments in pulsed-field gel electrophoresis (Maniloff, 1989; Hall and Hink, 1990), with the predicted sizes of the final (full-length) consensus sequence and putative "restriction fragments". Observed sizes of the whole genome and the restriction fragments were compatible with those from the final consensus sequences (data not shown).

Third, we entrusted ACEV genome sequencing by using a 454/Roche GS FLX pyrosequencing system to Beckman Coulter Genomics (Danvers, MA). The 222-kb sequence that resulted from pyrosequencing corresponded to an inner region of the final consensus sequence generated by using the Sanger method and contained the full poxvirus central region bounded by the two ITRs; the sequence was identical to that with the Sanger method except that pyrosequencing led to a single-nucleotide gap in each of a few homopolymer regions.

Analysis

ORFs in ACEV were identified through a two-step process. The first step was to select all candidate ORFs that were both methionine-initiated and greater than 50 amino acids by using a program for 6-frame ORF detection (Genaris, Yokohama, Japan) that is based on the work of Stajich et al. (2002). From this pool of the candidates, we selected ORFs that had no or minimum overlap with other ORFs, with consideration of the status showing the codon skew and codon bias rules of true ORFs in poxviruses indicated by Da Silva and Upton (2005) and favorable results in similarity and domain research performed by using blastp and PSI-BLAST in NCBI, and with InterProScan (Zdobnov and Apweiler, 2001).

Acknowledgments

We thank Ayako Sakairi of our institute for her support in our experiments and Masahiko Muraji and Yuichi Katayose of our institute for their advice on the present study. We also thank Bernard Moss of the National Institutes of Health, USA, for his advice on the present study. This work was supported in part by JSPS KAKENHI 21380044.

References

Afonso, C.L., Tulman, E.R., Delhon, G., Lu, Z., Viljoen, G.J., Wallace, D.B., Kutish, G.F., Rock, D.L., 2006. Genome of crocodilepox virus. J. Virol. 80, 4978–4991.
Afonso, C.L., Tulman, E.R., Lu, Z., Oma, E., Kutish, G.F., Rock, D.L., 1999. The genome of *Melanoplus sanguinipes* entomopoxvirus. J. Virol. 73, 533–552.
Afonso, C.L., Tulman, E.R., Lu, Z., Zsak, L., Kutish, G.F., Rock, D.L., 2000. The genome of fowlpox virus. J. Virol. 74, 3815–3831.
Afonso, C.L., Tulman, E.R., Lu, Z., Zsak, L., Osorio, F.A., Balinsky, C., Kutish, G.F., Rock, D.L., 2002. The genome of swinepox virus. J. Virol. 76, 783–790.
Alaoui-Ismaili, M.H., Richardson, C.D., 1996. Identification and characterization of a filament associated protein encoded by the *Amsacta moorei* entomopoxvirus. J. Virol. 70, 2697–2705.
Arif, B.M., 1995. Recent advances in the molecular biology of entomopoxviruses. J. Gen. Virol. 76, 1–13.
Arslan, D., Legendre, M., Seltzer, V., Abergel, C., Claverie, J.M., 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc. Natl. Acad. Sci. USA 108, 17486–17491.
Bardwell, V.J., Treisman, R., 1994. The POZ domain: a conserved protein-protein interaction motif. Genes Dev. 8, 1664–1677.
Baroudy, B.M., Venkatesan, S., Moss, B., 1982. Incompletely base-paired flip-flop terminal loops link the two DNA strands of the vaccinia virus genome into one uninterrupted polynucleotide chain. Cell 28, 315–324.
Bawden, A.L., Glassberg, K.J., Diggans, J., Shaw, R., Farmerie, W., Moyer, R.W., 2000. Complete genomic sequence of the *Amsacta moorei* entomopoxvirus: analysis and comparison with other poxviruses. Virology 274, 120–139.
Becker, M.N., Greenleaf, W.B., Ostrov, D.A., Moyer, R.W., 2004. *Amsacta moorei* entomopoxvirus expresses an active superoxide dismutase. J. Virol. 78, 10265–10275.
Becker, M.N., Moyer, R.W., 2007. Subfamily Entomopoxvirinae. In: Mercer, A.A., Schmidt, O., Weber, O. (Eds.), Poxviruses. Birkhauser VerlagBasel, pp. 253–271.
Brunetti, C.R., Amano, H., Ueda, Y., Qin, J., Miyamura, T., Suzuki, T., Li, X., Barrett, J.W., McFadden, G., 2003. Complete genomic sequence and comparative

analysis of the tumorigenic poxvirus Yaba monkey tumor virus. J. Virol. 77, 13335–13347.
Cameron, C., Hota-Mitchell, S., Chen, L., Barrett, J., Cao, J.X., Macaulay, C., Willer, D., Evans, D., McFadden, G., 1999. The complete DNA sequence of myxoma virus. Virology 264, 298–318.
Culyba, M.J., Hwang, Y., Minkah, N., Bushman, F.D., 2009. DNA binding and cleavage by the fowlpox virus resolvase. J. Biol. Chem. 284, 1190–1201.
Da Silva, M., Upton, C., 2005. Using purine skews to predict genes in AT-rich poxviruses. BMC Genomics 6, 22.
Dall, D., Sriskantha, A., Vera, A., Lai-Fook, J., Symonds, T., 1993. A gene encoding a highly expressed spindle body protein of *Heliothis armigera* entomopoxvirus. J. Gen. Virol. 74, 1811–1818.
DeLange, A.M., Reddy, M., Scraba, D., Upton, C., McFadden, G., 1986. Replication and resolution of cloned poxvirus telomeres in vivo generates linear minichromosomes with intact viral hairpin termini. J. Virol. 59, 249–259.
Furuta, Y., Mitsuhashi, W., Kobayashi, J., Hayasaka, S., Imanishi, S., Chinzei, Y., Sato, M., 2001. Peroral infectivity of non-occluded viruses of *Bombyx mori* nucleopolyhedrovirus and polyhedrin-negative recombinant baculoviruses to silkworm larvae is drastically enhanced when administered with *Anomala cuprea* entomopoxvirus spindles. J. Gen. Virol. 82, 307–312.
Gershon, P.D., Black, D.N., 1989. A capripoxvirus pseudogene whose only intact homologs are in other poxvirus genomes. Virology 172, 350–354.
Goebel, S.J., Johnson, G.P., Perkus, M.E., Davis, S.W., Winslow, J.P., Paoletti, E., 1990. The complete DNA sequence of vaccinia virus. Virology 179, 247–266.
González, A., Talavera, A., Almendral, J.M., Vinuela, E., 1986. Hairpin loop structure of African swine fever virus DNA. Nucl. Acids Res. 14, 6835–6844.
Gruidl, M.E., Hall, R.L., Moyer, R.W., 1992. Mapping and molecular characterization of a functional thymidine kinase from *Amsacta moorei* entomopoxvirus. Virology 186, 507–516.
Hautaniemi, M., Ueda, N., Tuimala, J., Mercer, A.A., Lahdenperä, J., McInnes, C.J., 2010. The genome of Pseudocowpoxvirus: comparison of a reindeer isolate and a reference strain. J. Gen. Virol. 91, 1560–1576.
Hall, R.L., Hink, W.F., 1990. Physical mapping and field inversion gel electrophoresis of *Amsacta moorei* entomopoxvirus DNA. Arch. Virol. 110, 77–90.
Hernández-Crespo, P., Veyrunes, J.C., Cousserans, F., Bergoin, M., 2000. The spheroidin of an Entomopoxvirus isolated from the grasshopper *Anacridium aegyptium* (AaEPV) shares low homology with spheroidins from lepidopteran or coleopteran EPVs. Virus Res. 67, 203–213.
Hochstrasser, M., 2009. Origin and function of ubiquitin-like proteins. Nature 458, 422–429.
Ikeda, M., Yanagimoto, K., Kobayashi, M., 2004. Identification and functional analysis of *Hyphantria cunea* nucleopolyhedrovirus *iap* genes. Virology 321, 359–371.
Jongeneel, C.V., Bouvier, J., Bairoch, A., 1989. A unique signature identifies a family of zinc-dependent metallopeptidases. FEBS Lett. 242, 211–214.
King, L.A., Wilkinson, N., Miller, D.P., Marlow, S.A., 1998. Entomopoxviruses In: Miller, L.K., Ball, L.A. (Eds.), The Insect Viruses. Plenum Press, New York and London, pp. 1–29.
Kyte, J., Doolittle, R.F., 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.
Laidlaw, S.M., Skinner, M.A., 2004. Comparison of the genome sequence of FP9, an attenuated, tissue culture-adapted European strain of Fowlpox virus, with those of virulent American and European viruses. J. Gen. Virol. 85, 305–322.
Langridge, W.H.R., Roberts, D.W., 1977. Molecular weight of DNA from four entomopoxviruses determined by electron microscopy. J. Virol. 21, 301–308.
Lasko, D.D., Tomkinson, A.E., Lindahl, T., 1990. Eukaryotic DNA ligases. Mutat. Res. 236, 277–287.
Lefkowitz, E.J., Wang, C., Upton, C., 2006. Poxviruses: past, present and future. Virus Res. 117, 105–118.
Legendre, M., Santini, S., Rico, A., Abergel, C., Claverie, J.M., 2011. Breaking the1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. Virol. J. 8, 99.
Lindahl, T., Barnes, D.E., 1992. Mammalian DNA ligases. Annu. Rev. Biochem. 61, 251–281.
Lytvyn, V., Fortin, Y., Banville, M., Arif, B., Richardson, C., 1992. Comparison of the thymidine kinase genes from three entomopoxviruses. J. Gen. Virol. 73, 3235–3240.
Maniloff, J., 1989. Anomalous values of *Mycoplasma* genome sizes determined by pulse-field gel electrophoresis. Nucl. Acid Res. 17, 1268.
Massung, R.F., Liu, L.I., Qi, J., Knight, J.C., Yuran, T.E., Kerlavage, A.R., Parsons, J.M., Venter, J.C., Esposito, J.J., 1994. Analysis of the complete genome of smallpox variola major virus strain Bangladesh-1975. Virology 201, 215–240.
Means, J.C., Penabaz, T., Clem, R.J., 2007. Identification and functional characterization of *AMVp33*, a novel homolog of the baculovirus caspase inhibitor *p35* found in *Amsacta moorei* entomopoxvirus. Virology 358, 436–447.
Mitsuhashi, W., Asano, S., Miyamoto, K., Wada, S., 2014. Further research on the biological function of inclusion bodies of *Anomala cuprea* entomopoxvirus, with special reference to effect on the insecticidal activity of a *Bacillus thuringiensis* formulation. Pest Manag. Sci. 70, 46–54, http://dx.doi.org/10.1002/ps.3521.
Mitsuhashi, W., Furuta, Y., Sato, M., 1998a. The spindles of an entomopoxvirus of *Coleoptera* (*Anomala cuprea*) strongly enhance the infectivity of a nucleopolyhedrovirus in *Lepidoptera* (*Bombyx mori*). J. Invertebr. Pathol. 71, 186–188.
Mitsuhashi, W., Saito, H., Sato, M., Nakashima, N., Noda, H., 1998b. Complete nucleotide sequence of spheroidin gene of *Anomala cuprea* entomopoxvirus. Virus Res. 55, 61–69.
Mitsuhashi, W., Sato, M., Hirai, Y., 2000. Involvement of spindles of an entomopoxvirus (EPV) in infectivity of the EPVs to their host insect. Arch. Virol. 145, 1465–1471.
Mitsuhashi, W., Kawakita, H., Murakami, R., Takemoto, Y., Saiki, T., Miyamoto, K., Wada, S., 2007. Spindles of an entomopoxvirus facilitate its infection of the host insect by disrupting the peritrophic membrane. J. Virol. 81, 4235–4243.
Mitsuhashi, W., Miyamoto, K., 2003. Disintegration of the peritrophic membrane of silkworm larvae due to spindles of an entomopoxvirus. J. Invertebr. Pathol. 82, 34–40.
Mitsuhashi, W., Saito, H., Sato, M., 1997. Complete nucleotide sequence of the fusolin gene of an entomopoxvirus in the cupreous chafer, *Anomala cuprea* Hope (Coleoptera: Scarabaeidae). Insect Biochem. Mol. Biol. 27, 869–876.
Moss, B., 2001. Poxviridae: The Viruses and their Replication. In: Fields, B.N., Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J., Monath, T.P., Roizman, B., Straus, S.E. (Eds.), Virology. Lippincott-Raven Publishers, Philadelphia, pp. 2849–2883.
Moss, B., 2007. Poxviridae: The Viruses and their Replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology, Volume 2. Lippincott Williams & Wilkins, Philadelphia, pp. 2905–2946.
Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., Darwin, K.H., 2008. Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis. Science 322, 1104–1107.
Perera, S., Li, Z., Pavlik, L., Arif, B., 2010. Entomopoxviruses. In: Asgari, S., Johnson, K. N. (Eds.), Insect Virology. Caister Academic Press, Essex, UK, pp. 83–102.
Pickup, D.J., Ink, B.S., Parsons, B.L., Hu, W., Joklik, W.K., 1984. Spontaneus deletions and duplications of sequences in the genome of cowpox virus. Proc. Natl. Acad. Sci. USA 81, 6817–6821.
Rawlings, N.D., Barrett, A.J., 1995. Evolutionary families of metallopeptidases. Methods Enzymol. 248, 183–228.
Reilly, L.M., Guarino, L.A., 1996. The viral ubiquitin gene of *Autographa californica* nuclear polyhedrosis virus is not essential for viral replication. Virology 218, 243–247.
Rodriguez, J.M., Salas, M.L., Vinuela, E., 1992. Genes homologous to ubiquitin-conjugating proteins and eukaryotic transcription factor SII in African swine fever virus. Virology 186, 40–52.
Salvesen, G.S., Duckett, C.S., 2002. IAP proteins: blocking the road to death's door. Nat. Rev. Mol. Cell Biol. 3, 401–410.
Sambrook, L., Russell, D.W., 2001. Molecular Cloning, third ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.
Senkevich, T.G., Koonin, E.V., Bugert, J.J., Darai, G., Moss, B., 1997. The genome of molluscum contagiosum virus: analysis and comparison with other poxviruses. Virology 233, 19–42.
Shchelkunov, S.N., Blinov, V.M., Resenchuk, S.M., Totmenin, A.V., Olenina, L.V., Chirikova, G.B., Sandakhchiev, L.S., 1994. Analysis of the nucleotide sequence of 53 kbp from the right terminus of the genome of variola major virus strain India-1967. Virus Res. 34, 207–236.
Shchelkunov, S.N., Totmenin, A.V., 1995. Two types of deletions in orthopoxvirus genomes. Virus Genes 9, 231–245.
Shchelkunov, S.N., Totmenin, A.V., Sandakhchiev, L.S., 1996. Analysis of the nucleotide sequence of 23.8 kbp from the left terminus of the genome of variola major virus strain India-1967. Virus Res. 40, 169–183.
Skinner, M.A., Buller, R.M., Damon, I.K., Lefkowitz, E.J., McFadden, G., McLnnes, C.J., Mercer, A.A., Moyer, R.W., Upton, C., 2011. Family Poxviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy IX. Elsevier Inc., San Diego and London, pp. 291–309.
Smith, G.L., 2007. Genus Orthopoxvirus: Vaccinia virus. In: Mercer, A.M., Schmidt, A., Weber, O. (Eds.), Poxviruses. Birkhauser Verlag, Basel, pp. 1–45.
Smith, I.R.L., Crook, N., 1988. In vivo isolation of baculovirus genotypes. Virology 166, 240–244.
Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., Fuellen, G., Gilbert, J.G., Korf, I., Lapp, H., Lehväslaiho, H., Matsalla, C., Mungall, C.J., Osborne, B.I., Pocock, M.R., Schattner, P., Senger, M., Stein, L.D., Stupka, E., Wilkinson, M.D., Birney, E., 2002. The Bioperl toolkit: Perl modules for the life sciences. Genome Res. 12, 1611–1618.
Striebel, F., Imkamp, F., Sutter, M., Steiner, M., Mamedov, A., Weber-Ban, E., 2009. Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. Nat. Struct. Mol. Biol. 16, 647–651.
Takemoto, Y., Mitsuhashi, W., Murakami, R., Konishi, H., Miyamoto, K., 2008. The N-terminal region of an entomopoxvirus fusolin is essential for the enhancement of peroral infection, whereas the C-terminal region is eliminated in digestive juice. J. Virol. 82, 12406–12415.
Taylor, J.M., Barry, M., 2006. Near death experiences: poxvirus regulation of apoptotic death. Virology 344, 139–150.
Thézé, J., Takatsuka, J., Li, Z., Gallais, J., Doucer, D., Arif, B., Nakai, M., Herniou, E.A., 2013. New insights into the evolution of *Entomopoxvirinae* from the complete genome sequences of four entomopoxviruses infecting *Adoxophyes honmai*, *Choristoneura biennis*, *Choristoneura rosaceana* and *Mythimna separata*. J. Virol. 87, 7992–8003.
Turner, P.C., Moyer, R.W., 1998. Control of apoptosis by poxviruses. Semin. Virol. 8, 453–469.
Upton, C., Slack, S., Hunter, A.L., Ehlers, A., Roper, R.L., 2003. Poxvirus orthologous clusters: toward defining the minimum essential poxvirus genome. J. Virol. 77, 7590–7600.
Upton, C., Macen, J.L., Maranchuk, R.A., Delange, A.M., McFadden, G., 1988. Tumorigenic poxviruses: Fine analysis of the recombination junctions in

malignant rabbit fibroma virus, a recombinant between Shope fibroma virus and myxoma virus. Virology 166, 229–239.
Van Oers, M.M., Vlak, J.M., 1997. The baculovirus 10-kDa protein. J. Invertebr. Pathol. 70, 1–17.
Vialard, J.E., Yuen, L., Richardson, C.D., 1990. Identification and characterization of a baculovirus occlusion body glycoprotein which resembles spheroidin, an entomopoxvirus protein. J. Virol. 64, 5804–5811.
Wijonarko, A., Hukuhara, T., 1998. Detection of a virus enhancing factor in the spheroid, spindle, and virion of an entomopoxvirus. J. Invertebr. Pathol. 72, 82–86.
Willer, D.O., McFadden, G., Evans, D.H., 1999. The complete genome sequence of shope (rabbit) fibroma virus. Virology 264, 319–343.
Zdobnov, E.M., Apweiler, R., 2001. InterProScan-an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17, 847–848.