University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Virology Papers | Virology, Nebraska Center for

9-1-2006

Genome of Invertebrate Iridescent Virus Type 3 (Mosquito Iridescent Virus)

Gustavo A. Delhon
University of Nebraska - Lincoln, gdelhon3@Unl.edu

Edan R. Tulman
Plum Island Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Greenport, New York

Claudio L. Afonso
Plum Island Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Greenport, New York

Zhiqiang Lu
Plum Island Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Greenport, New York

James J. Becnel
CMAVE, Agricultural Research Service, U.S. Department of Agriculture, Gainesville, Florida

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/virologypub

Part of the Virology Commons

Delhon, Gustavo A.; Tulman, Edan R.; Afonso, Claudio L.; Lu, Zhiqiang; Becnel, James J.; Moser, Bettina A.; Kutish, Gerald F.; and Rock, Daniel L., "Genome of Invertebrate Iridescent Virus Type 3 (Mosquito Iridescent Virus)" (2006). *Virology Papers.* Paper 57.
http://digitalcommons.unl.edu/virologypub/57

This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Gustavo A. Delhon, Edan R. Tulman, Claudio L. Afonso, Zhiqiang Lu, James J. Becnel, Bettina A. Moser, Gerald F. Kutish, and Daniel L. Rock

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/virologypub/57

JOURNAL OF VIROLOGY, Sept. 2006, p. 8439–8449 Vol. 80, No. 17
0022-538X/06/$08.00+0 doi:10.1128/JVI.00464-06
Copyright © 2006, American Society for Microbiology. All Rights Reserved.

Genome of Invertebrate Iridescent Virus Type 3 (Mosquito Iridescent Virus)

Gustavo Delhon,[1,3,9]* Edan R. Tulman,[1,4,5] Claudio L. Afonso,[1,6] Zhiqiang Lu,[1] James J. Becnel,[2] Bettina A. Moser,[2,7,8] Gerald F. Kutish,[1,5,6] and Daniel L. Rock[1,9]

Plum Island Animal Disease Center, Agricultural Research Service, U.S. Department of Agriculture, Greenport, New York 11944[1]; CMAVE, Agricultural Research Service, U.S. Department of Agriculture, Gainesville, Florida 32608[2]; Area of Virology, School of Veterinary Sciences, University of Buenos Aires, 1427 Buenos Aires, Argentina[3]; Department of Pathobiology and Veterinary Science[4] and Center of Excellence for Vaccine Research,[5] University of Connecticut, Storrs, Connecticut 06269; Southeast Poultry Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Athens, Georgia 30605[6]; U.S. Army Veterinary Laboratory Europe, APO AE 09180[7]; Microbiology Department, Veterinary Laboratory Europe, Geb. 3810, Zi. 122B, 66849 Landstuhl-Kirchberg, Germany[8]; and Department of Pathobiology, College of Veterinary Medicine, University of Illinois, Urbana, Illinois 61802[9]

Received 6 March 2006/Accepted 1 June 2006

Iridoviruses (IVs) are classified into five genera: *Iridovirus* and *Chloriridovirus*, whose members infect invertebrates, and *Ranavirus*, *Lymphocystivirus*, and *Megalocytivirus*, whose members infect vertebrates. Until now, *Chloriridovirus* was the only IV genus for which a representative and complete genomic sequence was not available. Here, we report the genome sequence and comparative analysis of a field isolate of *Invertebrate iridescent virus type 3* (IIV-3), also known as mosquito iridescent virus, currently the sole member of the genus *Chloriridovirus*. Approximately 20% of the 190-kbp IIV-3 genome was repetitive DNA, with DNA repeats localized in 15 apparently noncoding regions. Of the 126 predicted IIV-3 genes, 27 had homologues in all currently sequenced IVs, suggesting a genetic core for the family *Iridoviridae*. Fifty-two IIV-3 genes, including those encoding DNA topoisomerase II, NAD-dependent DNA ligase, SF1 helicase, IAP, and BRO protein, are present in IIV-6 (*Chilo iridescent virus*, prototype species of the genus *Iridovirus*) but not in vertebrate IVs, likely reflecting distinct evolutionary histories for vertebrate and invertebrate IVs and potentially indicative of genes that function in aspects of virus-invertebrate host interactions. Thirty-three IIV-3 genes lack homologues in other IVs. Most of these encode proteins of unknown function but also encode IIV3-053L, a protein with similarity to DNA-dependent RNA polymerase subunit 7; IIV3-044L, a putative serine/threonine protein kinase; and IIV3-080R, a protein with similarity to poxvirus MutT-like proteins. The absence of genes present in other IVs, including IIV-6; the lack of obvious colinearity with any sequenced IV; the low levels of amino acid identity of predicted proteins to IV homologues; and phylogenetic analyses of conserved proteins indicate that IIV-3 is distantly related to other IV genera.

Iridoviruses (IVs) (family *Iridoviridae*) are large, icosahedral, double-stranded DNA (dsDNA) viruses that cause subclinical to lethal infections in invertebrates and poikilothermic vertebrates. IV particles are 120 to 300 nm in diameter and consist of a central core of nucleic acid and proteins, an intermediate lipid membrane, a $T = 147$ capsid comprised of copies of the major capsid protein (MCP), and, in virions that bud from the plasma membrane, an external envelope (90, 95). IV genomic DNA is a linear molecule which is circularly permuted and terminally redundant (30). Although IV replication includes nuclear and cytoplasmic stages, virion assembly occurs exclusively in the cytoplasm (28). In contrast to most dsDNA insect viruses, insect IVs are not occluded in a protein matrix.

Based on particle size, host range, DNA cross-hybridization, the presence of a methyl transferase, and the MCP gene sequence, IVs are classified into five genera, with viruses of two genera (*Iridovirus* and *Chloriridovirus*) infecting invertebrates, mostly insects (invertebrate iridescent viruses [IIVs]), and viruses of the other three (*Ranavirus*, *Lymphocystivirus*, and *Megalocytivirus*) infecting cold-blooded vertebrates (vertebrate iridoviruses [VIVs]) (47, 84, 91). Complete genome sequences have been determined for viruses representing four of the five genera, including lymphocystis disease virus 1 (LCDV-1) and LCDV-China (LCDV-C) of the genus *Lymphocystivirus*; tiger frog virus (TFV), frog virus 3 (FV-3), *Ambystoma tigrinum* virus (ATV), and Singapore grouper iridovirus (SGIV) of the genus *Ranavirus*; rock bream iridovirus (RBIV) and infectious spleen and kidney necrosis virus (ISKNV) of the genus *Megalocytivirus*; and IIV-6 (or Chilo iridescent virus) of the genus *Iridovirus* (22, 34, 35, 42, 43, 69, 73, 76, 98). Also available are the genome sequences of grouper iridovirus (GIV) and orange-spotted GIV (OSGIV), viruses currently lacking taxonomic classification (53, 78). IV genomes range in size from 105 to 212 kbp and contain 96 to 234 largely nonoverlapping open reading frames (ORFs), a G+C content ranging from 27 to 55%, and complex repeat sequences mostly located between coding regions. Genomes exhibit little to no colinearity among genera (see Table 1).

The type species and currently sole member of the genus

* Corresponding author. Mailing address: Department of Pathobiology, College of Veterinary Medicine, University of Illinois, 2001 South Lincoln Ave., Urbana, IL 61802. Phone: (217) 244-3989. Fax: (217) 244-7421. E-mail: gadelhon@uiuc.edu.

TABLE 1. Characteristics of IV genomes

Virus	Genus	Genome size (bp)	GC content (%)	No. of putative genes	Coding density (%)	Protein size range (aa)	Accession no.	Reference
IIV-3	*Chloriridovirus*	190,132	48	126	68	60–1,377	DQ643392	This work
IIV-6	*Iridovirus*	212,482	29	234	85	40–2,432	AF303741	42
ATV	*Ranavirus*	106,332	54	96	79	32–1,294	AY150217	43
TFV	*Ranavirus*	105,057	55	105	94	40–1,294	AF389451	35
FV3	*Ranavirus*	105,903	55	98	80	50–1,293	AY548484	73
SGIV	*Ranavirus*	140,131	48	162	98	41–1,268	AY521625	69
LCDV-1	*Lymphocystivirus*	102,653	29	110	82	40–1,199	L63545	76
LCDV-C	*Lymphocystivirus*	186,250	27	176	67	40–1,193	AY380826	98
ISNKV	*Megalocytivirus*	111,362	55	124	93	40–1,208	AF371960	34
RBIV	*Megalocytivirus*	112,080	53	118	86	50–1,253	AY532606	22
GIV	Unassigned	139,793	49	120	83	62–1,268	AY666015	78
OSGIV	Unassigned	112,636	54	121	91	40–1,168	AY894343	53

Chloriridovirus is *Invertebrate iridescent virus type 3* (IIV-3), also known as mosquito iridescent virus (MIV). IIV-3 is characterized by its restricted host range (mosquitoes [*Diptera*]) and relatively large particle size (180 nm) (17, 90). In contrast, IIVs of the *Iridovirus* genus have been isolated from diverse hosts of the orders *Diptera*, *Lepidoptera*, *Hemiptera*, and *Coleoptera* and are approximately 120 nm in diameter (89). *Iridovirus* species include IIV-1, the first reported IV; IIV-6, the type species of the genus; and several tentative species which remain to be fully characterized (25, 89, 94). Notably, IIV-6 can also infect mosquitoes but causes only sublethal (covert) infections and reduced fitness in *Aedes aegypti* relative to noninfected conspecifics (55).

IIV-3 was originally isolated from larvae of the salt marsh mosquito *Ochlerotatus* (formerly *Aedes*) *taeniorhynchus* (Wiedemann) (19), with successful transmission to an additional mosquito host, *Ochlerotatus sollicitans* (93). IIV-3 has been isolated from several other mosquito species, including *Aedes vexans*, *Psorophora ferox*, *Culiseta annulata*, and *Culex territans*, which are important pests of both humans and domestic animals (14, 17, 93). Early mosquito larval stages are most susceptible to IIV-3 infection, but clinical disease (yellow-green iridescence beneath the epidermis) and death rates are highest in the fourth instar (93). IIV-3 infection of *O. taeniorhynchus* results in virus replication in the fat body and to a lesser extent in the dermis, imaginal disks, trachea, gonads, and hemocytes (33). Oral and transovarian transmission of IIV-3 have been documented for larval mosquitoes (19, 51, 52, 93). Two IIV-3 strains have been described, a field isolate referred to as regular strain (RMIV) and a laboratory isolate referred to as turquoise strain (TMIV) (93); these cause orange and blue-green iridescence, respectively, in infected larvae.

Despite the role of mosquitoes as significant vectors of human disease, chloriridoviruses are the least studied members of the family *Iridoviridae*. To date, IIV-3 genomic data are limited to the sequence of the DNA polymerase gene of RMIV (71, 83) and estimations that the TMIV genome contains about twice as much repetitive DNA as that of RMIV (82). Here, we report the complete genomic sequence, with analysis, of IIV-3, RMIV strain. These data, combined with previous work on viruses belonging or related to the other four IV genera, provide a first comparative overview of IV genomics.

MATERIALS AND METHODS

IIV-3 DNA isolation, cloning, and sequencing. The RMIV strain was originally obtained from field-collected *Ochlerotatus taeniorhynchus* (*Diptera*: *Culicidae*) larvae and maintained by horizontal transmission in the laboratory. RMIV-infected *O. taeniorhynchus* larvae were ground with a Tekmar Tissuemizer in deionized water and filtered through 400 mesh to remove large insect parts. The filtrate was layered on a continuous HS-40 Ludox gradient and spun for 30 min at 16,000 × *g*. The band containing RMIV was collected, resuspended in deionized water, and spun for 30 min at 16,000 × *g*, and genomic DNA was extracted from the pellet as previously described (87). Genomic DNA was incompletely digested with TacI endonuclease (New England Biolabs, Beverly, Mass.), and fragments larger than 1.0 kbp were cloned into the dephosphorylated EcoRI site of pUC19 plasmid vector and grown in *Escherichia coli* DH10B cells (Gibco-BRL, Gaithersburg, Md.). Plasmids were purified by alkaline lysis according to the manufacturer's instructions (Eppendorf-5 Prime, Boulder, Colo.). DNA templates were sequenced from both ends with M13 forward and reverse primers using dideoxy chain terminator sequencing chemistries (65) and the Applied Biosystems PRISM 3700 automated DNA sequencer (Applied Biosystems, Foster City, Calif.). Chromatogram traces were basecalled with Phred (23) and linearly assembled with Phrap (http://www.phrap.org), with the quality files and default settings used to produce a consensus sequence which was manually edited with Consed (31). An identical sequence was assembled by using the Cap3 assembler with quality files and clone length constraints (38). While the IIV-3 genome assembled essentially in a circular fashion consistent with circular permutation, specifically, randomly located but overlapping termini were observed in independent linear assemblies. For clarity, the IIV-3 genome is presented here in a linear fashion, with overlapping sequence removed from the right terminus of one assembly and the first base of the left terminus arbitrarily numbered base one. The final DNA consensus sequence represented on average 10-fold redundancy at each base position (5,483 reactions) and had a Consed estimated error rate of 0.04/10 kbp.

Sequence analysis. DNA composition, structure, repeats, and restriction enzyme patterns were analyzed as previously described (1). ORFs longer than 30 codons with a methionine start codon were evaluated for coding potential by the Hexamer (ftp.sanger.ac.uk/pub/rd), Glimmer (64), and Framefinder (http://www.ebi.ac.uk/~guy) computer programs. Minor ORFs contained within larger ORFs were excluded. Here, 126 ORFs are annotated as potential genes and numbered from left to right. Given the predicted nature of all IIV-3 genes and gene products, ORF names are used throughout this work to indicate both the predicted gene and its putative protein product. DNA and protein comparisons with entries in genetic databases (PROSITE, Pfam, Prodom, Sbase, Blocks, and GenBank) were performed with the BLAST (4), PsiBlast (5), FASTA, TFASTA (63), and HMMER (70) programs. The GCG (20), MEMSAT (44), Psort (61), and SAPS (13) programs were used for general analysis, transmembrane prediction, and physical characterization of predicted proteins. DNA repeats were analyzed with Lalign (http://workbench.sdsc.edu) and Dot Plot (GCG package). Sequence quality, sequence depth, and the lack of obvious polymorphism observed in IIV-3 repeat regions were consistent with that of the rest of the IIV-3 genome assembly. Multiple alignments were performed with Clustal, Dialign-T, and Kalign (49, 72, 75). Phylogenetic analysis was performed with the PHYLO_WIN, TREE-PUZZLE, IQPNNI, PHYLIP, PHYML, and MRBAYES software pack-

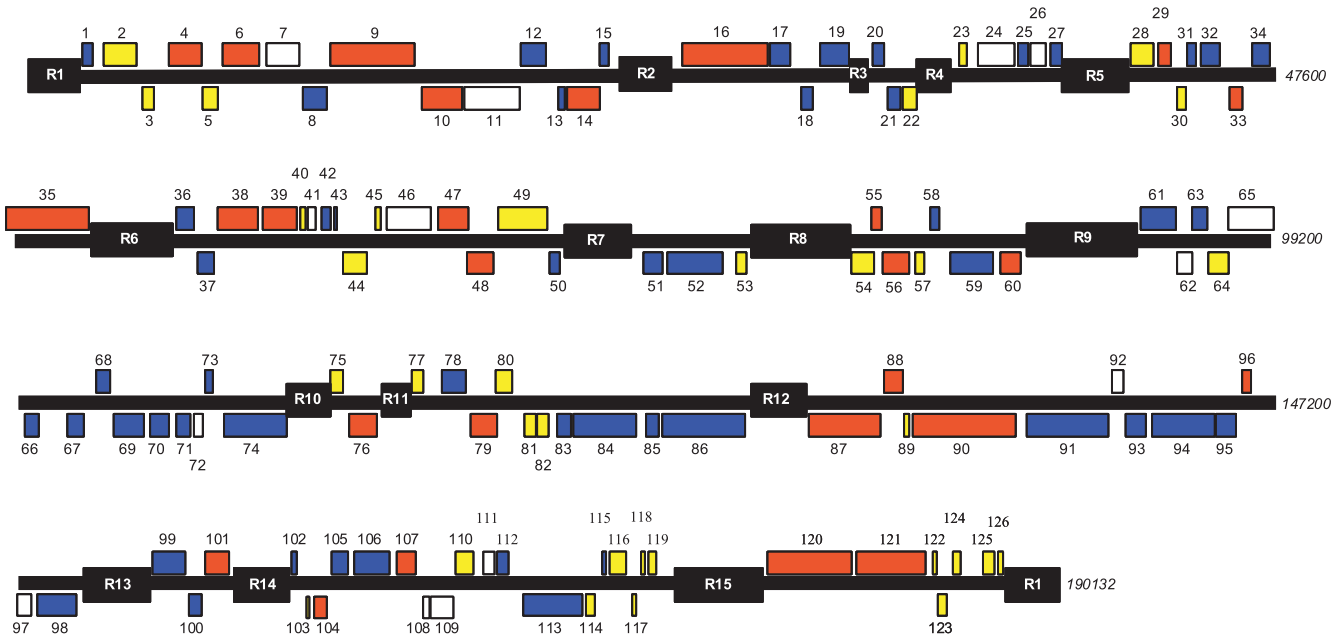

FIG. 1. Linear map of the IIV-3 genome. ORFs are numbered from left to right. ORFs transcribed to the right or to the left are located above and below the horizontal line, respectively. Red boxes represent genes encoded by all currently sequenced IV genomes. Blue, yellow, and white boxes are genes unique for IIVs, IIV-3, and selected IVs, respectively. Repetitive DNA sequences are shown as numbered black boxes.

ages (16, 24, 26, 32, 39, 66), and evolutionary models were selected with ModelGenerator (http://bioinf.nuim.ie/software/modelgenerator). Additional analyses were conducted on alignments where poorly aligned regions were removed with Gblocks (16).

Nucleotide sequence accession number. The IIV-3 genome sequence has been deposited in the GenBank database under accession no. DQ643392.

RESULTS AND DISCUSSION

Genome organization. IIV-3 DNA was assembled in a contiguous sequence of 190,132 bp, a figure that differs significantly from previous size estimates based on restriction endonuclease analysis (135,000 bp) and sucrose gradient centrifugation (383,000 bp) (82, 88). IIV-3 nucleotide composition averaged 48% G+C, which is lower than previously reported (54%) (82) and closer to the ranavirus and megalocytivirus genomes than to IIV-6, the other fully sequenced IIV (Table 1). The IIV-3 base composition was not uniformly distributed throughout the genome, with lower G+C values (42%) found in regions containing repeated sequences.

Based on coding potential analysis and similarity to known proteins, 126 IIV-3 ORFs were annotated here (Fig. 1; Table 2). These ORFs, which encode proteins of 60 to 1,377 amino acids (aa), represent a coding density of 68%, one of the lowest among fully sequenced IVs (Table 1). IIV-3 contains 33 unique genes, 27 homologues of genes present in all sequenced IVs, and 52 genes present in IIV-6 but not in VIVs. Consistent with the lack of gene colinearity observed between IIV-6 and VIV and even between LCDV-1 and LCDV-C, viruses likely belonging to the same genus, *Lymphocystivirus*, the IIV-3 genome exhibited no obvious colinearity with any other completely sequenced IV genome.

Repeated sequences. The IIV-3 genome contained highly repetitive DNA sequences in 15 distinct regions (designated R1 to R15 from left to right) that were 0.8 to 4.6 kbp in length, irregularly distributed, and apparently noncoding (Fig. 1). IIV-3 DNA repeats were complex, mostly direct and imperfect, and they could be divided into two groups based on sequence similarity, with group I repeats located at R2, R4, R6 to R8, R11 to R13, and R15 and group II repeats located at R1, R3, R5, R9, R10, and R14. Group I repeats were comprised of a 100-nucleotide (nt) sequence present in 2 to 10 copies which had 83 to 100% nucleotide identity. Group II repeats were also comprised of a sequence of approximately 100 nt and were present in two to six copies which had 80 to 99% nucleotide identity. Group I and group II repeats had only 62 to 68% nucleotide identity. Repeated sequences in each group contained invariant motifs (e.g., TAAATTTC, AATC, and GCAT in group I repeats; GAGTT, ATGCGT, and GAAATTT in group II repeats) flanked by less-conserved sequences.

DNA repeats are present in the intergenic regions of all fully sequenced IVs; however, their extent, arrangement, localization, and repeated sequence motifs differ between genera (22, 34, 35, 42, 43, 53, 69, 73, 76, 78, 98). In contrast to previous reports, IIV-3 repeated sequences were extensive (20% of the genome), resembling IIV-9 and IIV-16, in which 25% and 39% of the genomes, respectively, are made up of repetitive DNA (10, 58, 89). Although late transcription has been detected from IIV-9 repeated DNA regions, the role of IV DNA repeats is unknown (59). Repeated sequences are known to be involved in genome replication and gene transcription in other DNA viruses. For example, in nucleopolyhedroviruses (NPVs), repetitive DNA sequences function as enhancers of transcription and as origins of genomic DNA replication (3, 74).

Notable IIV-3 genes. (i) Viral transcription and DNA replication. The IIV-3 genome contained several genes with predicted roles in viral transcription and RNA processing,

TABLE 2. IIV-3 ORFs

ORF	Nucleotide position	Length (codons)	Best match[a]				Predicted structure and/or function[b]
			Protein	Accession no.	Blast score	% aa identity	
IIV3-001R	2620–3063	148	IIV-6 395R	AF303741	97	31	
IIV3-002R	3603–4976	458	*Drosophila* CG4416 product	AE014298	122	31	
IIV3-003L	5638–5171	156					
IIV3-004R	5671–7023	451	IIV-6 067R	AF303741	608	39	
IIV3-005L	7867–7217	217	*Myxococcus* unknown	AF448145	120	29	Cys-rich protein
IIV3-006R	8063–9544	494	IIV-6 118L	AF303741	546	30	Myristylated membrane protein
IIV3-007R	9683–11023	447	TFV 025R	AF389451	151	26	FV-3 31-kDa-like early protein
IIV3-008L	12292–11252	347	IIV-6 443R	AF303741	172	33	
IIV3-009R	12359–15778	1,140	IIV-6 428L	AF303741	1977	41	Rpb2
IIV3-010L	17538–15862	559	IIV-6 380R	AF303741	582	34	S/T protein kinase
IIV3-011L	19897–17645	751	IIV-6 380R	AF303741	287	26	
IIV3-012R	20020–21138	373	IIV-6 302L	AF303741	171	28	C2H2 Zn finger protein
IIV3-013L	21570–21301	90	IIV-6 141R	AF303741		32	
IIV3-014L	23106–21709	466	IIV-16 MCP	AF025775	1792	58	Major capsid protein
IIV3-015R	23196–23621	142	IIV-22 hypothetical 15.9 kDa	P25097	266	53	
IIV3-016R	26289–29705	1,139	IIV-6 295L	L63545	307	25	
IIV3-017R	29876–30742	289	IIV-6 335L	AF303741	133	28	
IIV3-018L	31615–31121	165	IIV-6 415R	AF303741	275	44	
IIV3-019R	31881–33098	406	IIV-6 201R	AF303741	656	40	Bro family protein
IIV3-020R	33982–34476	165	IIV-6 196R	AF303741	255	36	Thioredoxin
IIV3-021L	35147–34530	206	*Spodoptera* IAP	AF186378	139	26	C3HC4 RING finger protein
IIV3-022L	35868–35194	225					TM
IIV3-023R	37466–37783	106					
IIV3-024R	38120–39592	491	IIV-6 224L	AF303741	828	40	Cysteine protease
IIV3-025R	39626–40159	178	IIV-6 111R	AF303741	86	26	TM
IIV3-026R	40206–40883	226	IIV-6 350L	AF303741	256	36	VV A1L-like transcription factor
IIV3-027R	40990–41502	171	IIV-6 157L	AF303741	111	26	RING finger protein
IIV3-028R	44228–45238	337					
IIV3-029R	45281–45859	193	IIV-6 143R	AL390114	404	40	Thymidine kinase
IIV3-030L	46385–45954	144	*Arabidopsis thaliana*	AY087324	77	26	Thioredoxin-like protein
IIV3-031R	46406–46831	142	IIV-6 115R	AF303741	97	26	
IIV3-032R	46927–47763	279	*Xenopus* unknown	BC041236	152	31	
IIV3-033L	48462–47881	194	IIV-6 307L	AF303741	401	49	
IIV3-034R	48593–49390	266	IIV-6 077L	AF303741	120	32	Zn finger domain
IIV3-035R	49517–52810	1,098	IIV-6 179R	AF303741	1359	34	Protein kinase active-site signature
IIV3-036R	56099–56872	258	IIV-6 219L	AF303741	245	34	
IIV3-037L	57696–56938	253	IIV-6 060L	AF303741	148	27	TM
IIV3-038R	57775–59412	546	IIV-6 098R	AF303741	515	28	
IIV3-039R	59602–60930	443	IIV-6 393L	AF303741	314	28	FV-3 IE ICP46-like
IIV3-040R	61046–61294	83					
IIV3-041R	61328–61714	129	IIV-6 453L	AF303741	172	35	Thioredoxin
IIV3-042R	61812–62288	159	IIV-6 136R	AF303741	280	45	Zn binding signature
IIV3-043R	62326–62517	64	IIV-6 010R	AF303741	169	53	
IIV3-044L	63715–62711	335	*Saccharomyces* YOLO44W	Z74786	212	29	S/T protein kinase
IIV3-045R	63932–64201	90					
IIV3-046R	64376–66172	599	IIV-6 229L	AF303741	418	27	
IIV3-047R	66434–67654	407	IIV-6 337L	AF303741	583	37	Virion-associated membrane protein
IIV3-048L	68853–67726	376	IIV-6 376L	AF303741	763	46	Ribonucleotide reductase small subunit
IIV3-049R	69009–71057	683	CPXV 158	AF482758	175	24	
IIV3-050L	71580–71125	152	IIV-6 145L	AF303741	110	25	
IIV3-051L	75421–74576	282	IIV-6 213R	AF303741	109	35	
IIV3-052L	77801–75534	756	IIV-6 205R	AF303741	559	31	NAD-dependent DNA ligase
IIV3-053L	78645–78220	142	*Encephalitozoon* Rbp 19 kD	AL590447	73	25	Rpb7
IIV3-054L	83584–82658	309					
IIV3-055R	83754–84164	137	IIV-6 349L	AF303741	184	33	Transcription factor SII
IIV3-056L	85251–84220	344	IIV-6 287R	AF303741	319	27	
IIV3-057L	85749–85360	130					
IIV3-058R	85937–86386	150	IIV-6 391R	AF303741	233	32	
IIV3-059L	88441–86762	560	IIV-6 012L	AF083915	582	34	Exonuclease II
IIV3-060L	89678–88827	284	PBCV-1 A193L	U42580	130	25	PCNA-like protein
IIV3-061R	94254–95717	488	IIV-6 467R	AF303741	210	25	
IIV3-062L	96241–95546	232	ATV hypothetical protein	AY150217	152	56	Repetitive protein
IIV3-063R	96289–96924	212	IIV-6 309L	AF303741	80	32	
IIV3-064L	97827–96988	280					
IIV3-065R	97890–99779	630	*Rickettsia* RR alpha chain	AE008624	1167	57	Ribonucleotide reductase large subunit

Continued on facing page

TABLE 2—*Continued*

ORF	Nucleotide position	Length (codons)	Best match[a]				Predicted structure and/or function[b]
			Protein	Accession no.	Blast score	% aa identity	
IIV3-066L	100438–99860	193	IIV-6 357R	AF303741	97	25	TM
IIV3-067L	102380–101661	240	IIV-6 197R	AF303741	262	44	S/Y protein phosphatase
IIV3-068R	102821–103423	201	IIV-6 401R	AF303741	404	44	HMG-like
IIV3-069L	104770–103502	423	IIV-6 198R	AF303741	407	30	
IIV3-070L	105749–104904	282	IIV-6 306R	AF303741	171	31	SWIB/MDM2 domain
IIV3-071L	106597–105935	221	IIV-6 259R	AF303741	165	27	
IIV3-072L	107142–106675	156	IIV-6 374R	AF303741	103	28	
IIV3-073R	107238–107768	177	IIV-6 234R	AF303741	93	26	TM
IIV3-074L	110351–107850	834	IIV-6 268L	AF303741	496	30	
IIV3-075R	112290–112763	158	*Debaryomyces* plasmid protein	AJ011124	132	35	DNA Pol III box
IIV3-076L	114217–113078	380	IIV-6 369L	AF303741	466	36	*Xeroderma pigmentosum* protein
IIV3-077R	115523–115963	147					
IIV3-078R	116034–117074	347	IIV-6 244L	AF303741	691	49	Phosphoesterase
IIV3-079L	118306–117122	395	IIV-6 282R	AF303741	688	43	
IIV3-080R	118504–119226	241	CMLV112R	AF438165	136	28	MutT-like protein
IIV3-081L	120310–119786	175					
IIV3-082L	120800–120315	162					
IIV3-083L	121653–121078	192	IIV-6 358L	AF303741	87	28	
IIV3-084L	124239–121708	844	IIV1 98-kDa protein	P22856	1096	38	OTU domain
IIV3-085L	125066–124584	161	IIV-6 203L	AF303741	106	22	TM
IIV3-086L	128555–125217	1,113	IIV-6 045L	AF083915	2177	42	DNA topoisomerase II
IIV3-087L	134049–131110	980	IIV-6 022L	AF083915	1127	31	NTPase SNF2
IIV3-088R	134089–134871	261	IIV-6 075L	AF303741	630	47	ATPase
IIV3-089L	135221–134928	98					
IIV3-090L	139445–135315	1,377	IIV-6 176R	AF303741	1723	41	Rpb1
IIV3-091L	142900–139613	1,096	IIV-6 443R	AF303741	945	34	Repetitive protein
IIV3-092R	142975–143487	171	IIV-6 454R	AF303741		34	Rpb5
IIV3-093L	144466–143564	301	IIV-6 019R	AF303741	271	28	
IIV3-094L	147069–144583	829	IIV-6 050L	AF303741	382	27	P-loop nucleoside triphosphatase
IIV3-095L	148265–147177	363	Mouse type 2 metalloprotease	D86332	375	40	Metalloprotease
IIV3-096R	148391–148834	148	IIV-6 347L	AF303741	221	42	Erv1/Air-like thiol oxidoreductase
IIV3-097L	150653–150039	205	IIV-6 170L	AF303741	191	30	RuvC-like resolvase
IIV3-098L	152364–150805	520	IIV-6 439L	AF303741	535	33	Protein kinase-like protein
IIV3-099R	155290–156675	462	IIV-6 329R	AF303741	314	41	
IIV3-100L	157359–156724	212	IIV-6 378R	AF303741	460	44	
IIV3-101R	157391–158389	333	IIV-6 142R	AF303741	431	39	Ribonuclease III
IIV3-102R	161028–161390	121	*Buchnera* pyrophosphorylase	AE014078	95	38	
IIV3-103L	161631–161452	60					
IIV3-104L	162248–161691	186	IIV-6 355R	AF303741	390	47	CTD phosphatase
IIV3-105R	162408–163145	246	IIV-6 359L	AF303741	90	29	
IIV3-106R	163297–164715	473	IIV-6 030L	AF083915	817	42	Helicase
IIV3-107R	164859–165644	262	IIV-6 117L	AF303741	222	33	TM
IIV3-108L	166065–165754	104	IIV-6 161L	AF303741	311	57	Helicase, C-terminal domain
IIV3-109L	167108–166122	329	IIV-6 161L	AF303741	444	39	Helicase, N-terminal domain
IIV3-110R	167222–167992	257					
IIV3-111R	168970–169476	169	IIV-6 414L	AF303741	184	34	MutT-like protein
IIV3-112R	169526–169861	112	IIV-6 466R	AF303741	95	33	TM
IIV3-113L	172363–169943	807	IIV-6 155L	AF303741	300	32	
IIV3-114L	172916–172482	145	*Cyprinus carpio* MMp	AB057407	91	40	Metalloprotease prodomain
IIV3-115R	173073–173300	76	IIV-6 342R	AF303741	142	50	
IIV3-116R	173407–174168	254					
IIV3-117L	174511–174308	68					
IIV3-118R	174533–174754	74					
IIV3-119R	174807–175184	126	*Pneumocystis* endoprotease	AF009223	140	42	Proline-rich protein
IIV3-120R	179618–183040	1,141	IIV-3 DNA polymerase	AJ312708	4702	89	DNA polymerase
IIV3-121R	183279–186101	941	IIV-6 184R	AF303741	1367	34	VV D5R NTPase/helicase
IIV3-122R	186194–186394	67					
IIV3-123L	186865–186449	139					
IIV3-124R	186910–187617	236	*Xanthomonas* ribonuclease	AE011846	106	27	
IIV3-125R	187711–188601	297					
IIV3-126R	188648–188962	105					

[a] CPXV, cowpox virus; CMLV, camelpox virus. Accession numbers are numbers of homologous sequences from the GenBank database.
[b] Function was deduced from the degree of similarity to known genes and from Prosite signatures.TM, transmembrane.

```
Mosquito                    -----MFYHISLEHEILLHPRYFGPQLIETVKQKLYTEVEGTCTGKYGFV
Rat                         -----MFYHISLEHEILLHPRYFGPNLLNTVKQKLFTEVEGTCTGKYGFV
Saccharomyces_cerevisiae    -----MFFIKDLSLNITLHPSFFGPRMKQYLKTKLLEEVEGSCTGKFGYI
Encephalitozoon_cuniculi    -----MFFVRDFVQNIHLSPNYLGPNIQALIEEYLLSKVEGSCS-SSGYV
Mimivirus_367L              MAQQSLYFQTKLEDKVSLLPSQMVGNMENYLLENLEAKVKDKVT-EHGIV
IIV3-053L                   --MEQYLQAFEFVEEMVVLPKYLSWELYHHLAVLLREKYPKTYKNK-GYI
                                      .:  :: : *  :  .:   :   *  :   . . . * :

Mosquito                    IAVTTID--DIGSGTIQPGQGFVVYPVKYKAIVFRPFKGEVLDA--TVKQ
Rat                         IAVTTID--NIGAGVIQPGRGFVLYPVKYKAIVFRPFKGEVVDA--VVTQ
Saccharomyces_cerevisiae    LCVLDYDNIDIQRGRILPTDGSAEFNVKYRAVVFKPFKGEVVDG--TVVS
Encephalitozoon_cuniculi    VMVLSID--EISESRIILT-GETIFTVKYKALTLKPLKGEVIDA--NVVE
Mimivirus_367L              LKVNRIIEYDYGIISKNNFSGTAIYRVKYECLICSPVKNLSIIC--LVEN
IIV3-053L                   FNIKVKSILDNR----ITPTGQIVLVVMFQSDLYVPQVGHVFTERIRVNS
                            . :      :          *     * :..    *  .  .     * .

Mosquito                    VNKVGMFAEIGPLSCFISHHSIPADMQFCPNGAPPCYRAINGESVIAAED
Rat                         VNKVGLFTEIGPMSCFISRHSIPSEMEFDPNSNPPCYKTMDEDIVIQQDD
Saccharomyces_cerevisiae    CSQHGFEVQVGPMKVFVTKHLMPQDLTFNAGSNPPSYQSSED--VITIKS
Encephalitozoon_cuniculi    TNKMGVFASVGPLTVFISNHQIPNFLLEN---------------EITKNV
Mimivirus_367L              IVKGYIIAKNGPVIVAIPFNNIDSDKFQLTNGNIVYKNNSNN---IQKGD
IIV3-053L                   VDDRYQWITIEPLTVFLRSN------------------------IPYKP
                              .        *:   :  :                         *

Mosquito                    KIRLKIVGTRVDATGIFAIGTLMDDYLGLVGS------------------
Rat                         EIRLKIVGTRVDKNDIFAIGSLMDDYLGLVS-------------------
Saccharomyces_cerevisiae    RIRVKIEGCISQVSSIHAIGSIKEDYLGAI--------------------
Encephalitozoon_cuniculi    MIRLKIIGTKIDSTRIYAVGTLNDDSLGIIS-------------------
Mimivirus_367L              YVKVSIINIKTN---------LNEKKITTIAKLLDMATNDEIKSYDNDQL
IIV3-053L                   NTLVTVQICSIKMDNTLCFGTILD--------------------------
                               :.:     .         : :

Mosquito                    --------------
Rat                         --------------
Saccharomyces_cerevisiae    --------------
Encephalitozoon_cuniculi    --------------
Mimivirus_367L              LIVNGDVDDEQEFI
IIV3-053L                   --------------
```

FIG. 2. Multiple amino acid sequence alignment of IIV3-053L with selected Rpb7 proteins. Asterisks, colons, and single dots denote identical, conserved, and semiconserved residues, respectively (74). Rpb7 protein sequence names correspond to the following accession numbers: mosquito (*Anopheles gambiae*), XM_320916; rat, NM_053948; *Saccharomyces cerevisiae*, NC_001136; *E. cuniculi*, NC_003234; mimivirus (*Acanthamoeba polyphaga* mimivirus), NC_006450.

including five RNA polymerase II subunit (Rpb) homologues, and in viral DNA replication, metabolism, and maintenance (Table 2).

IIV3-053L is a novel IV protein with limited similarity to Rpb7, a DNA-dependent RNA polymerase II subunit highly conserved among eukaryotes and having roles in DNA repair, transcription, and RNAi-directed chromatin silencing (18, 21, 56) (Fig. 2). Cellular Rpb7 forms a complex with Rpb4 near the transcript exit groove and the C-terminal domain (CTD) linker region of RNA polymerase II, interacts with various transcription factors and CTD phosphatase FCP1, and is essential for cell viability in yeast (18). IIV3-053L was most similar to *Encephalitozoon cuniculi* Rpb7 (25% amino acid identity over 142 amino acids) (Fig. 2), particularly in the N-terminal half of the protein and including G64 in the predicted tip loop involved in binding of Rpb7 to the RNA polymerase II core (6). Less conserved was the C-terminal half of IIV3-053L, which lacked 20 aa and a predicted three-stranded antiparallel β-sheet relative to *E. cuniculi* Rpb7 (6, 77). IIV3-053L lacked recognizable Rpb7 Pfam signatures, suggesting that it, like homologues in *Giardia*, *Methanosarcina*, and mimivirus, is a highly divergent Rpb7 whose function in transcription remains to be determined.

IIV3-092R had limited similarity with predicted Rpb5-like proteins encoded by African swine fever virus (ASFV), mimivirus, and several IVs. Differences between IIV3-092R and other viral Rpb5-like proteins were greatest in the N-terminal two-thirds of the protein, a region in cellular Rpb5 proposed to interact with transcriptional regulators (85).

IIV3-104L was similar to the FCP1 phosphatase catalytic domain (FCPH), with orthologues in all fully sequenced IVs. FCP1 is responsible for dephosphorylating Rpb1 CTD and, together with specific kinases, driving the transition of RNA polymerase II from the initiation to elongation modes. Like SCP1, a recently described FCP1-like phosphatase with similar roles in transcription regulation, IIV3-104L and IV homologues lack most sequences N-terminal of the FCPH domain, including BRCT and TFIIF binding domains (96). Given the lack of recognizable CTD heptapeptide repeats in IV Rpb1 homologues, IV IIV3-104L-like proteins may conceivably target host Rpb1, a protein which is involved in the nuclear phase of IV transcription (29). Whether IIV3-104L-like proteins represent the IV virion-associated protein (VATT) proposed to modify host RNA polymerase II and to initiate iridoviral immediate-early transcription remains to be determined (92).

IIV3-080R and IIV3-111R contained domains characteristic

of MutT-like proteins, including the signature motif [G-X_5-E (D in IIV3-080R)-X_{4-5}-C/T-L/A-X-RE-F/L-X-EE-X-G/T] at positions 136 to 157 and 76 to 98, respectively. The MutT (or nudix) domain is found in certain phosphohydrolases which are believed to eliminate toxic nucleotide derivatives and to regulate the levels of signaling nucleotides in both eukaryotes and prokaryotes (57). While IIV3-111R had additional similarity with IIV-6 and LCDV MutT-like proteins, IIV3-080R is a novel IV ORF more similar to poxvirus MutT proteins (28% amino acid identity to variola minor virus F10R over 169 aa) (Table 2). A poxvirus MutT-like protein, vaccinia virus (VV) D10R, is essential for virus infectivity and appears to function as a repressor of host and viral transcription and translation (67). Whether MutT proteins play a similar role in IV remains to be determined.

IIV3-060L had similarity with viral homologues of proliferating cell nuclear antigen (PCNA). While most similar to the *Paramecium bursaria* chlorella virus (PBCV-1) PCNA-like protein A193L (25% amino acid identity, over 239 aa), IIV3-060L was less similar to PCNA-like proteins encoded by all currently sequenced IVs (17 to 21% amino acid identity) and by NPVs and mimivirus. Cellular PCNA, an acidic protein (pI 4.5) with critical roles in DNA replication and repair, assembles as a ring-shaped homotrimer around dsDNA and is highly conserved among eukaryotes. Nine R and K residues within each monomer confer a net positive electrostatic potential to the central channel, a feature thought to be important for PCNA sliding clamp function (48). In contrast, IV PCNA-like proteins are highly divergent between genera, have predicted pI values ranging from 6.4 (IIV-6) to 9.2 (IIV-3) and, with the exception of conserved residues K37, R102, K168, and K231 in IIV3-060L, contain a pattern and distribution of positive charges distinct from that of cellular PCNA. VV G8R, one of the most divergent viral PCNA-like proteins, is an essential protein that affects late transcription rather than viral DNA replication (46). While the function of IV PCNA-like proteins is unknown, they also may have a novel role during viral replication.

(ii) IIV-3 genes lacking VIV homologues. IIV-3 shared with IIV-6 52 genes absent in VIVs (Fig. 1). These potentially IIV-specific genes encoded several proteins with predicted functions involving DNA replication and maintenance, including DNA topoisomerase IIA (IIV3-086L), Pif1-like SFI helicase (IIV3-106R), NAD-dependent DNA ligase (IIV3-052L), HMGB-like protein (IIV3-068R), and Swib/mdm2 homology domain-containing protein (IIV3-070L), and several predicted to function in protein modification, including OTU-like cysteine protease (IIV3-084L), type 2 metalloprotease (IIV3-095L), and RING finger proteins (IIV3-021L and IIV3-027R). Additionally, a protein containing the BRO motif found in proteins encoded by a range of other insect viruses (IIV3-019R) was identified.

IIV-specific proteins predicted to manipulate DNA contained novel features relative to cellular or viral homologues. IIV3-086L, similar to topoisomerase IIA encoded by other viruses, lacked the C-terminal domain important for nuclear targeting and for interaction with other proteins (50). Similarly, IIV3-106R contained all seven motifs (I, Ia, and II to VI) characteristic of cellular Pif1-like helicases but lacked most of the N-terminal sequences implicated in Pif1/chromatin factor interaction (reviewed in reference 8). The IIV3-052L NAD-dependent DNA ligase contained a putative C-terminal BRCT domain (position 648 to 720) absent in entomopoxviral homologues, suggesting that IIV DNA ligases, similar to cellular homologues, might direct multimeric complex assemblies through this domain (27).

IIV3-068R contained A and B boxes (positions 70 to 138 and 143 to 185, respectively) similar to those found in non-histone chromatin proteins of the high mobility group B (HMGB), proteins that bind and distort DNA and interact with a number of transcription factors and DNA repair and recombination proteins (reviewed in reference 2). IIV HMGB-like proteins lack the acidic tail found in cellular HMGBs. Interestingly, a cellular HMGB mutant lacking the acidic tail exhibits 100-fold-higher DNA binding affinity than the wild-type protein, leading to a block in nucleosome sliding (11). Iridovirus HMGB may conceivable play a structural role in IIV genome conformation.

Several IIV-specific proteins shared similarity or functional motifs with proteins involved in modification of other viral or host proteins. Similar to cellular and viral members of the OTU (for *Drosophila* "ovarian tumor gene") cysteine protease superfamily, IIV3-084L contained in its C terminus the four conserved motifs of an OTU homology domain (position 596 to 719), including residues in motif I (C601) and motif IV (H718) thought to be important for cleavage (54, 68). IIV3-084L also contained a highly positively charged region (position 137 to 218) and a putative bipartite NLS (position 410), features shared by IIV-1 (Tipula iridescent virus) late protein L96 (35% amino acid identity) and IIV-6 232R (22% amino acid identity) (37). The classification of OTU proteins as cysteine proteases was originally based solely on similarity to the arterivirus NSP2 cysteine protease; however, a number of deubiquitinating enzymes exhibiting proteolytically active OTU domains, including a potent downregulator of NF-κB, were recently described (7, 12, 54, 86). IIV3-084L may similarly affect ubiquitin-mediated protein degradation to counteract antiviral cellular responses.

IIV3-095L, a protein with homologues in viruses infecting insects, including IIV-6, granuloviruses, and entomopoxviruses, was similar to mammalian type 2 matrix metalloproteinases (MMp). IIV3-095L contained an N-terminal signal peptide, the MMp prodomain consensus sequence (or cysteine switch) (PRCXXPD [position 117 to 123]), the catalytic domain signature (HEXXHXXGXXH [position 275 to 285]), a conserved Met at position 293 (reviewed in reference 62), and RRRR and RTRR motifs similar to the equivalently located RRKR furin cleavage target motif of mammalian MMp, indicating that IIV3-095L is likely activated by furin-mediated proteolysis of the prodomain (81). IIV3-095L lacked predicted transmembrane domains and the hemopexin-like domain which affects substrate recognition in many mammalian MMps. Relative to IIV3-095L, the MMp homologue in IIV-6 (165R) lacks 100 N-terminal amino acids, including the signal peptide, the cysteine switch signature, and the potential furin cleavage site, suggesting differences in compartmentalization and activation mechanisms between IIV MMps. IIV3-095L could conceivably function as is the case for entomopathogenic bacterial proteins, targeting receptors, antimicrobial peptides, or other factors involved in insect responses to pathogens, thus facilitating virus replication. Alternatively, viral MMp could act on the

extracellular matrix (e.g., peritrophic matrix of lepidopterous larvae), facilitating virus spread within the host (80). Notably, an MMp-like prodomain is also present in IIV3-114L; however, lack of the prodomain Cys residue critical for MMp function and other MMp features make IIV3-114L an unlikely MMp.

IIV3-021L and IIV3-027R contained C-terminal C3HC4-type RING finger motifs which are characteristic of a diverse array of single-subunit E3 ubiquitin ligases and critical for E3 recruitment of ubiquitin-conjugating enzymes to specific substrates (36). Given the importance of the RING domain in cellular E3 ubiquitin ligase function, its presence in IIV3-021L and IIV3-027R suggests that they too may affect protein ubiquitination to mediate specific virus-host interactions. Notably, the IIV3-021L RING motif ($CX_2CX_4RX_5PCXHX_3CX_2CX_9C$ PXC) exhibits particular similarity to the RING motif present in the E3-like baculovirus inhibitor of apoptosis protein 3 (IAP-3) and is preceded by a single, truncated domain (CXC $X_{10}EX_5H$ [position 126 to 145]) similar to the baculovirus IAP repeat (BIR) domain involved in caspase inhibition by viral and mammalian IAPs (60). These features suggest that IIV3-021L may also participate in caspase inhibition to affect apoptosis in the host cell.

IIV3-019R contained an N-terminal baculovirus repeated ORF (BRO)-family homology domain (Bro-N domain, position 24 to 132), 62% identical to the IIV-6 201R Bro-N domain, and included a single-stranded DNA binding motif at position 24 to 57. The C terminus of IIV3-019R, however, is most similar to the C-terminal domains of proteins lacking Bro-N domains (entomopoxvirus AMV207 and AMV209; 27% amino acid identity), consistent with the extensive domain shuffling found in BRO family proteins (40). BRO family proteins are encoded, often in multiple copies, by many invertebrate dsDNA viruses and by bacteriophages. While three BRO genes have been identified in IIV-6, IIV-3, like IIV-31, contains only one (9, 41). *Bombyx mori* NPV BRO proteins (BRO-a to BRO-e) are expressed early during infection, with BRO-a and BRO-c proteins interacting with chromatin (45, 97). The requirement of BRO proteins for virus replication seems to vary with the specific virus-host system. Disruption of the single BRO gene in the *Autographa californica* multicapsid NPV did not affect virus replication or virus pathogenicity in instar larvae (9). Conversely, inability to recover *B. mori* NPV BRO-d deletion mutants suggested a function essential for virus replication (45).

(iii) IIV-3 genes encoding conserved IV domains. IIV3-016R, IIV3-033L, IIV3-056L, and IIV3-107R are proteins of unknown function similar to proteins encoded by all currently sequenced IVs. IIV3-016R contained in the central region of the protein a conserved 90-aa domain which included the motif $YXCX_{8-9}GX_3NX_{11-12}PCCY$ (position 502 to 533), also found in entomopoxvirus proteins (MSV063 and AMV105) which have similarity with VV A7L early transcription factor subunit. IIV3-033L contained a putative N-terminal transmembrane domain and a 100-aa domain also present in proteins of unknown function encoded by ascovirus (accession no. CAC19143) and *Symbiobacterium thermofilum*, an uncultivable bacterium that depends on bacterial commensalism for growth (79). IIV3-056L and IIV3-107R, predicted to contain an N-terminal transmembrane domain, shared with IV homologues a 50-aa domain which included the motifs F/Y-X_4-V/I-R-G-X_{11}-A-X_2-h-h-X_{13-14}-G-X-P-X-P (position 144 to 187, where "h" is hydrophobic) and R-X_5-D-P/F-IRGD-L/V-X-I-X-P-X_5-F-X_5-P-X_3-L-X_2-G (position 116 to 151), respectively. While conserved, the functional relevance of these protein domains is unknown.

Comparison between IIV-3 and other IV genomes. IIV-3 overall resembled other IVs in genome size, DNA composition, and gene complement. Data also indicated that among completely sequenced IVs, IIV-3 most closely resembled IIV-6, the only other sequenced IIV. Despite these similarities, IIV-3 contained novel genomic features that indicate its distant relationship to other IV genera and confirm its unique position within the family *Iridoviridae*.

IIV-3 contained homologues of 27 genes present in all currently sequenced IVs (Fig. 1, red boxes), indicating that these genes play critical and likely essential roles in aspects of IV biology. Approximately half of the conserved genes are involved in viral transcription/DNA replication (IIV3-009R, IIV3-029R, IIV3-048L, IIV3-055R, IIV3-060L, IIV3-076L, IIV3-087L, IIV3-090L, IIV3-101R, IIV3-104L, IIV3-120R, and IIV3-121R). Others encode homologues of protein kinase, MCP, FV3 IE ICP46-like protein, ATPase, and Erv1/Alr-like protein (IIV3-010L, IIV3-014L, IIV3-039R, IIV3-088R, and IIV3-096R, respectively). The functions of nine IIV-3 genes conserved among IVs, including two genes encoding putative transmembrane proteins (IIV3-006R and IIV3-107R), are unknown.

Fifty-two IIV-3 genes have IIV-6 but not VIV counterparts (Fig. 1, blue boxes), suggesting that these genes (IIV-3/IIV-6 genes) function in infection of invertebrate hosts. Although approximately a third of the IIV-3/IIV-6 genes were clustered in three regions in the IIV-3 genome (positions 29876 to 35147, 99860 to 106597, and 143564 to 157359), even these regions lacked discernible colinearity. Only 12 of 52 IIV-3/IIV-6 genes have a predicted function/activity, with 7 genes likely involved in DNA replication/maintenance or gene expression (IIV3-026R, IIV3-052L, IIV3-059L, IIV3-068R, IIV3-070L, IIV3-078R, IIV3-086L, and IIV3-111R), 4 genes encoding protein modification enzymes (IIV3-020R, IIV3-067L, IIV3-095L, and IIV3-098L), and 1 gene encoding an apoptosis regulator (IIV3-021L). Of the IIV-3/IIV-6 genes with unknown function, six encode predicted transmembrane proteins (IIV3-025R, IIV3-037L, IIV3-066L, IIV3-073R, IIV3-085L, and IIV3-112R), three encode Zn finger proteins (IIV3-012R, IIV3-027R and IIV3-034R), and one encodes a Bro-like protein (IIV3-019R). An additional similarity between IIV-3 and IIV-6 was the lack of specific genes present in VIVs, including those encoding DNA methyltransferase and proteins proposed to play roles in evading host responses (e.g., β-OH steroid oxidoreductase, eIF-2α, caspase recruitment domain-containing protein). Notably, the only homologue of IIV3-015R is a gene similarly located adjacent to the MCP gene in IIV-22, an IIV infecting the dipteran *Simulium variegatum* (15).

While certain features of the IIV-3 genome were most similar to IIV-6, others made clear the unique nature of IIV-3 and its distant relationship to IIV-6. These included differences in the nature and extent of repetitive DNA, genome composition, gene colinearity, and gene complement. Thirty-three predicted IIV-3 proteins are not encoded by other IVs (Fig. 1, yellow boxes) and, with the exception of IIV3-044L (protein kinase),

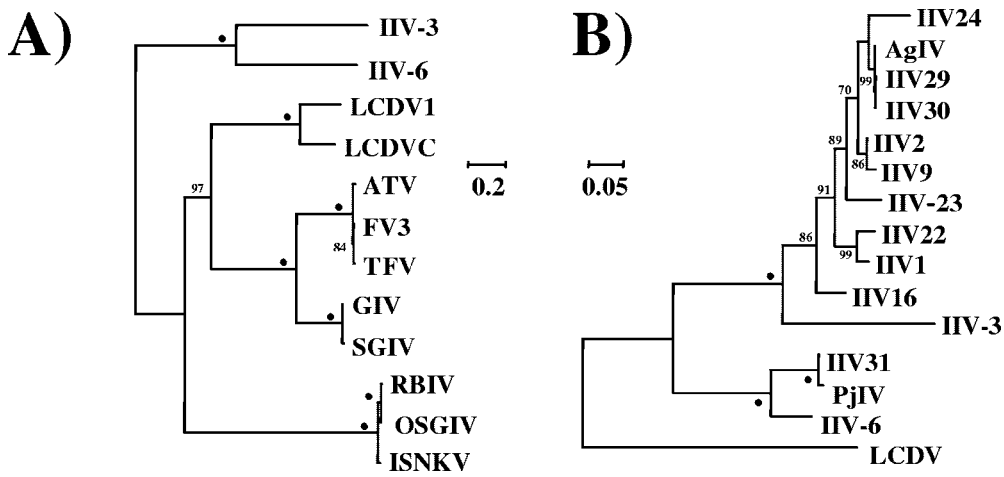

FIG. 3. Phylogenetic analysis of IIV-3 proteins (A) and IIV MCP (B). (A) Eleven conserved IIV-3 proteins (IIV3-009R, IIV3-014L, IIV3-048L, IIV3-055R, IIV3-076L, IIV3-087L, IIV3-088R, IIV3-090L, IIV3-101R, IIV3-120R, and IIV3-121R) were concatenated and aligned with similar data sets from other iridoviruses by using Kalign. The unrooted tree for 9,190 aligned characters was generated by the maximum likelihood tree search strategy, including the WAG model for correction for multiple substitutions, the four-category discrete gamma model for correction for among-site rate variation, 100 bootstrap replicates, and default settings, as implemented in Phyml. Bootstrap values greater than 70 are indicated at the appropriate nodes, and dots indicate values of 100. Sequences from the following viruses and accession numbers were compared: IIV-3, DQ643392; IIV-6, AF303741; LCDV-1, L63545; LCDV-C, AY380826; GIV, AY666015; SGIV, AY521625; FV-3, AY548484; ATV, AY150217; TFV, AF389451; ISKNV, AF371960; RBIV (strain KOR-TY1), AY532606; OSGIV, AY894343. (B) IIV3-014L was aligned with available IIV MCP sequences and LCDV MCP as an outgroup with Clustal, and the tree was generated as described above. Sequences from the following viruses and accession numbers were compared: IIV-24, AF042340; IIV-30, AF042336; IIV-29, AF042339; AgIV, AF042343; IIV-2, AF042335; IIV-9, AF025774; IIV-23, AF042342; IIV-22, M32799; IIV-1, M33542; IIV-16, AF025775; IIV-3, DQ643392; IIV-31, AF042337; PjIV, AF042338; IIV-6, AF303741; LCDV, AY849392. Scales indicate estimated distances. Similar topologies were obtained with maximum likelihood as implemented in MRBAYES and/or TREE-PUZZLE and with neighbor-joining and maximum-parsimony algorithms as implemented in PHYLO_WIN.

IIV3-053L (Rpb7), and IIV3-080R (MutT-like protein), they lacked similarity to any other protein. VIV gene homologues present in either IIV-3 or IIV-6, but not both, include IIV3-007R, which encodes a homologue of the FV-3 IE 31-kDa protein absent in IIV-6, and thymidylate synthase and dUTPase genes, present in IIV-6 and many VIVs but absent in IIV-3. While predicted IIV-3 proteins were most similar to IIV-6 homologues in pairwise comparisons, amino acid identities ranged only from 22 to 53%. Finally, phylogenetic analyses clearly separated IIV-3 from IIV-6 (Fig. 3). Analysis of a large, concatenated protein data set indicated that IIV-3 and IIV-6, while grouping together within the *Iridoviridae*, have genetic distances comparable to those between other IV genera, consistent with their classification into separate genera (Fig. 3A). Analysis of available IIV MCP data also indicated a clear separation between IIV-3 and IIV-6 and indicated that IIV-3, while unique, was more similar to a group of viruses which included IIV-22, consistent with IIV-3/IIV-22 colinearity at the IIV3-015/016 locus (Fig. 3B). Notably, several other members of this group (IIV-1, IIV-9, IIV-16) are currently classified as members of the genus *Iridovirus*. Given that here IIV-1, IIV-9, and IIV-16 were as phylogenetically distinct from IIV-6 (genus *Iridovirus*) as was IIV-3 (genus *Chloriridovirus*), reconsideration of current IIV taxonomy may be in order (Fig. 3B).

Conclusions. The genome sequence of IIV-3, sole member of the genus *Chloriridovirus*, has been determined. IIV-3 contains a number of novel genes which are likely important for infection of the natural mosquito host. Comparison between IIV-3, IIV-6, and available non-IIV-6 IIV gene sequences suggests that diversity within IIV may be greater than previously recognized. Genes common to all four IV genera, which include those encoding enzymes involved in RNA transcription, DNA replication, and protein processing, likely define the genetic core of the family *Iridoviridae*.

ACKNOWLEDGMENTS

We thank A. Lakowitz and A. Zsak for providing excellent technical assistance.

REFERENCES

1. **Afonso, C. L., E. R. Tulman, Z. Lu, E. Oma, G. F. Kutish, and D. L. Rock.** 1999. The genome of Melanoplus sanguinipes entomopoxvirus. J. Virol. **73:**533–552.
2. **Agresti, A., and M. E. Bianchi.** 2003. HMGB proteins and gene expression. Curr. Opin. Genet. Dev. **13:**170–178.
3. **Ahrens, C. H., M. N. Pearson, and G. F. Rohrmann.** 1995. Identification and characterization of a second putative origin of DNA replication in a baculovirus of Orgyia pseudotsugata. Virology **207:**572–576.
4. **Altschul, S. F., and D. J. Lipman.** 1990. Protein database searches for multiple alignments. Proc. Natl. Acad. Sci. USA **87:**5509–5513.
5. **Altschul, S. F., A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
6. **Armache, K. J., S. Mitterweger, A. Meinhart, and P. Cramer.** 2005. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J. Biol. Chem. **280:**7131–7134.
7. **Balakirev, M. Y., S. O. Tcherniuk, M. Jaquinod, and J. Chroboczek.** 2003. Otubains: a new family of cysteine proteases in the ubiquitin pathway. EMBO Rep. **4:**517–522.
8. **Bessler, J. B., J. Z. Torres, and V. A. Zakian.** 2001. The Pif1 subfamily of helicases: region-specific DNA helicases? Trends Cell Biol. **11:**60–65.
9. **Bideshi, D. K., S. Renault, K. Stasiak, B. A. Federici, and Y. Bigot.** 2003. Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. J. Gen. Virol. **84:**2531–2544.
10. **Bigot, Y., K. Stasiak, F. Rouleux-Bonnin, and B. A. Federici.** 2000. Characterization of repetitive DNA regions and methylated DNA in ascovirus genomes. J. Gen. Virol. **81:**3073–3082.
11. **Bonaldi, T., G. Langst, R. Strohner, P. B. Becker, and M. E. Bianchi.** 2003.

The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J. **21:**6865–6873.
12. **Borodovsky, A., H. Ovaa, N. Kolli, T. Gan-Erdene, K. D. Wilkinson, H. L. Ploegh, and B. M. Kessler.** 2002. Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. Chem. Biol. **9:**1149–1159.
13. **Brendel, V., P. Bucher, I. R. Nourbakhsh, B. E. Blaisdell, and S. Karlin.** 1992. Methods and algorithms for statistical analysis of protein sequences. Proc. Natl. Acad. Sci. USA **89:**2002–2006.
14. **Buchatsky, L. P.** 1977. An iridovirus from larvae of Culisetta annulata and Culex territans. Acta Virol. **21:**85–86.
15. **Cameron, I. R.** 1990. Identification and characterization of the gene encoding the major structural protein of insect iridescent virus type 22. Virology **178:**35–42.
16. **Castresana, J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17:**540–552.
17. **Chapman, H. C., T. B. Clark, and D. B. Woodard.** 1966. Additional mosquito hosts of the mosquito iridescent virus. J. Invertebr. Pathol. **8:**545–546.
18. **Choder, M.** 2004. Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. Trends Biochem. Sci. **29:**674–681.
19. **Clark, T. B., W. R. Kellen, and P. T. M. Lum.** 1965. A mosquito iridescent virus (MIV) from Aedes taeniorhynchus (Wiedemann). J. Invertebr. Pathol. **7:**519–521.
20. **Devereux, J., P. Haeberly, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
21. **Djupedal, I., M. Portoso, H. Spahr, C. Bonilla, C. M. Gustafsson, R. C. Allshire, and K. Ekwall.** 2005. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. Genes Dev. **19:**2301–2306.
22. **Do, J. W., C. H. Moon, H. J. Kim, M. S. Ko, S. B. Kim, J. H. Son, J. S. Kim, E. J. An, M. K. Kim, S. K. Lee, M. S. Han, S. J. Cha, M. S. Park, M. A. Park, Y. C. Kim, J. W. Kim, and J. W. Park.** 2004. Complete genomic DNA sequence of rock bream iridovirus. Virology **325:**351–363.
23. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
24. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5:**164–166.
25. **Fukaya, M., and S. Nasu.** 1966. A Chilo iridescent virus from the rice stem borer, Chilo suppresalis Walker (Lepidoptera:Pyralidae). Appl. Entomol. Zool. **1:**69.
26. **Galtier, N., M. Gouy, and C. Gautier.** 1996. SEAVIEW and PHYLO WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biol. Sci. **12:**543–554.
27. **Glover, J. N., R. S. Williams, and M. S. Lee.** 2004. Interactions between BRCT repeats and phosphoproteins: tangled up in two. Trends Biochem. Sci. **29:**579–585.
28. **Goorha, R.** 1982. Frog virus 3 DNA replication occurs in two stages. J. Virol. **43:**519–528.
29. **Goorha, R., G. Murti, A. Granoff, and R. Tirey.** 1978. Macromolecular synthesis in cells infected by frog virus 3. VIII. The nucleus is a site of frog virus 3 DNA and RNA synthesis. Virology **84:**32–50.
30. **Goorha, R., and K. G. Murti.** 1982. The genome of frog virus 3, an animal DNA virus, is circularly permuted and terminally redundant. Proc. Natl. Acad. Sci. USA **79:**248–252.
31. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
32. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
33. **Hall, D. W., and D. W. Anthony.** 1971. Pathology of a mosquito iridescent virus (MIV) infecting Aedes taeniorhynchus. J. Invertebr. Pathol. **18:**61–69.
34. **He, J. G., M. Deng, S. P. Weng, Z. Li, S. Y. Zhou, Q. X. Long, X. Z. Wang, and S.-M. Chan.** 2001. Complete genomic analysis of the mandarin fish infectious spleen and kidney necrosis iridovirus. Virology **291:**126–139.
35. **He, J. G., L. Lu, M. Deng, H. H. He, S. P. Weng, X. H. Wang, S. Y. Zhou, Q. X. Long, X. Z. Wang, and S.-M. Chan.** 2002. Sequence analysis of the complete genome of an iridovirus isolated from the tiger frog. Virology **292:**185–197.
36. **Hershko, A., and A. Ciechanover.** 1998. The ubiquitin system. Annu. Rev. Biochem. **67:**425–479.
37. **Home, W. A., S. Tajbakhsh, and V. L. Seligy.** 1990. Molecular cloning and characterization of a late Tipula iridescent virus gene. Gene **94:**243–248.
38. **Huang, X., and A. Madan.** 1999. CAP3: a DNA sequence assembly program. Genome Res. **9:**868–877.
39. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755.
40. **Iyer, L. M., E. V. Koonin, and L. Aravind.** 2002. Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSES transcription factors. Genome Biol. **3:**RESEARCH 0012.1-0012.11. [Online.] http://genomebiology.com/2002/3/3/RESEARCH/0012.
41. **Jakob, N. J., and G. Darai.** 2002. Molecular anatomy of Chilo iridescent virus genome and the evolution of viral genes. Virus Genes **25:**299–316.
42. **Jakob, N. J., K. Muller, U. Bahr, and G. Darai.** 2001. Analysis of the first complete DNA sequence of an invertebrate iridovirus: coding strategy of the genome of Chilo iridescent virus. Virology **286:**182–196.
43. **Jancovich, J. K., J. Mao, V. G. Chinchar, C. Wyatt, S. T. Case, S. Kumar, G. Valente, S. Subramanian, E. W. Davidson, J. P. Collins, and B. L. Jacobs.** 2003. Genomic sequence of a ranavirus (family Iridoviridae) associated with salamander mortalities in North America. Virology **316:**90–103.
44. **Jones, D. T., W. R. Taylor, and J. M. Thornton.** 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry **33:**3038–3049.
45. **Kang, W., M. Suzuli, Z. Evgueni, K. Okano, and S. Maeda.** 1999. Characterization of baculovirus repeated open reading frames (*bro*) in Bombyx mori nucleopolyhedrovirus. J. Virol. **73:**10339–10345.
46. **Keck, J. G., C. J. Baldick, and B. Moss.** 1990. Role of DNA replication in vaccinia virus gene expression: a naked template is required for transcription of three late trans-activator genes. Cell **61:**801–809.
47. **Kelly, D. C., M. D. Ayres, T. Lescott, J. S. Robertson, and G. M. Happ.** 1979. A small iridescent virus (type 29) isolated from Tenebro molitor: a comparison of its proteins and antigens with six other iridescent viruses. J. Gen. Virol. **42:**95.
48. **Krishna, T. S., X.-P. Kong, S. Gary, P. M. Burgers, and J. Kuriyan.** 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell **79:**1233–1243.
49. **Lassmann, T., and E. L. Sonnhammer.** 2005. Kalign—an accurate and fast multiple sequence alignment algorithm. BMC Bioinformatics **6:**298.
50. **Lavrukhin, O. V., J. M. Fortune, T. G. Wood, D. E. Burbank, J. L. Van Etten, N. Osheroff, and S. Lloyd.** 2000. Topoisomerase II from Chlorella virus PBCV-1. J. Biol. Chem. **275:**6915–6921.
51. **Linley, J. R., and H. T. Nielsen.** 1968. Transmission of mosquito iridescent virus in Aedes taeniorhynchus. I. Laboratory experiments. J. Invertebr. Pathol. **12:**7–16.
52. **Linley, J. R., and H. T. Nielsen.** 1968. Transmission of mosquito iridescent virus in Aedes taeniorhynchus. II. Experiments related to transmission in nature. J. Invertebr. Pathol. **12:**17–24.
53. **Lu, L., S. Y. Zhou, C. Chen, S. P. Weng, S.-M. Chan, and J. G. He.** 2005. Complete genome sequence analysis of an iridovirus isolated from the orange-spotted grouper, Epinephelus coiodes. Virology **339:**81–100.
54. **Marakova, K. S., L. Aravind, and E. V. Kooning.** 2000. A novel superfamily of predicted cysteine proteases from eukaryotes, viruses and Chlamydia pneumoniae. Trends Biochem. Sci. **25:**50–52.
55. **Marina, C. F., J. I. Arredondo-Jimenez, A. Castillo, and T. Williams.** 1999. Sublethal effects of iridovirus disease in a mosquito. Oecologia **119:**383–388.
56. **McKune, K., K. L. Richards, A. M. Edwards, R. A. Young, and N. A. Woychik.** 1993. Rpb7, one of two dissociable subunits of yeast RNA polymerase II, is essential for cell viability. Yeast **9:**295–299.
57. **McLennan, A. G.** 1999. The MutT motif family of nucleotide phosphohydrolases in man and human pathogens. Int. J. Mol. Med. **4:**79–89.
58. **McMillan, N. A., S. Davison, and J. Kalmakoff.** 1990. Comparison of the genomes of two sympatric iridescent viruses (types 9 and 16). Arch. Virol. **114:**277–284.
59. **McMillan, N. A., and J. Kalmakoff.** 1994. RNA transcript mapping of the Wiseana iridescent virus genome. Virus Res. **32:**343–352.
60. **Nachmias, B., Y. Ashhab, and D. Ben-Yehuda.** 2004. The inhibitor of apoptosis protein family (IAPs): an emerging therapeutic target in cancer. Semin. Cancer Biol. **14:**231–243.
61. **Nakai, K., and P. Horton.** 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. **24:**34–36.
62. **Parks, W. C., C. L. Wilson, and Y. S. Lopez-Boado.** 2004. Matrix metalloproteinases as modulators of inflammation and innate immunity. Nat. Rev. **4:**617–629.
63. **Pearson, W. R.** 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. **183:**63–98.
64. **Salzberg, S. L., A. L. Delcher, S. Kasif, and O. White.** 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. **26:**544–548.
65. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
66. **Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler.** 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18:**502–504.
67. **Shors, T., J. G. Keck, and B. Moss.** 1999. Down regulation of gene expression by the vaccinia virus D10 protein. J. Virol. **73:**791–796.
68. **Snijder, E. J., A. L. Wassenaar, W. J. Spaan, and A. E. Gorbalenya.** 1995. The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. J. Biol. Chem. **270:**16671–16676.
69. **Song, W. J., Q. W. Qin, J. Qiu, C. H. Huang, F. Wang, and C. L. Hew.** 2004. Functional genomics analysis of Singapore grouper iridovirus: complete sequence determination and proteomic analysis. J. Virol. **78:**12576–12590.
70. **Sonnhammer, E. L., S. R. Eddy, E. Birney, A. Bateman, and R. Durbin.** 1998.

Pfam: multiple sequence alignments and HMM-profiles of protein domains. Nucleic Acids Res. **26:**320–322.
71. **Stasiak, K., M.-V. Demattei, B. A. Federici, and Y. Bigot.** 2000. Phylogenetic position of the DpAV-4a ascovirus DNA polymerase among viruses with a large double-stranded DNA genome. J. Gen. Virol. **81:**3059–3072.
72. **Subramanian, A. R., J. Weyer-Menkhoff, M. Kaufmann, and B. Morgenstern.** 2005. DIALIGN-T: an improved algorithm for segment-based multiple sequence alignment. BMC Bioinformatics **6:**66.
73. **Tan, W. G. H., T. J. Barkman, V. G. Chinchar, and K. Essani.** 2004. Comparative genomic analysis of frog virus 3, type species of the genus Ranavirus (family Iridoviridae). Virology **323:**70–84.
74. **Theilmann, D. A., and S. Stewart.** 1992. Tandemly repeated sequence at the 3′ end of the IE-2 gene of the baculovirus Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus is an enhancer element. Virology **187:**97–106.
75. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
76. **Tidona, C. A., and G. Darai.** 1997. The complete DNA sequence of lymphocystis disease virus. Virology **230:**207–216.
77. **Todone, F., P. Brick, F. Werner, R. O. J. Weinziert, and S. Onesti.** 2001. Structure of an archeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. Mol. Cell **8:**1137–1143.
78. **Tsai, C.-T., J.-W. Ting, M.-H. Wu, M.-F. Wu, I.-C. Guo, and C.-Y. Chang.** 2005. Complete genomic sequence of the grouper iridovirus and comparison of genomic organization with those of other iridoviruses. J. Virol. **79:**2010–2023.
79. **Ueda, K., A. Yamashita, J. Ishikawa, M. Shimada, T. O. Watsuji, Morimura, H. Ikeda, M. Hattori, and T. Beppu.** 2004. Genome sequence of Symbiobacterium thermophilum, an uncultivable bacterium that depends on microbial commensalism. Nucleic Acids Res. **32:**4937–4944.
80. **Undeen, A. H., and T. Fukuda.** 1994. Effects of host resistance and injury on the susceptibility of Aedes taeniorhynchus to mosquito iridescent virus. J. Am. Mosq. Control Assoc. **10:**64–66.
81. **Van Wart, H. E., and H. Birkedal-Hansen.** 1990. The cystein switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. Proc. Natl. Acad. Sci. USA **87:**5578–5582.
82. **Wagner, G. W., and J. D. Paschke.** 1977. A comparison of the DNA of R and T strains of mosquito iridescent virus. Virology **81:**298–308.
83. **Wang, J. W., R. Q. Deng, X. Z. Wang, Y. S. Huang, K. Xing, J. H. Feng, J. G. He, and Q. X. Long.** 2003. Cladistic analysis of iridoviruses based on protein and DNA sequences. Arch. Virol. **148:**2181–2194.
84. **Webby, R. J., and J. Kalmakoff.** 1998. Sequence comparison of the major capsid protein gene from 18 diverse iridoviruses. Arch. Virol. **143:**1949–1966.
85. **Wei, W., D. Dorjsuren, Y. Lin, W. Qin, T. Nomura, N. Hayashi, and S. Murakami.** 2001. Direct interaction between the subunit RAP30 of transcription factor IIF (TFIIF) and RNA polymerase subunit 5, which contributes to the association between TFIIF and RNA polymerase II. J. Biol. Chem. **276:**12266–12273.
86. **Wertz, I. E., K. M. O'Rourke, H. Zhou, M. Eby, L. Aravind, S. Seshagiri, P. Wu, C. Wiesmann, R. Baker, D. L. Boone, A. Ma, E. V. Koonin, and V. M. Dixit.** 2004. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-κB signalling. Nature **430:**694–699.
87. **Wesley, R. D., J. C. Quintero, and C. A. Mebus.** 1984. Extraction of viral DNA from erythrocytes of swine with acute African swine fever. Am. J. Vet. Res. **45:**1127–1131.
88. **Williams, T.** 1994. Comparative studies of iridoviruses: further support for a new classification. Virus Res. **33:**99–121.
89. **Williams, T.** 1998. Invertebrate iridescent viruses. Plenum Press, New York, N.Y.
90. **Williams, T.** 1996. The iridoviruses. Adv. Virus Res. **46:**345–412.
91. **Williams, T., and J. S. Cory.** 1994. Proposals for a new classification of iridescent viruses. J. Gen. Virol. **75:**1291–1301.
92. **Willis, D. B., and A. Granoff.** 1985. *Trans* activation of an immediate-early frog virus 3 promoter by a virion protein. J. Virol. **56:**495–501.
93. **Woodard, D. B., and H. C. Chapman.** 1968. Laboratory studies with the mosquito iridescent virus (MIV). J. Invertebr. Pathol. **11:**296–301.
94. **Xeros, N.** 1954. A second virus disease of the leather jacket, Tipula paludosa. Nature **174:**562.
95. **Yan, X., N. H. Olson, J. L. Van Etten, M. Bergoin, M. G. Rossmann, and T. S. Baker.** 2000. Structure and assembly of large lipid-containing dsDNA viruses. Nat. Struct. Biol. **7:**101–103.
96. **Yeo, M., P. S. Lin, M. E. Dahmus, and G. N. Gill.** 2003. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. J. Biol. Chem. **278:**26078–26085.
97. **Zemskov, E. A., W. Kang, and S. Maeda.** 2000. Evidence for nucleic acid binding ability and nucleosome association of *Bombyx mori* nucleopolyhedrovirus BRO proteins. J. Virol. **74:**6784–6789.
98. **Zhang, Q.-Y., F. Xiao, J. Xie, Z.-Q. Li, and J.-F. Gui.** 2004. Complete genomic sequence of lymphocystis disease virus isolated from China. J. Virol. **78:**6982–6994.