ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

VIROLOGY

Virology 316 (2003) 90–103

www.elsevier.com/locate/yviro

# Genomic sequence of a ranavirus (family *Iridoviridae*) associated with salamander mortalities in North America

James K. Jancovich,[a] Jinghe Mao,[b] V. Gregory Chinchar,[c] Christopher Wyatt,[d] Steven T. Case,[d] Sudhir Kumar,[a] Graziela Valente,[a] Sankar Subramanian,[a] Elizabeth W. Davidson,[a] James P. Collins,[a] and Bertram L. Jacobs[a,*]

[a] *School of Life Sciences, Arizona State University, Tempe, AZ 85287-4601, USA*
[b] *Division of Natural Sciences, Tougaloo College, Jackson, MS 39174, USA*
[c] *Department of Microbiology, University of Mississippi Medical Center, Jackson, MS 39216, USA*
[d] *Department of Biochemistry, University of Mississippi Medical Center, Jackson, MS 39216, USA*

Received 2 May 2003; returned to author for revision 3 July 2003; accepted 1 August 2003

**Abstract**

Disease is among the suspected causes of amphibian population declines, and an iridovirus and a chytrid fungus are the primary pathogens associated with amphibian mortalities. *Ambystoma tigrinum* virus (ATV) and a closely related strain, Regina ranavirus (RRV), are implicated in salamander die-offs in Arizona and Canada, respectively. We report the complete sequence of the ATV genome and partial sequence of the RRV genome. Sequence analysis of the ATV/RRV genomes showed marked similarity to other ranaviruses, including tiger frog virus (TFV) and frog virus 3 (FV3), the type virus of the genus *Ranavirus* (family *Iridoviridae*), as well as more distant relationships to lymphocystis disease virus, *Chilo* iridescent virus, and infectious spleen and kidney necrosis virus. Putative open reading frames (ORFs) in the ATV sequence identified 24 genes that appear to control virus replication and block antiviral responses. In addition, >50 other putative genes, homologous to ORFs in other iridoviral genomes but of unknown function, were also identified. Sequence comparison performed by dot plot analysis between ATV and itself revealed a conserved 14-bp palindromic repeat within most intragenic regions. Dot plot analysis of ATV vs RRV sequences identified several polymorphisms between the two isolates. Finally, a comparison of ATV and TFV genomic sequences identified genomic rearrangements consistent with the high recombination frequency of iridoviruses. Given the adverse effects that ranavirus infections have on amphibian and fish populations, ATV/RRV sequence information will allow the design of better diagnostic probes for identifying ranavirus infections and extend our understanding of molecular events in ranavirus-infected cells.


*Keywords:* Ambystoma tigrinum virus; Regina ranavirus; Amphibian; Iridovirus; Dot plot

## Introduction

Infectious disease is among the suspected causes of amphibian population declines (Collins and Storfer, 2003). Several iridoviruses (genus *Ranavirus*) and a chytrid fungus (*Batrachochytrium dendrobatides*) are the primary pathogens associated with amphibian mortalities (Daszak et al., 1999). The family *Iridoviridae* has four genera that infect invertebrates (*Chloriridovirus, Iridovirus*) and vertebrates (*Lymphocystivirus, Ranavirus*) (Williams et al., 2000). Iridoviruses posess icosahedral symmetry and vary in size from 120 to 200 nm in diameter. While all iridoviruses possess circularly permutated, double-stranded DNA genomes that vary from ~140 to 303 kilobase pairs (kbp), only members of the ranavirus and lymphocystisvirus genera have highly methylated genomes (Williams et al., 2000). Four iridoviral genomes are completely sequenced: lymphocystis disease virus (LCDV) (genus *Lymphocystivirus;* Tidona and Darai, 1997), insect iridovirus IIV [*Chilo* iridescent virus (CIV)] (genus *Iridovirus;* Jakob et al., 2001.), tiger frog virus (TFV) (genus *Ranavirus;* He et al., 2002),

* Corresponding author. Fax: +1-480-965-0098.
*E-mail address:* bjacobs@asu.edu (B.L. Jacobs).

0042-6822/$ – see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2003.08.001

Table 1
Summary of genomic sequence information for five viral species representing four genera within the family *Iridoviridae*

| Genus: | *Iridovirus* | *Ranavirus* | *Ranavirus* | *Lymphocystivirus* | Unassigned |
|---|---|---|---|---|---|
| Species: | CIV | ATV | TFV | LCDV | ISKNV |
| Genome size (bp) | 212,482 | 106,332 | 105,057 | 102,653 | 111,362 |
| GC (%) | 28.6 | 54 | 55 | 29.1 | 54.8 |
| No. putative ORFs | 468 | 96 | 105 | 195 | 124 |
| ORF size (AA) | 40–2432 | 32–1294 | 40–1294 | 40–1199 | 40–1208 |
| Accession No. | AF303741 | AY150217 | NC003407 | NC001824 | NC003494 |

and infectious spleen and kidney necrosis virus (ISKNV) (genus unassigned; He et al., 2001). These iridoviruses infect insect, fish, and frogs. Because of their importance in amphibian epidemics we sequenced the genomes of two closely related viruses, *Ambystoma tigrinum* virus (ATV) and Regina ranavirus (RRV), linked to salamander deaths throughout western North America. We report our findings of the complete genomic sequence of ATV and then compare it to the partial sequence of RRV and the completed genomes of other iridoviruses.

## Results and discussion

### *Genomic properties of ATV*

Over 1500 ATV insert-containing clones were isolated and about 1.5 million bp were sequenced, for an average level of redundancy within ATV of about eightfold. About 1300 individual sequences were assembled into a continuous sequence using Sequencher. The completed ATV sequence was shown to be circularly permutated, a feature characteristic of iridoviruses (Williams et al., 2000), and contained 106,332 bp and a 54% G + C content. The genome size and G + C content of ATV (Table 1) are comparable to those of TFV (He et al., 2002) and ISKNV (He et al., 2001). In contrast, while the ATV genome is slightly larger than that of LCDV (102,653 bp), the G + C content of LCDV (29.1%) is much lower (Tidona and Darai, 1997). Moreover, ATV is far removed in size and G + C content from CIV, whose genome (212,482 bp) is almost twice as large as ATV's and whose G + C content (28.6%) is markedly lower (Jakob et al., 2001). The ATV genomic sequence was deposited in the NCBI database (http://www.ncbi.nlm.nih.gov/) and the Accession number is AY150217. RRV *Pst*I fragments were deposited in the NCBI database under Accession Nos. AF368228 (1.3 kbp), AF368230 (1.6 kbp), AF397203 (1.9 kbp), AF368231 (3.0 kbp), AY029323.1 (4.9 kbp), AF368229 (6.7 kbp), and AF367980 (8.1 kbp).

### *Open reading frame analysis*

Putative ORFs were identified in the ATV genome using the four methods described under Materials and methods. Only ORFs detected by all four annotation methods or with homology to other known iridovirus ORFs were considered for further analysis. A diagramatic representation of the ATV genome is shown in Fig. 1, and Table 2 lists the predicted ORFs and their sizes, homologous proteins, and predicted functions. The 96 ATV ORFs range in size from 32 to 1294 amino acids in length, which is similar to the size range for TFV (Table 1). The predicted sizes and numbers of ORFs are greater for CIV, while the genomes of iridoviruses associated with cold-blooded vertebrates are predicted to contain relatively the same size and number of ORFs as the ATV genome (Table 1). The smallest ATV ORF, 56R, is 8 amino acids smaller than the 40-amino-acid length requirement established in the methods for predicting ORFs; however, this ORF has homology with TFV ORF 26R, so it was included in our analyses. Consistent with the genetic organization of other iridoviruses, ATV ORFs are, for the most part, non-overlapping. However, 9 overlapping ORFs were predicted in the ATV genome based on the above criteria (Fig. 1, Table 2). Two of these, 6bL and 43bL, have homology with CIV sequences and are located within subunits of the RNA polymerase. Two other relatively large overlapping ORFs opposite 5 small ones are found between ATV ORFs 31R and 38R (Fig. 1, Table 2). Because it is difficult to determine which of the ATV ORFs within this region are "real," all of them are included in further analyses. Open reading frame 41R (82% homology with TFV ORF 69L) and ORF 40L, which contains a CARD-like caspase recruitment domain, overlap (Table 2). Both ORFs are included as they are found in ATV and TFV (He et al., 2002). Open reading frame 61bR is found on the same strand as 61R, but in a different reading frame. Both were identified by more than one method and are included as ATV ORFs. The start codons for ATV ORFs 73L and 86R have some overlap with ORFs 74L and 85R, respectively. All of these ATV ORFs have homology to TFV ORFs (Table 2). Finally, ORFs 72L and 72bR are identical in length but opposite in direction (Fig. 1). Homology between ATV ORF 72L and TFV ORF 48L or RRV Pst 3.0 is relatively low (56 and 78%, respectively; Table 2), and both ORFs contain repeat sequences making it difficult to determine the correct ORF. Functional analysis will be required to elucidate the role of the overlapping ORFs. In addition to the nine predicted overlapping genes, two relatively long intergenic regions, between ORFs 54R and 55R and ORFs

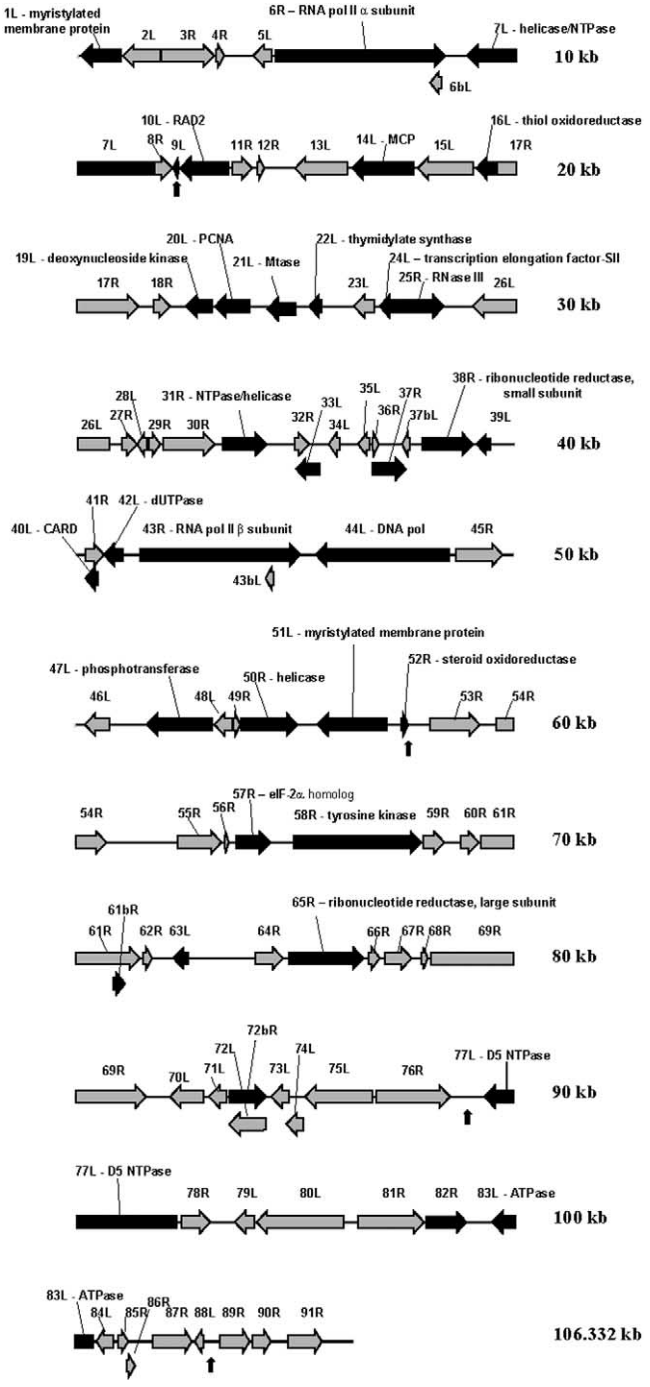


Fig. 1. ATV genome organization. Predicted ORFs are indicated by their location, orientation, and putative size. Dark arrows represent ORFs with predicted function or homology to other iridoviral ORFs, while light arrows represent predicted ORFs in ATV with unknown function or without homology to other iridoviral ORFs. Small vertical arrows indicate positions of inversions relative to the TFV genome.

63L and 64R, were identified (Fig. 1). Long intergenic regions were not observed in other vertebrate iridovirus genomic sequences (Tidona and Darai, 1997; He et al., 2001, 2002). These regions in ATV contain small potential ORFs; however, the ORFs were not consistently identified by the four annotation methods, and the ORFs were not conserved among iridoviruses. These regions may contain ORFs in ATV, but further analysis is required to identify coding sequences within these regions.

A comparison of codon usage by ATV and other iridoviruses is shown in Table 3. The principal codons used by ATV are identical to those used by TFV and ISKNV except for those encoding Arg and Ser (He et al., 2001, 2002). In contrast, marked changes are seen between ATV and LCDV, but these occur primarily at the third ("wobble") nucleotide position and reflect the G + C content of each virus. Thus while the third base in ATV codons is generally a G or C, in LCDV the third base is predominantly an A or T.

ATV showed >95% identity to all RRV ORFs except for one (72L). After RRV, the highest level of ORF identity was with TFV, followed by LCDV, CIV, and ISKNV, respectively (Table 2). Levels of sequence identity between ATV and TFV ORFs, the only complete ranavirus genomic sequence currently available, were generally >90%, although matches as low as 56% were detected. In contrast, levels of identity to LCDV, CIV, and ISKNV ORFs were markedly lower. For example, within the viral DNA polymerase gene identities of 98% (RRV), 97% (TFV), 37% (LCDV), 35% (RSIV), 36% (ISKNV), and 25% (CIV) were noted. The 96 predicted ATV ORFs can be divided into four groups: those with homology to putative viral or cellular replicative proteins; those potentially involved with immune modulation or pathogenesis; those with homology to other iridovirus ORFs, but of unknown function; and those of unknown function with no homology to other iridovirus sequences (Table 4).

The ATV genome contains homologs of most of the iridovirus genes associated with DNA replication and modification, nucleotide metabolism, protein synthesis, and virus structure (He et al., 2002). The functions of these genes were described previously (He et al., 2002) and will not be discussed in detail in this paper. However, despite general agreement among iridovirus ORFs, three marked differences were observed between the predicted ATV genes and those described for other ranaviruses. First, ORF 10L in ATV codes for the DNA repair enzyme, RAD2. The predicted AUG, or start site, for this ATV gene begins at amino acid position 28 in the predicted TFV RAD2 protein (ORF 101) (He et al., 2002). Alignment of the upstream sequence from the start site of the RAD2 gene between ATV and TFV reveals that the ATV sequence is identical to that of TFV except for the addition of 1 bp in the TFV sequence (data not shown). The addition of a basepair in the 5′ upstream region of ATV changes the reading frame to initiate RAD2 protein synthesis at the same start site as described in TFV. The ATV sequence in the region upstream of the RAD2 start site was confirmed by sequencing multiple clones containing this region and by sequencing both ATV and RRV clones. In addition, the RAD2 start site in ATV corresponds to start sites in LCDV, ISKNV, and CIV, suggesting that a

Table 2
Predicted open reading frames in *Ambystoma tigrinum* virus

| ORF | Nucleotide position | Number of amino acids | MW | Conserved region or signature | Accession No. of conserved regions[a] | Predicted function | Corresponding iridovirus ORFs[b] | Percentage of identity to iridovirus ORFs[c] | Accession No. of iridovirus ORFs[a] |
|---|---|---|---|---|---|---|---|---|---|
| 1L | 73–981 | 303 | 32,771 | DUF230 poxvirus protein—unknown function | pfam03003 | Myristylated membrane protein | TFV 2 | 97 | |
| | | | | | | | LCDV 29 | 36 | NP_078745.1 |
| | | | | | | | CIV 337L | 32 | NP_149800.1 |
| 2L[d] | 1,022–1,858 | 279 | 31,104 | | | | LCDV 14 | 25 | NP_078619.1 |
| | | | | | | | LCDV 25 | 28 | NP_078687.1 |
| 3R | 1,892–3,103 | 404 | 44,566 | | | | TFV 4 | 96 | |
| | | | | | | | CIV 229L | 23 | NP_149692.1 |
| 4R | 3,149–3,328 | 60 | 6,516 | | | | TFV 5 | 88 | |
| 5L | 3,997–4,416 | 140 | 14,876 | | | | TFV 7L | 73 | |
| 6R | 4,495–8,376 | 1,294 | 151,050 | RNA polymerase $\alpha$ subunit | pfam00623 | RNA pol II $\alpha$ subunit | TFV 8 | 96 | NP_571990.1 |
| | | | | | | | LCDV 1 | 43 | NP_078624.1 |
| | | | | RNA polymerase 1 A subunit N-terminal | smart00663 | | ISKNV 28L | 39 | NP_612250.1 |
| | | | | RNA polymerase A/β′/A″ subunit | pfam01854 | | | | |
| 6bL[d] | 8,040–8,300 | 87 | 9,051 | | | | CIV 344R | 40 | NP_149807.1 |
| 7L | 8,882–11,725 | 948 | 106,187 | SNF-2 family terminal domain | pfam00176 | Helicase NTPase | RRV Pst 8.1 | 99 | AAK53744.1 |
| | | | | DEAD-like helicase | smart00487 | | TFV 9L | 97 | NP_571991.1 |
| | | | | Helicase C-terminal domain | pfam00271 | | | | |
| 8R | 11,741–12,151 | 137 | 14,896 | | | | TFV 10 | 99 | |
| 9L | 12,161–12,307 | 49 | 6,033 | | | | | | |
| 10L | 12,328–13,419 | 364 | 40,626 | XPG-I region | pfam00867 | DNA repair enzyme RAD2 | RRV Pst 8.1 | 99 | AAK53745 |
| | | | | Xeroderma pigmentosum G–N region | smart00485 | | TFV 101 | 97 | NP_572012.1 |
| | | | | | | | LCDV 34 | 38 | NP_078767.1 |
| | | | | XPG N-terminal domain | | | ISKNV | 35 | NP_612249.1 |
| | | | | Xeroderma pigmentosum G–I region 5′-3′ exonuclease | pfam00752 | | CIV 369L | 26 | NP_149832.1 |
| | | | | | smart00484 | | | | |
| | | | | Helix-hairpin-helix class 2 (Pol1 family) motifs | smart00475 | | | | |
| | | | | | smart00279 | | | | |
| 11R | 13,512–13,976 | 155 | 17,725 | | | | RRVPst 8.1 | 100 | AAK53746 |
| | | | | | | | TFV 100 | 95 | |
| | | | | | | | ISKNV 86L | 50 | NP_612308.1 |
| | | | | | | | LCDV 71 | 43 | NP_078627.1 |
| | | | | | | | CIV 307L | 39 | NP_149770.1 |
| 12R | 14,091–14,243 | 51 | 5,167 | | | | TFV 99L | 87 | |
| 13L | 14,952–16,136 | 395 | 45,587 | | | | RRV Pst 8.1 | 99 | AAK53747.1 |
| | | | | | | | TFV 97 | 94 | NP_572011.1 |
| | | | | | | | FV3 ICP-46 | 89 | P14358 |
| | | | | | | | ISKNV 115R | 24 | NP_612337.1 |
| | | | | | | | LCDV 27 | 23 | NP_078648.1 |
| | | | | | | | CIV 393L | 20 | NP_149856.1 |
| 14L | 16,263–17,651 | 463 | 50,000 | | | Major capsid protein (MCP) | TFV 96 | 96 | NP_572010.1 |
| | | | | | | | FV3 | 95 | Q67473 |
| | | | | | | | LCDV 91 | 51 | NP_044812.1 |
| | | | | | | | CIV 274L | 46 | NP_149737.1 |
| | | | | | | | ISKNV 6L | 46 | NP_612228.1 |
| 15L | 17,747–19,018 | 424 | 48,821 | | | | TFV 95 | 80 | |
| 16L | 19,089–19,538 | 150 | 16,527 | | | Thiol oxidoreductase | TFV 94 | 96 | |
| | | | | | | | ISKNV 43L | 42 | NP_612265.1 |
| | | | | | | | LCDV 79 | 39 | NP_078699.1 |
| | | | | | | | CIV 347L | 30 | NP_149810.1 |
| 17R | 19,571–21,394 | 608 | 65,841 | | | | TFV 93L | 93 | |
| | | | | | | | LCDV 11 | 28 | NP_078643.1 |
| 18R | 21,742–22,116 | 125 | 14,258 | | | | TFV 92L | 78 | |
| 19L[d] | 22,482–23,066 | 195 | 22,142 | Deoxynucleoside kinase | pfam01712 | Deoxynucleoside kinase | LCDV 60 | 37 | NP_078725.1 |
| | | | | Thymidylate kinase | pfam02223 | | ISKNV 32R | 28 | NP_612254.1 |
| | | | | | | | CIV 143R | 26 | NP_149606.1 |
| 20L | 23,144–23,923 | 260 | 27,706 | | | Proliferating cell nuclear antigen (PCNA) | TFV 90 | 95 | |
| | | | | | | | LCDV 45 | 24 | NP_078615.1 |
| | | | | | | | ISKNV 112R | 24 | NP_612334.1 |
| 21L | 24,271–24,912 | 214 | 24,779 | | | Cytosine DNA methyltransferase | TFV 89 | 95 | NP_572009.1 |
| | | | | | | | FV3 | 95 | AAA86959.1 |
| | | | | | | | LCDV 51 | 56 | NP_078617.1 |
| | | | | | | | ISKNV 46L | 49 | NP_612268.1 |
| 22L | 25,280–25,564 | 95 | 10,975 | Thymidylate synthase | pfam00303 | Thymidylate synthase | TFV 88 | 95 | NP_572008.1 |
| 23L | 26,292–26,762 | 157 | 17,422 | | | | TFV 87 | 94 | NP_572007.1 |
| | | | | | | | FV3 p18K | 92 | P03298 |
| 24L | 26,893–27,168 | 92 | 10,422 | C2C2 Zinc finger; nucleic acid binding motif in transcriptional elongation factor TFIIS and RNA polymerases | smart00440 | Transcription elongation factor-SII | TFV 86 | 96 | NP_572006.1 |
| | | | | | | | LCDV 105 | 40 | NP_078754.1 |
| | | | | | | | CIV 349L | 40 | NP_149812.1 |
| | | | | Transcription factor S-II | pfam01096 | | | | |
| 25R | 27,224–28,342 | 373 | 40,647 | Ribonuclease III family | smart00535 | RNase III | TFV 85 | 95 | NP_572005.1 |
| | | | | RNase3 domain | pfam00636 | | LCDV 44 | 44 | NP_078726.1 |

*(continued on next page)*

Table 2 *(continued)*

| ORF | Nucleotide position | Number of amino acids | MW | Conserved region or signature | Accession No. of conserved regions[a] | Predicted function | Corresponding iridovirus ORFs[b] | Percentage of identity to iridovirus ORFs[c] | Accession No. of iridovirus ORFs[a] |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | CIV 142R | 32 | NP_149605.1 |
| | | | | | | | ISKNV 87R | 29 | NP_612309.1 |
| 26L | 29,002–30,729 | 576 | 63,947 | | | | TFV 84 | 86 | |
| 27R | 31,009–31,353 | 115 | 12,833 | | | | RRV Pst 1.9 | 99 | AAK84402.1 |
| | | | | | | | TFV 82L | 95 | |
| | | | | | | | LCDV 102 | 30 | NP_078638.1 |
| 28L | 31,356–31,574 | 73 | 7,942 | | | | TFV 81 | 95 | |
| 29R | 31,637–31,888 | 84 | 9,223 | Possible membrane-associated motif in LPS-induced tumor necrosis factor alpha factor (LITAF), also known as PIG7, and other animal proteins | smart00714 | | TFV 80 | 90 | |
| 30R | 31,947–33,125 | 393 | 41,934 | | | | TFV 79L | 80 | |
| 31R | 33,319–34,308 | 330 | 36,864 | | | NTPase/helicase | TFV 78L | 93 | NP_572004.1 |
| | | | | | | | LCDV 36 | 31 | NP_078700.1 |
| 32R | 34,970–35,290 | 107 | 11,331 | | | | TFV 77 | 95 | |
| 33L[d] | 34,979–35,533 | 185 | 20,217 | | | | | | |
| 34L | 35,745–35,978 | 78 | 8,502 | | | | TFV 76 | 91 | |
| 35L | 36,414–36,677 | 88 | 9,408 | | | | TFV 74 | 96 | |
| | | | | | | | LCDV 124 | 26 | NP_078755.1 |
| 36R | 36,736–36,888 | 51 | 5,975 | | | | TFV 73 | 82 | |
| 37R[d] | 36,752–37,507 | 252 | 27,525 | | | | | | |
| 37bL | 37,419–37,592 | 58 | 6,739 | | | | TFV 72 | 88 | |
| 38R | 37,876–39,036 | 387 | 43,970 | Ribonucleotide reductase, small chain | pfam00268 | Ribonucleotide reductase, small subunit | TFV 71L | 96 | NP_572003.1 |
| | | | | | | | LCDV 26 | 50 | NP_078636.1 |
| | | | | | | | CIV 376L | 34 | NP_149839.1 |
| 39L[d] | 39,115–39,459 | 115 | 12,428 | | | | | | |
| 40L[d] | 40,208–40,492 | 95 | 10,331 | CARD, caspase recruitment domain | pfam00619 smart00114 | Apoptosis signaling | | | |
| 41R | 40,224–40,631 | 136 | 14,658 | | | | TFV 69L | 82 | |
| 42L | 40,634–41,068 | 145 | 15,269 | dUTPase | pfam00692 | dUTPase | TFV 68 | 95 | NP_572002.1 |
| | | | | | | | CIV 438L | 42 | NP_149901.1 |
| 43R | 41,447–45,109 | 1,221 | 133,772 | RNA polymerase $\beta$ subunit | pfam00562 | RNA pol II $\beta$ subunit | RRV Pst 3.8 | 99 | AAK84400.1 |
| | | | | | | | TFV 65L | 95 | NP_572001.1 |
| | | | | | | | LCDV 3 | 45 | NP_078633.1 |
| | | | | | | | ISKNV 34R | 39 | NP_612256.1 |
| 43bL[d] | 44,323–44,514 | 64 | 6,747 | | | | CIV 430R | 53 | NP_149893.1 |
| 44L | 45,474–48,512 | 1,013 | 114,403 | DNA polymerase type B family | smart00486 pfam00136 | DNA polymerase | RRV Pst 3.8 | 98 | AAK84401.1 |
| | | | | | | | TFV 63 | 97 | NP_572000.1 |
| | | | | DNA polymerase family B, exonuclease domain | pfam03104 | | RRV Pst 1.6 | 95 | AAK54493.1 |
| | | | | | | | LCDV 5 | 37 | NP_078724.1 |
| | | | | DNA polymerase type B, organellar and viral | pfam03175 | | ISKNV 1L | 36 | NP_612241.1 |
| | | | | | | | CIV 37L | 33, 25 | NP_149500.1 |
| 45R | 48,676–49,731 | 352 | 40,043 | | | | TFV 62 | 94 | |
| 46L | 50,219–50,770 | 184 | 20,618 | | | | TFV 61 | 96 | |
| | | | | | | | CIV 170L | 30 | NP_149633.1 |
| 47L | 51,637–53,130 | 498 | 53,689 | | | Phosphotransferase | TFV 59 | 96 | |
| | | | | | | | ISKNV 13R | 28 | NP_612235.1 |
| | | | | | | | LCDV 17 | 24 | NP_078729.1 |
| 48L | 53,175–53,576 | 134 | 15,271 | | | | TFV 58 | 95 | |
| | | | | | | | LCDV 83 | 26 | NP_078686.1 |
| 49R | 53,613–53,759 | 49 | 5,238 | | | | TFV 57L | 95 | |
| 50R | 53,770–55,062 | 431 | 47,244 | DEAD-like helicase superfamily | smart00487 | Helicase | TFV 56L | 96 | NP_571999.1 |
| | | | | DEAD/DEAH box helicase | pfam00270 | | TFV 9L | 96 | " |
| | | | | SNF2 family N-terminal domain | pfam00176 | | CIV 161L | 28 | NP_149624.1 |
| 51L | 55,525–57,102 | 526 | 56,220 | Glycosylhydrolase family 3 N-terminal domain | pfam00933 | Myristylated membrane protein/glycosyl hydrolase | RRV Pst 8.1 | 99 | AAK54492.1 |
| | | | | | | | TFV 55 | 97 | |
| | | | | | | | CIV 118L | 34 | NP_149581.1 |
| | | | | | | | ISKNV 7L | 31 | NP_612229.1 |
| | | | | | | | LCDV 20 | 29 | NP_078665.1 |
| | | | | | | | CIV 458R | 22 | NP_149921.1 |
| 52R | 57,441–57,599 | 53 | 5,852 | | | 3-$\beta$-hydroxy-$\Delta$5-C27-steroid oxidoreductase-like | TFV 54L | 72 | NP_571998.1 |
| 53R | 58,082–59,227 | 382 | 42,777 | | | | TFV 23 | 92 | |
| 54R | 59,613–60,707 | 365 | 41,047 | | | | TFV 24 | 89 | |
| 55R | 62,328–63,332 | 335 | 38,365 | | | | TFV 25 | 97 | NP_571993.1 |
| | | | | | | | FV3 31KD | 96 | P18178 |
| | | | | | | | LCDV 49 | 34 | NP_078713.1 |
| | | | | | | | ISKNV 118L | 24 | NP_612340.1 |
| 56R | 63,402–63,497 | 32 | 3,590 | | | | TFV 26 | 73 | |
| 57R | 63,659–64435 | 259 | 28,418 | Ribosomal protein S1-like RNA binding domain factor-2 | smart00316 | Translational control-eif2$\alpha$ homolog | RRV Pst 1.3 | 100 | AAK54491.1 |
| | | | | | | | FV3 | 94 | AAD38359 |

Table 2 *(continued)*

| ORF | Nucleotide position | Number of amino acids | MW | Conserved region or signature | Accession No. of conserved regions[a] | Predicted function | Corresponding iridovirus ORFs[b] | Percentage of identity to iridovirus ORFs[c] | Accession No. of iridovirus ORFs[a] |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | TFV 27 | 92 | NP_571994.1 |
| 58R | 64,968–67,877 | 970 | 107,166 | CAP10, putative lipopolysaccharide-modifying enzyme | smart00672 | Tyrosine kinase | RRV Pst 1.3 | 98 | AAK54490.1 |
| | | | | | | | TFV 29 | 94 | NP_571995.1 |
| | | | | | | | LCDV 8 | 31 | NP_078770.1 |
| | | | | | | | CIV 179R | 31 | NP_149642.1 |
| 59R | 67,929–68414 | 162 | 18,201 | | | | TFV 30 | 94 | |
| | | | | | | | LCDV 75 | 33 | NP_078685.1 |
| 60R | 68,786–69,202 | 139 | 15,155 | | | | TFV 32 | 97 | |
| 61R | 69,255–71,468 | 738 | 82,214 | | | | TFV 33 | 76 | |
| 61bR[d] | 70,858–71,124 | 89 | 8,955 | | | | | | |
| 62R | 71,555–71,743 | 63 | 6,688 | | | | TFV 34 | 93 | |
| 63L[d] | 72,223–72,576 | 118 | 13,020 | | | | | | |
| 64R | 74,110–74,736 | 209 | 23,196 | Catalytic domain of ctd-like phosphatases | smart00577 | | TFV 40 | 96 | NP_078678.1 |
| | | | | | | | LCDV 64 | 36 | NP_612227.1 |
| | | | | NIF, NLI interacting factor | pfam03031 | | ISKNV 5L | 30 | NP_149818.1 |
| | | | | Haloacid dehalogenase-like hydrolase | pfam00702 | | CIV 355R | 29 | |
| 65R | 74,878–76,572 | 565 | 62,080 | Ribonucleotide reductase, barrel domain | pfam02867 | Ribonucleoside-diphosphate reductase | TFV 41 | 96 | NP_571996.1 |
| | | | | | | | LCDV 12 | 56 | NP_078756.1 |
| | | | | | | | CIV 85L | 45 | NP_149548.1 |
| | | | | | | | | 47 | |
| 66R | 76,682–76,945 | 88 | 9,579 | | | | TFV 42 | 90 | |
| 67R | 77,048–77,668 | 207 | 22,733 | | | | TFV 43 | 75 | |
| 68R | 77,899–78,036 | 46 | 5,706 | | | | TFV 44 | 92 | |
| 69R | 78,111–81,605 | 1,165 | 129,215 | Polycoat protein domain | | | RRV Pst 4.9 | 99 | AAK37740.1 |
| | | | | | | | TFV 45 | 96 | |
| | | | | | | | LCDV 2 | 28 | NP_078748.1 |
| | | | | | | | ISKNV 76L | 26 | NP_612298.1 |
| | | | | | | | CIV 295L | 25 | NP_149758.1 |
| 70L | 82,155–82,913 | 253 | 25,493 | | | | TFV 46L | 59 | |
| 71L | 83,043–83,450 | 136 | 15,582 | | | | RRV Pst 3.0 | 97 | AAK54494.1 |
| | | | | | | | TFV 47L | 96 | NP_078640.1 |
| | | | | | | | LCDV 88 | 37 | |
| 72L | 83,507–84,331 | 275 | 30,831 | Neurofilament triplet H1-like | | | RRV Pst 3.0 | 78 | AAK54495.1 |
| | | | | | | | TFV 48L | 56 | |
| 72bR[d] | 83,517–84,341 | 275 | 30,760 | | | | | | |
| 73L | 84,461–84,874 | 253 | 25,493 | | | | TFV 49L | 93 | |
| 74L | 84,804–85,187 | 126 | 14,321 | | | | TFV 50L | 90 | |
| 75L | 85,238–86,776 | 513 | 56,983 | SAP-DNA binding domain | pfam02037 | | TFV 51L | 94 | |
| | | | | | smart00513 | | RRV Pst 3.0 | 90 | AAK54496.1 |
| | | | | | | | LCDV 58 | 38 | NP_078703.1 |
| 76R | 86,858–88,540 | 561 | 61,672 | | | | TFV 53 | 96 | |
| | | | | | | | LCDV 9 | 22 | NP_078649.1 |
| 77L | 89,329–92,253 | 975 | 108,871 | | | D5 family NTPase | TFV 22 | 97 | |
| | | | | | | | LCDV 6 | 36 | NP_078717.1 |
| | | | | | | | ISKNV 109L | 35 | NP_612331.1 |
| | | | | | | | CIV 184R | 35 | NP_149647.1 |
| | | | | | | | “ | 23 | |
| 78R | 92,383–93,039 | 219 | 25,345 | | | | TFV 21 | 97 | |
| | | | | | | | LCDV 70 | 44 | NP_078618.1 |
| | | | | | | | ISKNV 56L | 41 | NP_612278.1 |
| | | | | | | | CIV 67R | 25 | NP_149530.1 |
| 79L | 93,592–94,038 | 149 | 16,220 | | | | TFV 20 | 92 | |
| | | | | | | | LCDV 84 | 44 | NP_078769.1 |
| | | | | | | | CIV 117L | 43 | NP_149580.1 |
| 80L | 94,089–96,083 | 665 | 72,595 | | | | TFV 19 | 69 | |
| | | | | | | | LCDV 14 | 28 | NP_078619.1 |
| | | | | | | | LCDV 16 | 25 | NP_078744.1 |
| | | | | | | | LCDV 13 | 29 | NP_078677.1 |
| | | | | | | | CIV 380R | 24 | NP_149843.1 |
| | | | | | | | LCDV 25 | 25 | NP_078687.1 |
| | | | | | | | LCDV 15 | 26 | NP_078684.1 |
| | | | | | | | CIV 232R | 29 | NP_149695.1 |
| | | | | | | | “ | 34 | |
| | | | | | | | “ | 29 | |
| | | | | | | | ISKNV 55L | 25 | NP_612277.1 |
| | | | | | | | CIV 378R | 45 (×3) | NP_198481.1 |
| 81R | 96,410–97,915 | 502 | 53,426 | | | | TFV 18L | 97 | |
| 82R[d] | 97,955–98,899 | 315 | 34,685 | | | | | | |
| 83L | 99,477–100,400 | 308 | 34,797 | ATPase | smart00382 | ATPase | TFV 16 | 95 | NP_571992.1 |
| | | | | | pfam00004 | | FV3 | 84 | S27907.1 |
| | | | | ABC transporter | pfam00005 | | LCDV 46 | 53 | NP_078656.1 |
| | | | | Viral (superfamily 1) RNA helicase | pfam01443 | | ISKNV 122R | 51 | NP_612345.1 |
| | | | | | | | CIV 75L | 37 | NP_149538.1 |
| 84L | 100,485–100,856 | 124 | 14,025 | | | | TFV 15 | 94 | |
| | | | | | | | LCDV 92 | 28 | NP_078646.1 |
| 85R | 100,964–101,203 | 80 | 9,425 | | | | TFV 14L | 85 | |
| 86R | 101,169–101,360 | 64 | 7,564 | | | | TFV 13L | 84 | |
| 87R | 101,753–102,643 | 297 | 32,680 | | | | TFV 12L | 96 | |
| | | | | | | | LCDV 39 | 31 | NP_078701.1 |

*(continued on next page)*

Table 2 *(continued)*

| ORF | Nucleotide position | Number of amino acids | MW | Conserved region or signature | Accession No. of conserved regions[a] | Predicted function | Corresponding iridovirus ORFs[b] | Percentage of identity to iridovirus ORFs[c] | Accession No. of iridovirus ORFs[a] |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | CIV 287R | 23 | NP_149750.1 |
| | | | | | | | ISKNV 96L | 26 | NP_612318.1 |
| 88L | 102,715–102,924 | 70 | 7,856 | | | | TFV 11 | 94 | |
| 89R | 103,279–103,962 | 228 | 24,784 | | | | TFV 103 | 92 | |
| | | | | | | | LCDV 48 | 24 | NP_078768.1 |
| 90R | 104,031–104,441 | 137 | 15,264 | | | | TFV 104 | 85 | |
| 91R | 104,836–105,603 | 256 | 29,795 | Putative replication factor and/or DNA binding/packing protein | | | TFV 105 | 95 | |
| | | | | | | | LCDV 43 | 41 | NP_078747.1 |
| | | | | | | | CIV 282R | 34 | NP_149745.1 |
| | | | | | | | ISKNV 61L | 24 | NP_612283.1 |

[a] Accession numbers were derived from the NCBI database.

[b] Isolate abbreviations: RRV, Regina ranavirus; TFV, tiger frog virus; FV3, frog virus 3; LCDV, lymphocystis disease virus;, CIV, *Chilo* iridovirus; ISKNV, infectious spleen kidney necrosis virus.

[c] The percentage of amino acid identity is based on BlastP scores.

[d] These ATV ORFs are not found in the genome of TFV, the other *Ranavirus* genome that has been sequenced.

point mutation occurred in TFV that shifted the RAD2 start site upstream from the start sites present in all other iridoviruses. Second, ATV ORF 19L has sequence homology to a deoxynucleotide kinase or thymidylate kinase that was not described in the TFV genome. The homologous ORF was identified in the opposite strand of TFV ORF 91L and in all of the other iridoviral genomes sequenced (Table 2). This gene's function appears to be required for virus replication as it is conserved throughout the *Iridoviridae*. Finally, ATV appears to lack a functional integrase gene. Sequence analysis shows that ATV has a region homologous at the nucleotide level to the FV3 integrase gene. However, there are two stop codons within the ATV sequence that prevent synthesis of the full-size protein. While this finding suggests that the ATV integrase is not functional, we favor another interpretation. Identification of the homologous region in FV3 as an integrase was not based on functional studies, but on the presence of an integrase-like motif within the viral sequence (Rohoziniski and Goorha, 1992). Our recent work calls this interpretation into question as analysis of the putative FV3 integrase gene and other homologous genes show no homology to authentic integrase genes (data not shown). Whatever the function of the misidentified integrase gene, it appears as if its ATV counterpart is nonfunctional.

Three ATV ORFs are potential immune evasion or pathogenesis-related proteins. ATV ORF 40L encodes a protein with homology to human CARD-like caspases. This ORF was not described in TFV, but can be found in the TFV genome in the opposite strand of TFV ORF 70L. The function of this protein is unknown, but it may regulate apoptosis in viral-infected cells (Bouchier-Hayes and Martin, 2002). TFV ORF 54L encodes a 355-amino-acid protein with homology to human and mouse 3$\beta$-hydroxy-$\Delta$5-C27 steroid oxidoreductase. Comparing this ORF with the homologous ATV gene (ORF 52R) reveals that the predicted size of the ATV gene product is truncated to 53 amino acids. This protein has been suggested to be involved with the biosynthesis of hormonal steroids (He et al., 2002) and, by analogy to a similar gene in poxviruses, may be involved with host immune evasion (Chinchar, 2002). It is not clear if this protein is required for ATV replication or if the truncated version found in ATV is sufficient for in vivo protein function. Open reading frame 57R encodes a homolog of the alpha subunit of eukaryotic translation initiation factor 2 (eIF-2$\alpha$) (Table 2). Phosphorylation of eIF-2$\alpha$ leads to a subsequent block in translation initiation and is an effective host antiviral mechanism. The ATV 57R gene is conserved throughout the genus *Ranavirus* (Yu et al., 1999; Essbauer et al., 2001), and a similar gene has been identified among poxviruses, the K3L gene (Beatie et al., 1991). The poxvirus K3L gene is an interferon resistance gene thought to function as a competitive inhibitor for eIF-2$\alpha$ phosphor-

Table 3
Vertebrate iridovirus predominant codon usage

| Amino acid | | LCDV | ISKNV | TFV | ATV |
|---|---|---|---|---|---|
| A | Ala | GCA | | GCC | GCC (50.1) |
| C | Cys | | | | TGC (81.5) |
| D | Asp | GAT | GAC | GAC | GAC (86.2) |
| E | Glu | GAA | GAG | GAG | GAG (79.4) |
| F | Phe | | | | TTT (60.5) |
| G | Gly | GGT | GGC | | GGC (31.6) |
| H | His | CAT | CAC | CAC (86.7) | |
| I | Ile | | | | ATC (48.0) |
| K | Lys | AAA | AAG | AAG | AAG (77.2) |
| L | Leu | **TTA** | CTG | CTG | CTG (42.1) |
| N | Asn | AAT | AAC | AAC | AAC (83.9) |
| P | Pro | | | | CCC (47.9) |
| Q | Glu | CAA | CAG | CAG | CAG (80.4) |
| R | Arg | AGA | **CGC** | | AGG (65.0) |
| S | Ser | TCT | **AGC** | | TCC (29.5) |
| T | Thr | | | | ACC (44.2) |
| V | Val | GTT | GTG | GTG | GTG (43.1) |
| Y | Tyr | TAT | TAC | TAC | TAC (88.6) |

*Note.* The predominant codon used for each amino acid is as presented in references for LCDV (Tidona and Darai, 1997), ISKNV (He et al., 2001), and TFV (He et al., 2002). For ATV the predominant codons used are presented with the percentage of use of the predominant codon given in parentheses. Codons in boldface type differ from those in ATV at nucleotides other than in the third (wobble) position.

Table 4
TV ORF characteristics

| Category | ATV ORF |
|---|---|
| Genes with homology to putative viral/cellular replicative proteins | 1L (myristylated membrane protein), 6R (RNA pol II $\alpha$ subunit), 7L (helicase/NTPase), 10L (RAD2), 14L (MCP), 16L (thiol oxidoreductase), 19L (deoxynucleoside kinase), 20L (P CNA), 21L (methyltransferase), 22L (thymidylate synthase), 24L (transcription elongation factor-SII), 25R (RNase III), 31R (NTPase/helicase), 38R (ribonucleotide reductase, small subunit), 42L (dUTPase), 43R (RNA pol II $\beta$ subunit), 44L (DNA pol ), 47L (phosphotransferase), 50R (helicase), 51L (myristylated membrane protein/glycosyl hydrolase), 58R (tyrosine kinase), 65R (ribonucleotide diphosphate reductase), 77L (D5 family NTPase), 83L (ATPase) |
| Genes potentially involved with immune modulation/ pathogenesis | 40L (CARD-like caspase), 52R (steroid oxidoreductase), 57R (eif-2 homolog) |
| Genes with homology to other iridovirus ORFs, but of unknown function | 2L, 3R, 4R, 5L, 6bL, 8R, 11R, 12R, 13L, 15L, 17R, 18R, 23L, 26L, 27R, 28L, 29R, 30R, 32R, 34L, 35L, 36R, 37bL, 41R, 43bL, 45R, 46L, 48L, 49R, 53R, 54R, 55R, 56R, 59R, 60R, 61R, 62R, 64R, 66R, 67R, 68R, 69R, 70L, 71L, 72L, 73L, 74L, 75L, 76R, 78R, 79L, 80L, 81R, 84L, 85R, 86R, 87R, 88L, 89R, 90R, 91R |
| Genes of unknown function, and no homology | 9L, 33L, 37R, 39L, 61bR, 63L, 72bR, 82R |

ylation, thereby inhibiting this host antiviral response (Carroll et al., 1993).

Sixty ATV ORFs have homology to other iridovirus ORFs, but of unknown function. Fifty-seven (95%) have homology with TFV ORFs, 21 (35%) have homology to LCDV ORFs, 13 (21.7%) have homology with CIV ORFs, and 9 (15%) have homology with ISKNV ORFs (Table 2). Although little is known about many of these proteins, several of them encode genes (e.g., the 18-K early protein, the 31-kDa protein, and p40) sequenced in FV3 and other iridoviruses. Even though their functions are not known, homologies among iridoviral genes with unknown function suggest that they may encode proteins whose action is required for iridovirus replication. Further analysis will determine the role of these conserved viral genes in iridovirus replication.

In addition to the proteins of known function and/or homology, there are eight predicted ATV ORFs that encode proteins of unknown function that lack homology to iridovirus or other viral or cellular sequences. Their predicted gene products vary in size from 49 to 315 amino acids in length (Table 2). One ORF, 9L, has homology with the DNA sequence of TFV ORF 102; however, there is no homology between the translation product of TFV ORF 102 and ATV ORF 9L. The eight ORFs in ATV whose predicted gene products have no known function nor homology to other viral sequences may be genes required for pathogenesis in salamanders. Clearly, further analysis of these genes is required.

### *Dot plot analysis*

The genomic sequence of ATV was compared to itself, to the cloned RRV *Pst*I fragments, and to other completely sequenced iridoviruses (i.e., TFV, ISKNV, LCDV, and CIV). The ATV genome compared to itself revealed the expected −45° diagonal line (Fig. 2A), plus horizontal and vertical "dots" (thin arrows). These "dots" indicate repeat sequences throughout the genome (Figs. 2A and B). Sequence analysis of these horizontal and vertical "dots" revealed that these repeats are found within intergenic regions, are A + T rich, and contain, for the most part, a 14-bp palindromic sequence (Table 5). The 14-bp palindrome was observed 76 times within the ATV genome. A similar repeat sequence was observed 52 times within the genome of TFV but only 3 times in the LCDV genome, once in the CIV genome, and was not present within the ISKNV genome. Although the palindromic sequence was generally observed between ATV genes, the sequence was also observed within two putative TFV ORFs of unknown function (60L and 98R; He et al., 2002). The sequence of the palindrome is similar to the upstream sequence repeats observed in the FV3 p40 protein reported by Munnes et al., (1995). Goohra and Granoff (1974) reported that FV3 genes tend to end in regions of dyad symmetry, suggesting that this may serve as a transcriptional termination signal or, if transcribed into RNA, may protect the viral messages, which lack poly(A) tails, from degredation. We also examined the genome for a consensus promoter region associated with the palindromic sequence; however, a conserved sequence was not observed. The absence of a consensus promoter may reflect the fact that viral genes within different temporal classes (i.e., immediate early, delayed early, and late) are regulated differentially and may possess unique transcription factors and promoter regions. Promoter regions for putative immediately early genes in FV3 were used to search for a consensus promoter in the ATV genome, but a common promoter element was not observed. Further analysis may revel different temporal classes of promoter regions.

In addition to the repeat sequences in noncoding regions, there are "dots" along the −45° diagonal that represent

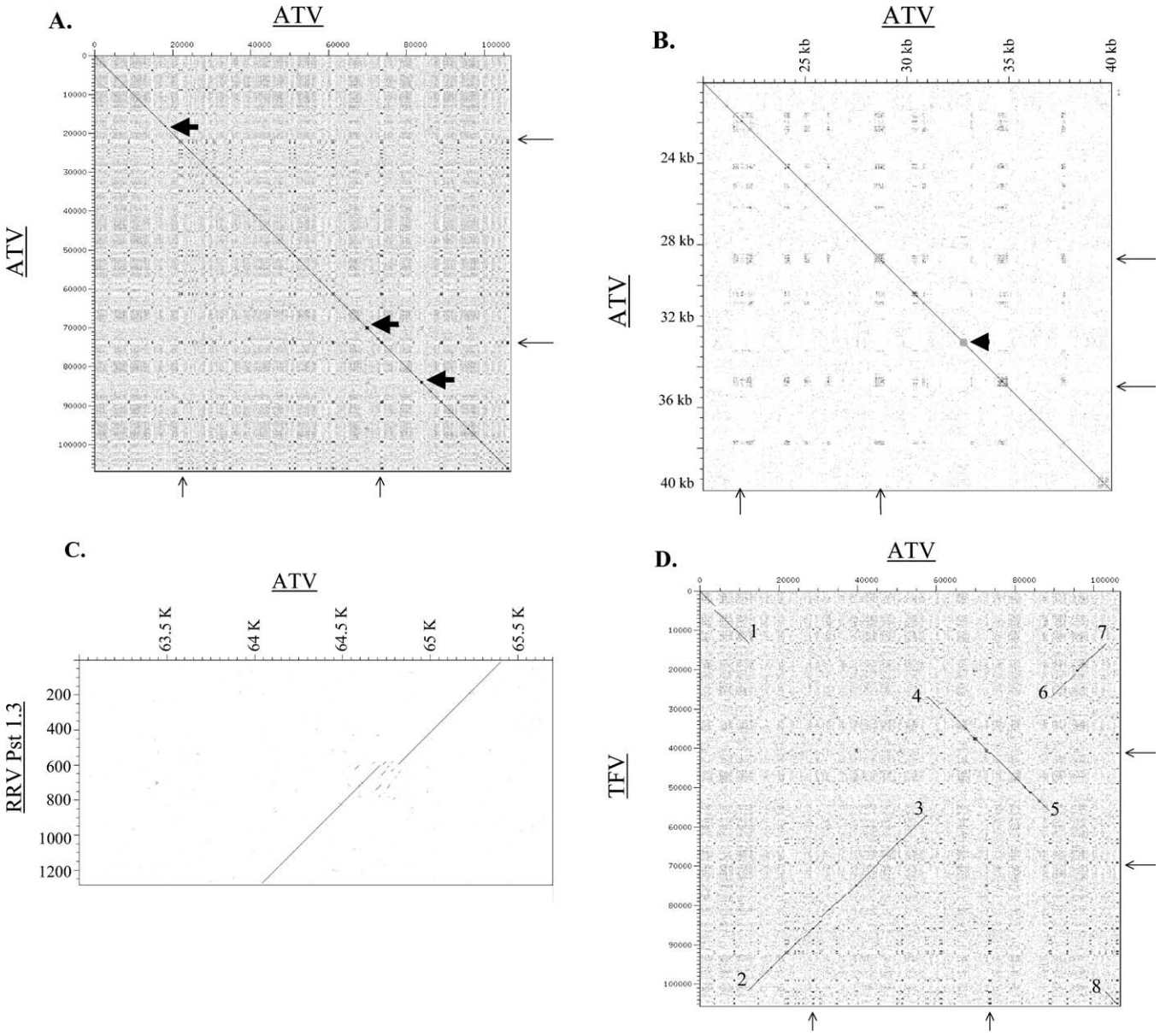


Fig. 2. Dot plot analysis of the ATV genome (horizontal axis) versus (vertical axis) (A) the ATV genome; (B) a 20-kbp region of the ATV genome; (C) the RRV *Pst*I 1.3-kbp fragment; (D) the TFV genome. Thin arrows indicate intergenic repeats, while thick arrows point to intragenic repeats. Breaks in the diagonal line indicate an insertion or deletion, while small lines offset from the diagonal represent local repeat sequences.

localized repeats within genes (Fig. 2, thick arrowheads). Fig. 2B is an enlargement of a repeat within ATV ORF 30R (Fig. 2B, thick arrowhead), a gene without predicted conserved domains or function, but with homology to TFV ORF 79L (Table 2). Analysis of this region revealed a 127-bp sequence repeated three times (Fig. 2B). The function of this and the other observed intragenic repeats is unknown. Repeat sequences were also found in the ISKNV genome (He et al., 2001), as well as in the genomes of FV3, LCDV, CIV, and RSIV (reviewed in Williams, 1996). Poxviruses (Wittek and Moss, 1980), herpesviruses (Wadsworth et al., 1975), baculoviruses (Hayakawa et al., 1999), adenoviruses (Arrand and Roberts, 1979), and retroviruses (Shoemaker et al., 1980) also have repeated elements. The repetitive sequences in these viruses function as replication regulatory sequences. We did not observe any homology

Table 5
Palindromic sequence consensus base frequency

| | Consensus sequence | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | T | A | T | C | T | T | A | A | G | A | T | A | T |
| G | 7 | 0 | 4 | 0 | 0 | 5 | 0 | 0 | 7 | **70** | 0 | 0 | 6 | 9 |
| C | 15 | 11 | 0 | 0 | **68** | 1 | 0 | 0 | 5 | 0 | 0 | 2 | 1 | 7 |
| A | **46** | 7 | **72** | 0 | 4 | 1 | 0 | **75** | **64** | 6 | **76** | 0 | **63** | 10 |
| T | 8 | **58** | 0 | **76** | 4 | **69** | **76** | 1 | 0 | 0 | 0 | **74** | 6 | **50** |

*Note*. Bold numbers indicate the consensus base.

between the repeat elements from other viruses and those from ATV; however, the location of the repeat elements in ATV suggests a role in replication regulation.

To exam the colinearity of ATV and RRV, the genome of ATV was plotted against the currently sequenced RRV *Pst*I fragments. RRV *Pst* fragments of 1.6, 3.8, and 4.9 kbp revealed colinearity with ATV, while the remaining RRV *Pst* fragments of 1.3, 1.9, 3.0 6.7, and 8.1 kbp showed colinear dot plots with polymorphisms located between ATV ORFs. Fig. 2C shows one example of the polymorphisms observed between ATV and RRV. The dot plot shows a repeat sequence, indicated by the short parallel lines offset from the consensus diagonal, as well as an insertion in ATV, indicated by a gap in the consensus line (Fig. 2C). Interestingly, all of the polymorphisms observed between ATV and RRV occurred in noncoding regions, indicating that potential ORFs were not disturbed. In addition, these polymorphic regions may help in differentiating ranavirus strains and may be useful markers in tracing individual viral isolates in the field. For example, we have already used one of these polymorphisms, a 16-bp repeat, to demonstrate that ATV isolates from AZ, UT, CO, and Canada can be readily distinguished by the presence or the absence of this 16-bp repeat (Jancovich et al., in preparation).

Comparing ATV to the LCDV, CIV, or ISKNV genomes did not reveal sequence colinearity (data not shown). In contrast, comparing ATV to TFV identified not only colinear regions, but also inversions and polymorphisms (Fig. 2D). Inversions occurred at basepair positions 12,222, 57,522, 89,316, and 103,128 in the ATV genome corresponding to basepair positions 101,456, 56,854, 26,726, and 13,191 in the TFV genomic sequence (Fig. 2D, postions 7–8, 5–6, 3–4, and 1–2, respectively). The inverted regions and polymorphisms between ATV and TFV reflect the deletions, insertions, or truncations of different ORFs between the two ranaviruses. For example, one ATV ORF (52R) that is homologous to the TFV ORF 54L encoding the putative 3$\beta$-hydroxy-$\Delta$5-C27 steroid oxidoreductase was truncated by an inversion occurring between positions 3 and 4 (Fig. 2D). The truncation of ATV ORF 52R suggests that the inversion event occurred in ATV as this virus diverged from a common ancestor. Perhaps the inversions detected between ATV and TFV reflect the high recombination rate of ranaviruses. In FV3 very high recombination rates have been detected (Granoff and Chinchar, 1986), and this may contribute to genomic rearrangements. Moreover, if the rearrangement occurs in a nonessential region, it may be propagated throughout the population. In addition, inversions may create more, or fewer, pathogenic viral strains as one or more genes are disrupted. For example, the limited host range of ATV (Jancovich et al., 2001) may reflect the observed genomic rearrangement between ATV and TFV.

### *Phylogenetic analysis*

The sequence of the ATV viral DNA polymerase (DNA pol) gene, viral RNA polymerase II $\beta$ subunit, and the MCP were used to determine the relationship of ATV with other iridoviruses as well as with non-iridoviruses and eukaryotic organisms. Phylogenic trees for all three genes support the view that ATV is a member of the genus *Ranavirus* within the family *Iridoviridae* (Figs. 3A–C). This result is supported with high bootstrap support. Analysis of the MCP gene (Fig. 3A) clearly shows that ATV is more closely related to FV3 and TFV (genus *Ranavirus*) than it is to LCDV, CIV, and ISKNV (members, respectively, of *Lymphocystivirus*, *Iridovirus*, and an unassigned genus) (bootstrap values $>$95%). Interestingly, analysis of the MCP gene suggests that a virus recently isolated from groupers may also be a new member of the genus *Ranavirus*. Our analysis of the viral DNA polymerase supports the conclusions made by Stasiak et al. (2000). The *Ascoviridae* and *Iridoviridae* families are clustered and are more distantly related to the eukaryotic polymerase sequences, which also cluster together with high statistical support (Fig. 3C). In addition, the bootstrap values of 100% that separate ATV from TFV in all three genes suggest that ATV has diverged from the FV3 and TFV-like viruses (Figs. 3A–C). Therefore, these three genes show congruent patterns and establish that ATV is a member of the genus *Ranavirus* within the *Iridoviridae*.

In the area of molecular biology, this sequence analysis provides the groundwork for identifying the function of key viral genes, e.g., the subunits of the DNA-dependent RNA polymerase, the 18-kDa early protein, the eIF-2 $\alpha$ homolog, etc., for ranaviruses associated with salamanders in North America. In addition, the number of polymorphisms observed between two closely related ranavirus isolates, ATV and RRV, may provide a way to track the progress and origin of ATV outbreaks. In this way we should be able to elucidate the molecular events underlying ATV replication, host restriction, viral pathogenesis, and immune evasion.

## Materials and methods

### *Construction of a genomic library and genomic sequencing: ATV*

ATV was isolated in 1996 from diseased tiger salamanders (*A. tigrinum stebbinsi*) found in the San Raphael Valley (Arizona) (Jancovich et al., 1997). ATV was grown in EPC cells (*epitheloma papilloma ciprini;* Fijan et al., 1983) in Eagle's minimal essential medium (Cellgro, USA) supplemented with 10% FBS (Hyclone, USA). Viral DNA was isolated by a modified Hirt extraction (Hirt, 1967). Twenty-five 75-cm$^2$ flasks (Corning, USA) of EPC cells were infected with ATV at a m.o.i. of about 0.1 PFU and the infection was allowed to proceed until cytopathic

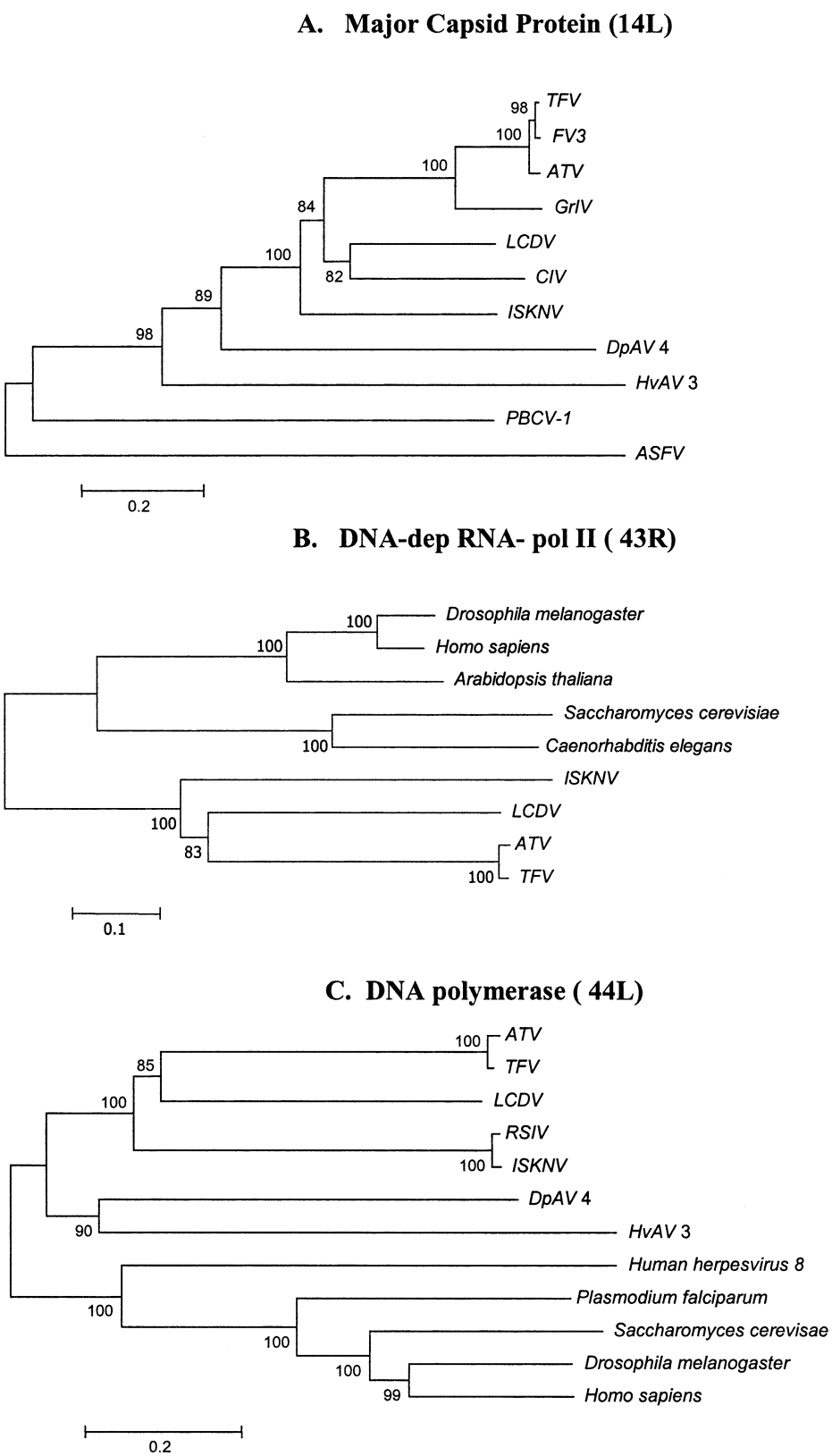


Fig. 3. Phylogenetic relationships of iridoviruses obtained using three protein sequence alignments: (A) major capsid protein, (B) RNA pol II $\beta$ subunit protein, (C) DNA polymerase protein. The neighbor-joining trees obtained using MEGA2 are shown with the statistical support indicating the robustness of the inferred branching pattern as assessed using the bootstrap test. Names, abbreviations, and amino acid GenBank Accession numbers for the sequences used for phyologenetic analysis: tiger frog virus (TFV) 18767719, 18767710, 18767709; frog virus 3 (FV3) 20137785; grouper iridovirus (GrIV) 19880315; lymphocystis disease virus (LCDV) 9695414, 13358406, 13358408; insect iridovirus IIV [*Chilo* iridescent virus (CIV)]15078986; infectious speen kidney necrosis virus (ISKNV) 19881411, 19881439, 19881424; *Diadromus pulchellus* ascovirus 4a (DpAV) 21668339, 11931713; *Heliothis virescens* ascovirus 3 (HvAV) 16265802, 21668320; *Paramecium bursaria* chlorella virus 1 (PBCV-1) 1620102; African swine fever virus (ASFV) 16905424; red sea bream iridovirus (RSIV) 6015024; human herpesvirus 8 2246468; *Arabidopsis thaliana* 15234572; *Plasmodium falaciparum* 23507969; *Saccharomyces cerevisiae* 6324781, 6320101; *Ceanorhabditis elegans* 20451233; *Drosophila melanogastar* 84936, 17647353; *Homo sapiens* (human) 4505941, 14250670.

effects were observed in about 80% of the cells. Cells and virus were removed from the flask and centrifuged at 60,000 × *g* for 60 min. The pellet was resuspended in 2.5 ml of a hypotonic lysis buffer (10 mM Tris–HCl, pH 8, 10 mM EDTA, 0.5% NP-40, 0.25% sodium deoxycholate) and incubated on ice for 5 min before centrifugation at 3000 × *g* at 4°C for 10 min to remove nuclei and cellular contamination. The virus-containing supernatant was removed, treated with a final concentration of 1% SDS and 1 mg/ml of proteinase K, and incubated at 50°C for 60 min. Viral DNA was extracted with phenol:chlorophorm:isoamyl alcohol and precipitated in 70% ethanol with 0.3 M sodium acetate at −20°C. Viral DNA was removed by spooling, washed twice in ice-cold 70% ethanol, and resuspended in sterile water.

Purified ATV genomic DNA was sheared using a probe sonicator (Branson SONICATOR Cell Disruptor 200, Branson Sonic Power Co., USA) at the highest microtip setting and continuous pulse. At 1-s intervals, 100-μl aliquots of sonicated DNA were removed and analyzed by gel electrophoresis. DNA between 500 and 2500 bp in length was isolated using DEAE–cellulose membrane extraction as described by the manufacturer (Schleicher & Schuell, USA). Sheared ATV genomic DNA was polished and ligated into the pSTBlue-1 vector using the Perfectly Blunt Cloning kit as described by the manufacturer (Novagen, USA), and cloned DNA was transformed into TOP10 chemically competent *Escherichia coli* cells (Invitrogen, USA). Transformed cells were plated onto LB agar plates containing 0.1 mg/ml of ampicillin and 50 μg/ml of X-gal and incubated at 37°C overnight. About 1500 white colonies were picked and plasmids containing ATV DNA were isolated using the Eppendorf 96-Spin Plasmid Isolation kit (Eppendorf, Germany). The genomic library was sequenced with primers complementary to the SP6 and T7 regions of the vector and an ABI System 377 automated DNA sequencer (Perkin–Elmer, USA) using dye-labeled dideoxynucleotide triphosphates (Sanger et al., 1977). The genome was assembled using Sequencher (Gene Codes, USA). Gaps in the genome were filled using primer walking techniques.

### *Construction of a genomic library and genomic sequencing: RRV*

RRV was initially isolated in Saskatchewan, Canada, from infected *A. tigrinum diaboli* (Bollinger et al., 1999). Virus was prepared from cultures of infected fathead minnow (FHM) cells and purified as follows. Twelve confluent 150-$cm^2$ flasks of FHM cells were infected with RRV at a low m.o.i. When CPE was marked, about 10 days postinfection, the cells were removed from the flasks by swirling and virions pelleted from clarified culture medium by high-speed centrifugation. The virus pellet was resuspended in RSB (10 mM Tris–HCl, pH 7.5, 10 mM NaCl, 1.5 mM $MgCl_2$) and frozen at −80°C. A 3-ml aliquot of concentrated virus was thawed and, after increasing the $MgCl_2$ concentration to 10 mM, DNase was added to a final concentration of 200 μg/ml, and the sample digested for 1 h at 37°C. At that time, EDTA was added to a final concentration of 50 mM, and the treated viral suspension was layered atop an 8-ml cushion of 20% (w/w) sucrose/RSB. The sample was centrifuged for 1.5 h in a SW41 rotor (Beckman, USA) at 30,000 rpm at 4°C. Following centrifugation, the overlay along with the upper portion of the sucrose cushion was removed by aspiration, and the remaining liquid poured off. The resulting pellet was resuspended in 1 ml of TE containing 1% SDS, 200 μg/ml of proteinase K, and 40 μg/ml of DNase-free RNase and digested overnight at 37°C. The following day viral DNA was isolated by phenol–chloroform extraction, and viral DNA precipitated with 2 vol of 95% ethanol in the presence of 0.3 M sodium acetate. Viral DNA was digested with *Pst*I and the resulting fragments were cloned into pGEM-7Z (Promega, USA).

RRV *Pst*I fragments were sequenced using the transposon-facilitated method of Strathmann et al. (1991). Prior to transposon insertion, RRV *Pst*I restriction fragments were subcloned into pCSOS-72 (GenBank Accession No. AF061788), and the Tn1000 transposon was inserted into pCSOS-72::RRV *Pst*I by bacterial conjugation. The location of the transposon was determined by long-range PCR, and a series of clones, containing transposons spaced at ~800-bp intervals along the inserted fragment, were identified and sequenced using a LiCor 4000L automated DNA sequencer and primers complementary to the flanking vector and transposon sequences. Altogether nine *Pst*I fragments encompassing 42 kbp of RRV genomic DNA were sequenced.

### *Genome analysis*

Putative ORFs were predicted using four different methods. First, the ATV genome was divided into 5-kbp fragments and each fragment was examined by BLASTX analysis (Gish and States, 1993). The genome was then examined using GeneMarkS (Besemer et al., 2001), Glimmer (Delcher et al., 1999), and Vector Nti software (InforMax Inc., USA); in all cases an ORF had to start with an AUG codon and be a minimum of 40 codons long to be considered a putative gene. Open reading frames predicted by all four programs, or with homology to any previously identified iridovirus ORFs, were considered in our analysis. Dot plot analysis of the ATV viral genome was performed using the Dotter program (Sonnhammer and Durbin, 1995) available at the Poxvirus Bioinformatics Resource Center (http://www.poxvirus.org). The genome of ATV was compared to itself, RRV, and other iridovirus genomic sequences (TFV, ISKNV, LCDV, and CIV).

### *Multiple sequence alignment and phylogenetics*

The viral DNA polmerase (DNA pol), major capsid protein (MCP), and β subunit of the virus-encoded DNA-

dependent RNA polymerase (DNA dep RNA pol) sequences of ATV and other iridoviruses were used to determine the taxonomic position of ATV. Amino acid sequence alignments were constructed using the default options in ClustalW (Thompson et al., 1994) and the phylogenetic relationships among species were determined using the neighbor-joining (NJ; Saitou and Nei, 1987) as well as the maximum parsimony methods as implemented in MEGA2 (Kumar et al., 2001). All sites containing alignment gaps and missing data were excluded from phylogenetic analyses. For the NJ method, pairwise protein sequence divergence was corrected for multiple hits by using the Poisson Correction (Nei and Kumar, 2000). The reliability of the NJ tree was inferred using the Felsenstein (1985) bootstrap method with 1000 replicates in MEGA2. Outcomes of the bootstrap test are occurrence frequencies of each interior branch in the original tree, which provide a measure of the robustness of the respective branching patterns. A bootstrap value $>95\%$ is considered statistically significant, but branches with bootstrap frequency $>80\%$ are also often considered well resolved. For the maximum parsimony (MP) analysis, equal weighting was used for protein sequence changes in MEGA2. All MP trees were identical in topology to the NJ trees reported here.

## Acknowledgments

We thank Dr. Scott Bingham at the Arizona State University DNA Laboratory and Dr. Charles Woodley at the University of Mississippi Medical Center DNA Sequencing Facility for their help with the sequencing of ATV and RRV, respectively, and Dr. Chris Upton for his help with the dot plot analyses. This research was funded by Grant IBN9977063 from the National Science Foundation Integrated Reseach Challanges in Environmental Biology.

## References

Arrand, J., Roberts, R.J., 1979. The nucleotide sequences at the termini of adenovirus 2 DNA. J. Mol. Biol. 128, 577–594.

Beatie, E., Tartaglia, J., Paoletti, E., 1991. Vaccinia virus-encoded eIF-2$\alpha$ homolog abrogates the anti-viral effect of interferon. Virology 183, 419–422.

Besemer, J., Lomsadze, A., Borodovsky, M., 2001. GeneMarkS: a self-training method for the prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29, 2606–2618.

Bollinger, T.K., Mao, J., Schock, D., Brigham, R.M., Chinchar, V.G., 1999. Pathology, isolation and molecular characterization of an iridovirus from tiger salamanders in Saskatchewan. J. Wildlife Dis. 35, 413–429.

Bouchier-Hayes, L., Martin, S.J., 2002. CARD games in apoptosis and immunity. EMBO Rep. 3, 616–621.

Carroll, K., Elroy-Stein, O., Moss, B., Jagus, R., 1993. Recombinant vaccinia virus K3L gene product prevents activation of double-stranded RNA-dependent, initiation factor 2 alpha-specific protein kinase. J. Biol. Chem. 268, 12837–12842.

Chinchar, V.G., Granoff, A., 1986. Temperature-sensitive mutants of frog virus 3: biochemical and genetic characterization. J. Virol. 58, 192–202.

Chinchar, V.G., 2002. Ranaviruses (family *Iridoviridae*): emerging cold-blooded killers. Arch. Virol. 147, 447–470.

Collins, J.P., Storfer, A., 2003. Global amphibian declines: sorting the hypotheses. Diversity and Distributions 9, 89–98.

Daszak, P., Berger, L., Cunningham, A.A., Hyatt, A.D., Green, D.E., Speare, R., 1999. Emerging infectious diseases and amphibian population declines. Emerg. Infect. Dis. 5, 735–748.

Delcher, A.L., Harmon, D., Kasif, S., White, O., Salzberg, S.L., 1999. Improved microbial gene identification with GLIMMER. Nuc. Acid Res. 27, 4636–4641.

Essbauer, S., Bremont, M., Ahne, W., 2001. Comparison of the eIF-2 alpha homologous proteins of seven ranaviruses (*Iridoviridae*). Virus Genes 23, 347–359.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using bootstrap. Evolution 39, 783–791.

Fijan, N., Sulimanovic, D., Bearzotti, M., Muzinic, D., Zwillenberg, L.O., Chilmonczyk, S., Vautherot, J.F., de Kinkelin, P., 1983. Some properties of the *Epithelioma papulosum cyrini* (EPC) cell line from carp *Cyprinus carpio*. Ann. Virol. (Inst Pasteur) 134, 207–220.

Gish, W., States, D.J., 1993. Identification of protein coding regions by database similarity search. Nature Genet. 3, 266–272.

Goorha, R., Granoff, A., 1974. Macromolecular synthesis in cells infected with frog virus 3 II. Evidence for post-transcriptional control of a viral structural protein. Virology 60, 251–259.

Hayakawa, T., Ko, R., Okano, K., Seong, S., Goto, C., Maeda, S., 1999. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. Virology 262, 277–297.

He, J.G., Deng, M., Weng, S.P., Li, Z., Zhou, S.Y., Long, Q.X., Wang, X.Z., Chan, S.M., 2001. Complete genome analysis of the mandarin fish infectious spleen and kidney necrosis iridovirus. Virology 291, 126–139.

He, J.G., L,L., Deng, M., He, H.H., Weng, S.P., Wang, X.H., Zhou, S.Y., Long, Q.X., Wang, X.Z., Chan, S.M., 2002. Sequence analysis of the complete genome of an iridovirus isolated from the tiger frog. Virology 292, 185–197.

Hirt, B., 1967. Sequence extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26, 365–369.

Jakob, N.J., Muller, K., Bahr, U., Darai, G., 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.

Jancovich, J.K., Davidson, E.W., Morado, J.F., Jacobs, B.L., Collins, J.P., 1997. Isolation of a lethal virus from the endangered tiger salamander, *Ambystoma tigrinum stebbinsi* Lowe. Dis. Aquat. Org. 31, 161–167.

Jancovich, J.K., Davidson, E.W., Seiler, A., Jacobs, B.L., Collins, J.P., 2001. Transmission of the *Ambystoma tigrinum* virus to alternative hosts. Dis. Aquat. Org. 46, 159–163.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetic analysis. Bioinform. Appl. Note 17, 1244–1245.

Munnes, M., Schetter, C., Holker, I., Doerfler, W., 1995. A fully 5′-CG-3′ but not a 5′-CCGG-3′ methylated late frog virus 3 promoter retains activity. J. Virol. 69, 2240–224.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford Univ. Press, New York.

Rohozinski, J., Goorha, R., 1992. A frog virus 3 gene codes for a protein containing the motif characteristic of the INT family of integrases. Virology 186, 693–700.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA 74, 5463–5467.

Shoemaker, C., Goff, S., Gilboa, E., Paskind, M., Mitra, S.W., Baltimore, D., 1980. Structure of a cloned circular Moloney murine leukemia virus DNA molecule containing an inverted segment: implications for retroviral integration. Proc. Natl. Acad. Sci. USA 77, 3932–3936.

Sonnhammer, E.L.L, Durbin, R., 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167, GC1–10.

Stasiak, K., Demattei, M.V., Federici, B.A., Bigot, Y., 2000. Phylogenetic position of the *Diadromus pulchellus* ascovirus DNA polymerase among viruses with large double-stranded DNA genomes. J. Gen. Virol. 81, 3059–3072.

Strathmann, M., Hamilton, B.A., Mayeda, C.A., Simon, M.I., Meyerowitz, E.M., Palazzolo, M.J., 1991. Transposon-facilitated DNA sequencing. Proc. Natl. Acad. Sci. USA 88, 1247–1250.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 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.

Tidona, C.A., Darai, G., 1997. The complete DNA sequence of lymphocystis disease virus. Virology 230, 207–216.

Wadsworth, S., Jacob, R.J., Roizman, B., 1975. Anatomy of herpes virus DNA II. Size, composition and arrangement of inverted terminal repetitions. J. Virol. 15, 1487–1497.

Williams, T., 1996. The Iridoviruses. Adv. Virus Res. 46, 345–412.

Williams, T., Chinchar, V.G., Darai, G., Hyatt, A., Kalmakoff, J., Seligy, V., 2000. *Iridoviridae*, in: van Regenmortel et al. (Eds.), Seventh Report of the International Committee on the Taxonomy of Viruses, pp. 167–182.

Wittek, R., Moss, B., 1980. Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. Cell 21, 277–284.

Yu, Y.X., Bearzotti, M., Vende, P., Ahne, W., Bremont, M., 1999. Partial mapping and sequencing of a fish iridovirus genome reveals genes homologous to the frog virus 3 p31, p40 and human eIF2 alpha. Virus Res. 63, 53–63.