JVI Journals.ASM.org

Comparative Genomics of Carp Herpesviruses

Andrew J. Davison,[a] Tomofumi Kurobe,[b] Derek Gatherer,[a] Charles Cunningham,[a] Ian Korf,[c] Hideo Fukuda,[d] Ronald P. Hedrick,[b] Thomas B. Waltzek[b]*

MRC—University of Glasgow Centre for Virus Research, Glasgow, United Kingdom[a]; Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, Davis, California, USA[b]; Genomics Center, University of California, Davis, Davis, California, USA[c]; Laboratory of Genome Science, Tokyo University of Marine Science and Technology, Tokyo, Japan[d]

Three alloherpesviruses are known to cause disease in cyprinid fish: cyprinid herpesviruses 1 and 3 (CyHV1 and CyHV3) in common carp and koi and cyprinid herpesvirus 2 (CyHV2) in goldfish. We have determined the genome sequences of CyHV1 and CyHV2 and compared them with the published CyHV3 sequence. The CyHV1 and CyHV2 genomes are 291,144 and 290,304 bp, respectively, in size, and thus the CyHV3 genome, at 295,146 bp, remains the largest recorded among the herpesviruses. Each of the three genomes consists of a unique region flanked at each terminus by a sizeable direct repeat. The CyHV1, CyHV2, and CyHV3 genomes are predicted to contain 137, 150, and 155 unique, functional protein-coding genes, respectively, of which six, four, and eight, respectively, are duplicated in the terminal repeat. The three viruses share 120 orthologous genes in a largely colinear arrangement, of which up to 55 are also conserved in the other member of the genus *Cyprinivirus*, anguillid herpesvirus 1. Twelve genes are conserved convincingly in all sequenced alloherpesviruses, and two others are conserved marginally. The reference CyHV3 strain has been reported to contain five fragmented genes that are presumably nonfunctional. The CyHV2 strain has two fragmented genes, and the CyHV1 strain has none. CyHV1, CyHV2, and CyHV3 have five, six, and five families of paralogous genes, respectively. One family unique to CyHV1 is related to cellular JUNB, which encodes a transcription factor involved in oncogenesis. To our knowledge, this is the first time that JUNB-related sequences have been reported in a herpesvirus.

The order *Herpesvirales* consists of the families *Herpesviridae*, *Alloherpesviridae*, and *Malacoherpesviridae* (1). Each family contains viruses that are associated with distinct host groups. Thus, members of the family *Herpesviridae* infect mammals, birds, or reptiles, members of the family *Alloherpesviridae* (referred to as alloherpesviruses) infect fish or frogs, and members of the family *Malacoherpesviridae* infect mollusks. The family *Alloherpesviridae* is divided into the four genera *Batrachovirus*, *Cyprinivirus*, *Ictalurivirus*, and *Salmonivirus*. The genus *Cyprinivirus* contains four species, three of which (*Cyprinid herpesvirus 1*, *Cyprinid herpesvirus 2*, and *Cyprinid herpesvirus 3*) are associated with common carp or goldfish (family Cyprinidae in the order Cypriniformes and one of which (*Anguillid herpesvirus 1*) is associated with freshwater eels (family Anguillidae in the order Anguilliformes).

Cyprinid herpesvirus 1 (CyHV1) is also known as carp herpesvirus, carp pox virus, and *Herpesvirus cyprini*. CyHV1 disease is most frequently characterized by mucoid to waxy epidermal growths on the skin of common carp (*Cyprinus carpio*) and koi carp (a variety of *Cyprinus carpio*). This disease is perhaps the oldest known in fish, having been recorded in the Middle Ages (2), and has been described by various names, the most prominent being carp pox (3). It has a broad geographical distribution (4). A viral etiology was first suggested in 1907 (5) and was supported over 5 decades later by the detection of particles in the growths that have the distinctive morphology of a herpesvirus (6). The virus was first isolated by Sano and colleagues, who also demonstrated the etiology of the disease (7), and it has since been isolated by others (8). Two phases of infection have been described, namely, an acute, generally lethal, systemic disease in young carp and a recurring, generally nonlethal, proliferative skin disease linked to periods of lower water temperature among survivors or fish exposed at older ages (7, 9–13). Young fish that survive initial infection may carry the virus in latent form (13).

Cyprinid herpesvirus 2 (CyHV2) is also known as goldfish hematopoietic necrosis virus. CyHV2 disease is manifested as epizootics in cultured goldfish (*Carassius auratus*). It has also been described as herpesviral hematopoietic necrosis and is distributed worldwide (14). The virus was first isolated from outbreaks in 1992 and 1993 in Japan and was shown to possess the morphology of a herpesvirus (15). Epizootics typically occur in spring and autumn, often with high mortality, and all sizes of goldfish are susceptible. Few gross external lesions are evident on infected fish, but the liver, spleen, and kidneys may appear pale and enlarged, and the spleen and other tissues often display white granular nodules (15). Histopathological findings may typically include mild to severe multifocal or diffuse coagulative necrosis in kidney and spleen tissue. Viral latency is suggested by the presence of CyHV2 DNA in healthy goldfish for lengthy periods (14, 16).

Cyprinid herpesvirus 3 (CyHV3) was originally named koi herpesvirus by Hedrick and colleagues (17) and has also been referred to as carp nephritis and gill necrosis virus (18). CyHV3 disease occurs as epizootics in common and koi carp and has also been described as koi herpesvirus disease. The first samples may have been collected as early as 1996 (19), but the first definitive description of the disease dates from outbreaks that occurred in

Received 17 November 2012 **Accepted** 18 December 2012

Published ahead of print 26 December 2012

Address correspondence to Andrew J. Davison, andrew.davison@glasgow.ac.uk.

* Present address: Thomas B. Waltzek, Department of Infectious Diseases and Pathology, College of Veterinary Medicine, University of Florida, Gainesville, Florida, USA.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JVI.03206-12.

doi:10.1128/JVI.03206-12

the United States and Israel in 1998 (20). This was followed by reports from Israel (21) and several other countries, making it evident that the disease had spread globally (19). The virus was first isolated and proven to be the etiological agent by Hedrick and colleagues (17, 20) and was shown to have the morphology of a herpesvirus in gill tissues from moribund koi (20, 22). Epizootics involving mass mortality occur in spring and autumn, and all ages of carp are susceptible (17, 23–26). Moribund fish display abnormal features of the gills and skin, and secondary infections are common. The kidney and spleen may be enlarged, and the heart may appear flaccid and mottled. The most consistent microscopic lesions are seen in the gills, with hyperplasia and hypertrophy of the branchial epithelium and fusion of the secondary lamellae (17). However, the major portal of entry of the virus is the skin and fins, with disease in gills occurring subsequently (27). Various lines of evidence suggest that CyHV3 may establish latency and that changes in water temperature play a part in reactivation (28–35).

DNA sequence information for CyHV1 and CyHV2 is limited and nonredundant data totaling 13 and 5 kbp, respectively, are available in GenBank (36, 37). In contrast, the complete genome sequences of three CyHV3 strains have been published (38). The CyHV3 reference strain KHV-U (GenBank accession number DQ657948.1) has a genome of 295,146 bp, consisting of a unique region (U) flanked by a terminal direct repeat (TR). The complete genome sequence of a noncyprinid member of the same genus, anguillid herpesvirus 1 (AngHV1; species *Anguillid herpesvirus 1*), which infects eels (*Anguilla anguilla* and *Anguilla japonica*), has also been published (39) and was recently updated as the result of a transcriptomic study (40). The AngHV1 genome (GenBank accession number FJ940765.3) is smaller than that of CyHV3, at 248,526 bp, and similar in structure. Phylogenetically, the three cyprinid herpesviruses (CyHVs) are related closely to each other (36), with CyHV2 and CyHV3 slightly more closely related to each other than either is to CyHV1 (37). AngHV1 is the next most closely related virus, with other alloherpesviruses much more distant (37, 39, 41).

In this paper, we report the complete genome sequences of CyHV1 and CyHV2 and analyze the comparative genomics of the three CyHVs.

MATERIALS AND METHODS

Growth of viruses. CyHV1 strain NG-J1 (Niigata Japan-1) was isolated in 1981 from ornamental Asagi koi carp (3 years old, approximately 2-kg body weight) bearing papillomas (7). CyHV2 strain ST-J1 (Saitama Japan-1) was isolated in 1992 from an ornamental goldfish (0 to 1 year old, 5- to 15-g body weight) suffering from herpesviral hematopoietic necrosis (H. Fukuda, unpublished data). The viruses were grown on a koi fin cell line (KF-1) as described previously (17), passaging the former via an infected cell medium and the latter via infected cells.

DNA sequencing. Preparations of CyHV1 and CyHV2 DNA were isolated from extracellular virions as described previously (36). The SacI restriction endonuclease profiles of these preparations and a virion DNA preparation from CyHV3 reference strain KHV-U (38) were compared by agarose gel electrophoresis and found to differ significantly, as expected (data not shown).

CyHV1 DNA sequence data were generated commercially (Macrogen, Seoul, South Korea) by cloning of randomly fragmented virion DNA into recombinant plasmids, followed by Sanger sequencing. The data were assembled and edited by standard methods, and gaps or ambiguous regions were closed or checked by sequencing PCR products directly or as plasmid clones. The final consensus sequence represented an average 9.2-fold redundancy at each nucleotide over the whole genome. Average redundancy in TR (14.8-fold) was greater than that in U (9.5-fold). The XbaI restriction endonuclease profile generated from the DNA preparation by agarose gel electrophoresis was closely similar to those reported for Japanese CyHV1 isolates (42) (data not shown).

CyHV2 DNA sequence data were generated at the Gene Pool (University of Edinburgh) by next-generation sequencing using standard instruments and methods (Illumina, San Diego, CA). The data comprised a single-end set (5,117,383 reads of 50 nucleotides) and a paired-end set (27,197,794 reads of 48 nucleotides). Both sets (totaling 32,315,177 reads) were treated as a combined, single-end set and assembled *de novo* by using Velvet 0.7.31 (43) as described previously (44). On the assumption that the CyHV2 genome is generally colinear with the CyHV3 genome, a tentative ordering of 25 large contigs was derived by monitoring for predicted amino acid sequence conservation. Iterative assemblies of the combined data sets against speculative junction sequences between contigs were constructed by using Maq 0.7.1 (45) and visualized by using Tablet 1.10.05.21 (46). This permitted the joining of several contigs, thus reducing the number to seven. Gaps or ambiguous regions were closed or checked by sequencing PCR products directly or as plasmid clones. Among these regions were approximately 50 containing tandem reiterations of initially uncertain lengths. Finally, the identity of each nucleotide and the continuity of sequence were confirmed by assembling the combined data sets against the complete genome sequence. A total of 10,250,990 reads (31.72% of the combined sets, consisting of 45.48% of the single-end set and 29.12% of the paired-end set) matched this sequence, the remainder presumably representing contaminating carp cellular DNA or poor-quality reads. The final alignment represented an average 1,802-fold redundancy at each nucleotide over the complete genome. Average redundancy in TR (3,981-fold) was greater than that in U (1,676-fold).

The potential genome termini of CyHV1 and CyHV2 were located on the basis of multiple sequence reads sharing the same end and sequence similarities to the CyHV3 termini and then mapped experimentally as described previously (47) (data not shown). Briefly, virion DNA was flush-ended and ligated to a partially double-stranded adaptor oligonucleotide, followed by PCR amplification utilizing an adaptor primer plus primers near the putative termini. Several plasmid clones of each product were sequenced.

Nucleotide sequence accession numbers. The GenBank accession numbers of the CyHV1 and CyHV2 genome sequences are JQ815363 and JQ815364, respectively. The European Nucleotide Archive accession number of the CyHV2 read data is ERP001303.

RESULTS AND DISCUSSION

Genome structure and composition. Maps of the CyHV1 and CyHV2 genomes are shown in Fig. 1 and 2, respectively, and an updated map of the CyHV3 genome is shown in Fig. 3. The sequences and other data on the locations of genome ends indicate that the CyHV1 and CyHV2 genomes (291,144 and 290,304 bp, respectively) have the same structure as that of CyHV3, with U flanked at each end by TR (Table 1). Thus, the CyHV3 genome remains the largest overall among the CyHVs and, indeed, all sequenced herpesviruses, although the CyHV2 genome has the greatest complexity in terms of the size of U plus one copy of TR. The CyHV2 genome contains an unusual 220-bp inverted repeat that is absent from CyHV1 and CyHV3, with copies to the right of ORF25C and ORF48 (Fig. 2). Experiments involving PCR across the inverted repeat confirmed that the intervening region is present in a single orientation in virion DNA (data not shown). Analyses of nucleotide composition indicate that the CG dinucleotide is strongly underrepresented in the CyHV1 genome (Table 1). In cellular genomes, CG dinucleotide depletion is an evolutionary phenomenon that is thought to have occurred via methylation of

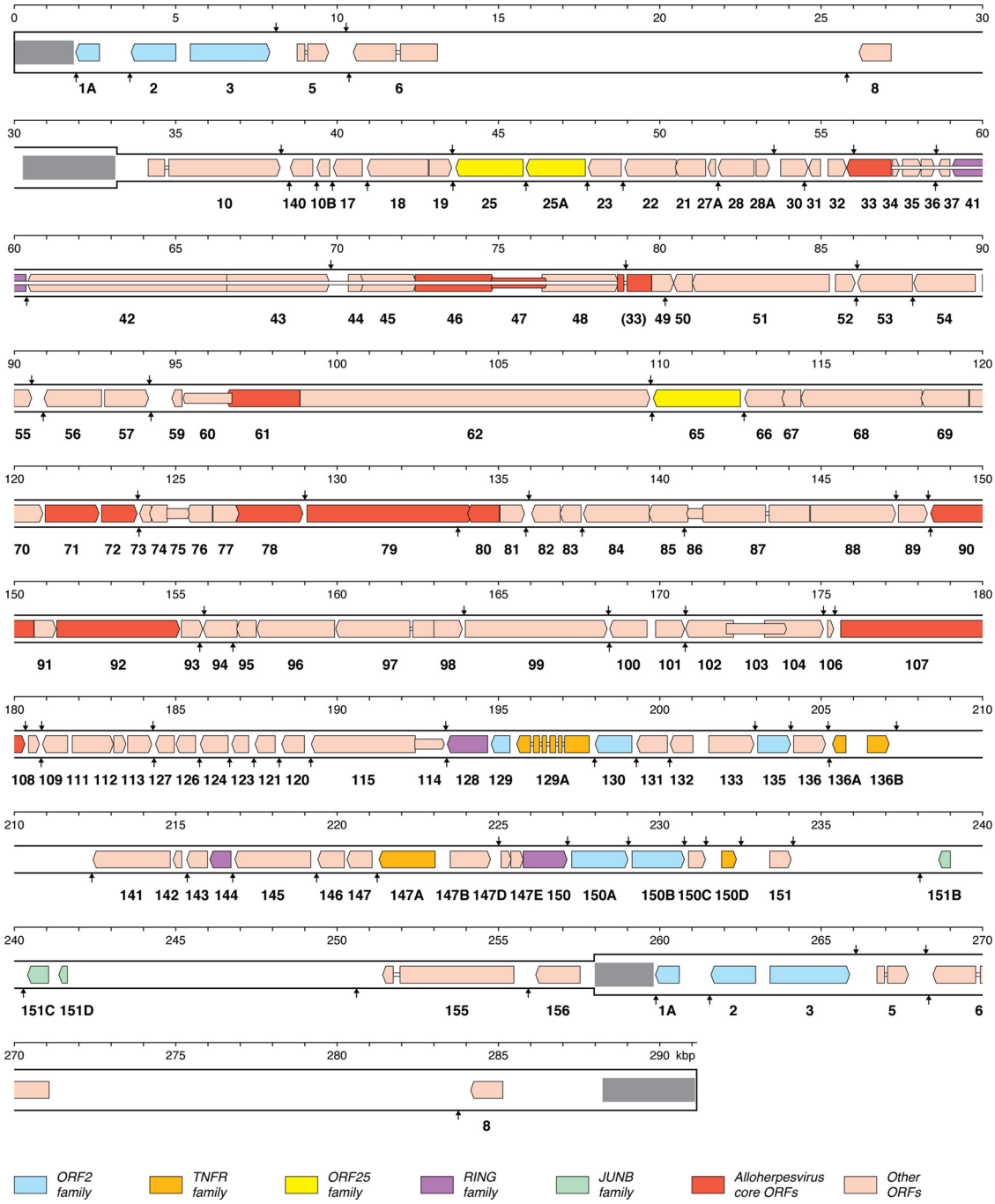

FIG 1 CyHV1 genome map. The terminal direct repeat (TR) is shown in a thicker format than the unique region (U). ORFs predicted to encode functional proteins are indicated by colored arrows (see the key at the foot), with nomenclature lacking the ORF prefix given below. Introns are shown as narrow white bars. Colors of protein-coding regions indicate core ORFs that are convincingly conserved among alloherpesviruses, families of related ORFs, and other ORFs. Telomere-like repeats at the ends of TR are shown by gray-shaded blocks. Predicted poly(A) sites are indicated by vertical arrows above and below the genome for rightward- and leftward-oriented ORFs, respectively.

the C residue followed by spontaneous deamination to yield a TG dinucleotide, which becomes fixed by DNA replication. When this phenomenon has been observed to a significant degree in herpesviruses, it has been attributed to methylation of latent genomes in dividing cell populations (48).

The derived genome sizes (Table 1) are those of the sequences obtained. They are unlikely to equate precisely to actual sizes because each CyHV sequence contains numerous tandem direct reiterations of short sequences, often in complex forms containing partial or dispersed repeats. Reiterations are characteristic of most herpesvirus genomes and are often variable in length, resulting in heterogeneous genome size. In the CyHVs, tandem reiterations are most common in TR, and the most prominent are the telomere-like repeats based on the element TTAGGG (marked in gray in

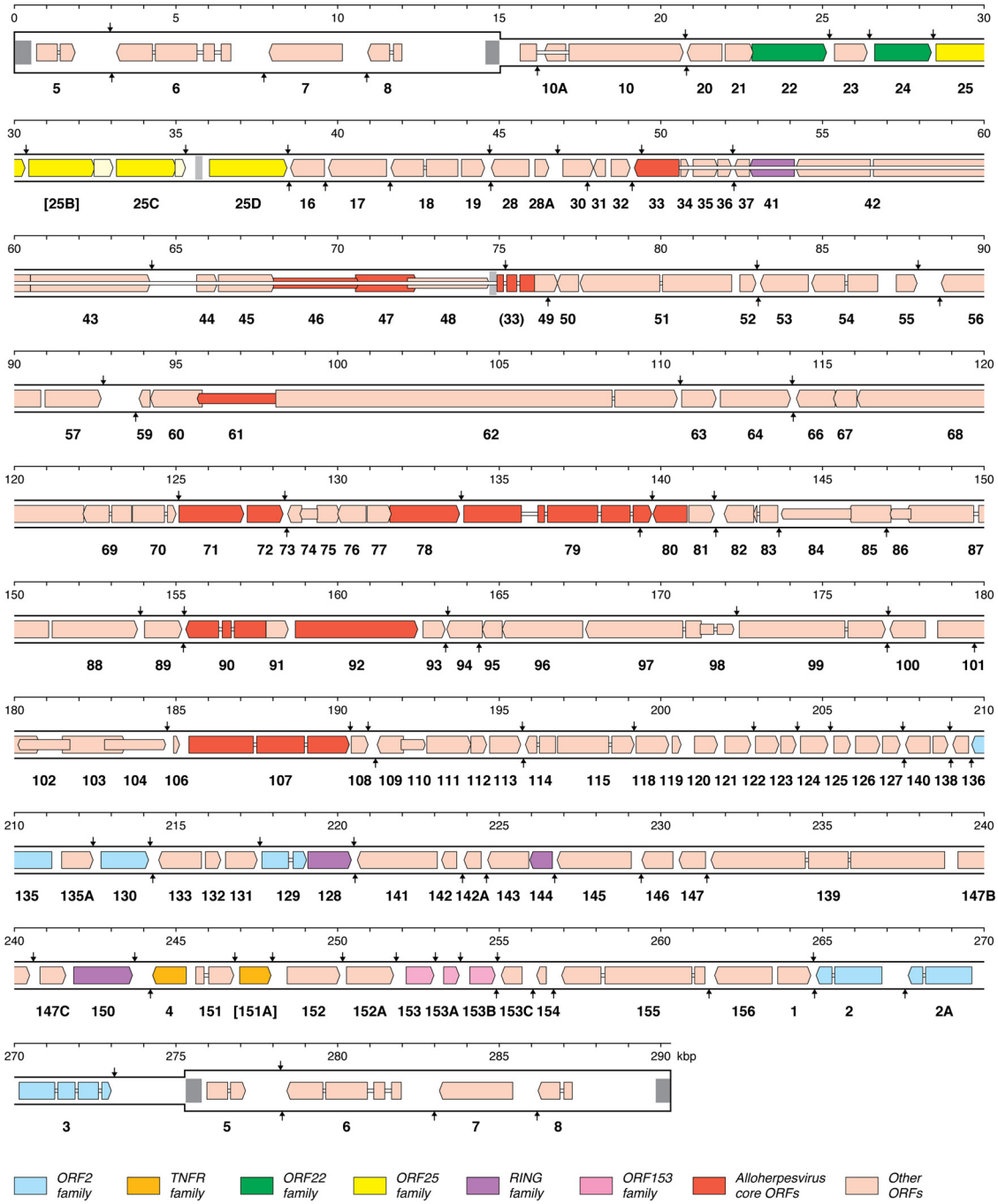

FIG 2 CyHV2 genome map. See the Fig. 1 legend for schematic details. Inverted repeats at approximately 36 and 75 kbp are indicated by light-gray-shaded blocks. Names of fragmented ORFs are given in square brackets, with the ORFs depicted as intact. Pale yellow ORFs downstream from ORF25B and ORF25C represent remnants of additional ORFs in the ORF25 family.

Fig. 1 to 3), which are located very close to the ends of TR and, hence, to the genome ends. CyHV1 has the largest telomere-like repeats, whereas AngHV1 lacks them. Similarly positioned telomere-like repeats are present in the genomes of some members of the family *Herpesviridae*, including Marek's disease virus of chickens (subfamily *Alphaherpesvirinae*) and human herpesviruses 6A, 6B, and 7 (subfamily *Betaherpesvirinae*). There is evidence that such reiterations may mediate integration of herpesvirus genomes into the telomeric regions of host chromosomes *in vitro* and *in vivo* (49, 50). It is not known whether this is the case for the CyHVs.

The CyHVs exhibit regions of similarity at their genome termini (Fig. 4) that may be involved in the generation of linear genomes from concatemeric intermediates presumed to be generated during DNA replication. The greater extent of similarity is near the right end. Conservation of terminal sequences is too low

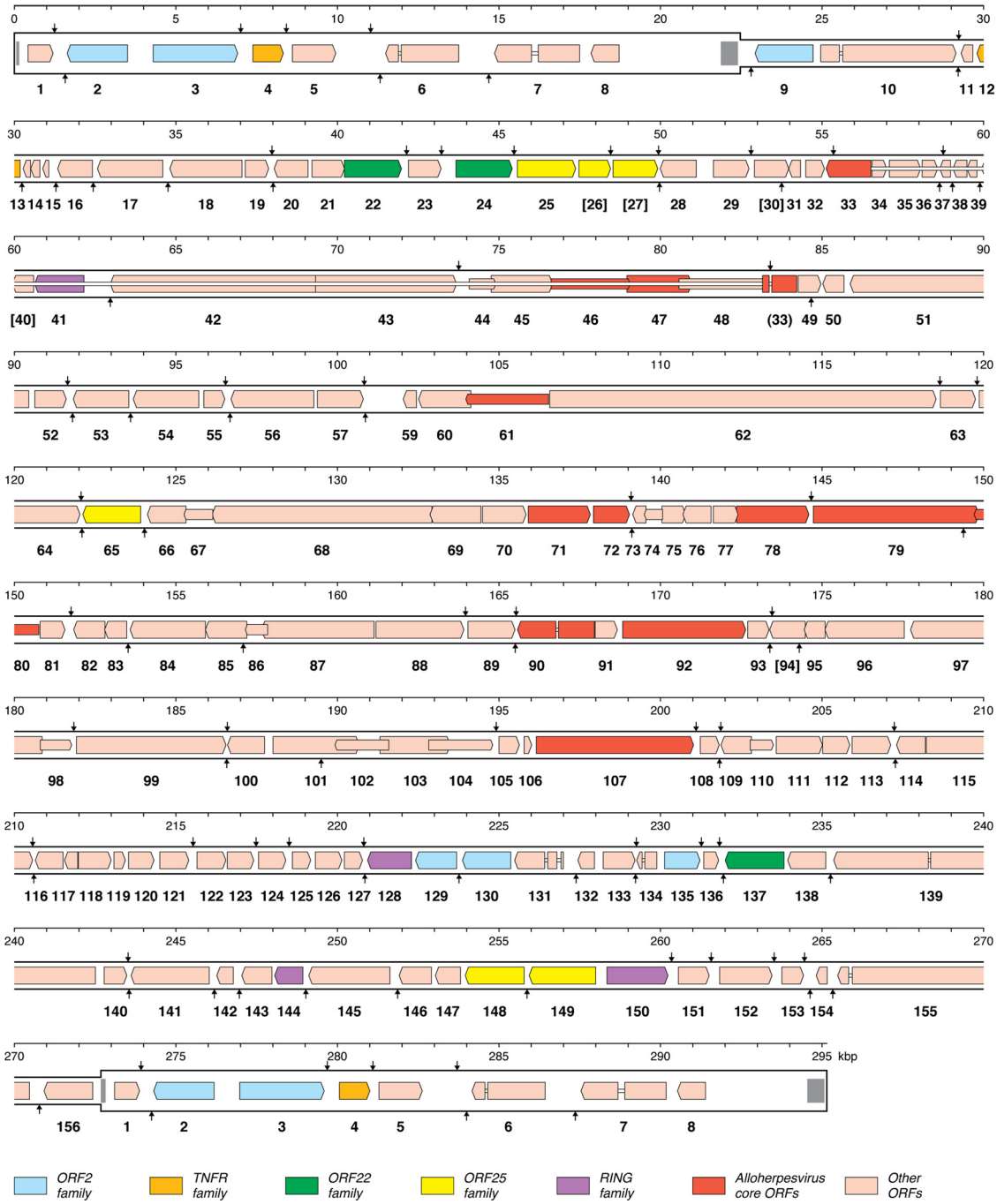

FIG 3 CyHV3 genome map. See the Fig. 1 and 2 legends for schematic details. Updated from reference 38.

to make alignments with AngHV1 rewarding, but there is perhaps some resemblance to CyHV sequences in an A+T-rich region about 40 bp from the right end and the few nucleotides at the right end (Fig. 4). Also, there is a short palindrome at the termini that is especially notable in CyHV3 (CCTTGCC and GGCAAGG) and AngHV1 (CgTTCACC and GGTGAAgG, where lowercase characters are noncomplementary). In relation to flanking genes, the right end is located in a similar context in each CyHV genome, to the right of ORF8. The left end is located in a similar context in CyHV1 and CyHV3, to the right of ORF156. However, the left end in CyHV2 is translocated, being situated between ORF3 and ORF5, at a position that is well within TR in CyHV1 and CyHV3. As a consequence, the counterparts of some open reading frames (ORFs) in TR in CyHV3 (ORF1, ORF2, and ORF3) or CyHV1 (ORF2 and ORF3) are located near the right end of U in CyHV2.

Arrangement of ORFs. A range of standard bioinformatic tools was employed to generate predictions of the locations and arrangements of functional protein-coding ORFs in the CyHV1 and CyHV2 sequences, as used previously for CyHV3 (38). These tools included programs directed at assessing codon preference,

TABLE 1 Genome features of CyHVs and AngHV1

Virus	Size (bp)			Nucleotide composition (%)		No. of ORFs			
	Genome	U	TR	G+C	CG obs/exp[a]	Genome[b]	Unique[c]	U	TR
CyHV1	291,144	224,784	33,180	51.3	57	143	137	131	6
CyHV2	290,304	260,238	15,033	51.7	87	154	150	146	4
CyHV3	295,146	250,208	22,469	59.2	90	163	155	147	8
AngHV1	248,526	227,258	10,634	53.0	94	134	129	124	5

[a] Observed value in relation to the value expected from nucleotide composition if the CG dinucleotide is not depleted.
[b] Number of ORFs in U plus two copies of TR.
[c] Number of ORFs in U plus one copy of TR.

third-position codon G+C bias, and amino acid sequence conservation. In further assessing the gene layout, several general features of herpesvirus genomes were taken into account, including the rarity of extensive overlap between ORFs and the locations of potential poly(A) signals (AATAAA or ATTAAA) downstream from individual ORFs or sets of ORFs transcribed 3′ coterminally. Putative initiating ATG codons were assigned to ORFs on the basis of various criteria, which included favorable codon preference and G+C bias, minimal ORF overlap, and, in appropriate cases, the encoding of predicted signal peptides or conserved amino acid sequences.

The maps (Fig. 1 to 3) were derived solely by bioinformatic means and represent informed interpretations that are susceptible to refinement in the future. Thus, it is possible that some of the ORFs included do not encode functional proteins. Moreover, it is possible that some ORFs encoding functional proteins might have been missed, such as any that are small or spliced, that are conserved weakly or not at all, that overlap other ORFs, or that utilize noncanonical (non-ATG) initiation codons. This caveat is important in light of the finding that many such ORFs are translated in cell culture by human cytomegalovirus (a member of the family *Herpesviridae*, subfamily *Betaherpesvirinae*), although it remains undetermined which, if any, of the proteins produced are functional (51). Our analysis would also not have identified genes specifying RNAs that do not encode functional proteins, as exemplified again by human cytomegalovirus (52, 53). However, no such RNAs were detected in a transcriptomic study of AngHV1 (40).

The initial ORF designations were modified from predictions of splicing, which were made on the basis of potential donor and acceptor sites in positions in which their use would result in the significant extension of an ORF or the joining together of two conserved ORFs located in different reading frames or separated by stop codons in one genome but represented as a single ORF in another. In some instances, introns within ORFs were predicted from amino acid sequence alignments, though these were necessarily more tentative (e.g., in CyHV2 ORF70). Splicing between protein-coding regions was predicted to be more common in CyHV2 (affecting 29 ORFs, not counting duplicates in TR) than CyHV1 and CyHV3 (eight and nine ORFs, respectively). It was not possible to assess with any confidence additional splicing that would contribute short, upstream exons to established ORFs, whether translated (thus extending the 5′ end of the ORF) or nontranslated (thus contributing a 5′ untranslated region). However, such splicing may exist, since several instances have been identified in AngHV1 (40). Indeed, there were some bioinformatic hints of splicing between nontranslated regions. One example involves ORF49, which in each CyHV genome is followed immediately by a potential splice donor site that may be linked to a potential splice acceptor site upstream from ORF52.

The nomenclature used for CyHV1 and CyHV2 ORFs is based on that for CyHV3 (38), in which ORFs are numbered along the genome from ORF1 to ORF156, with ORF1 to ORF8 duplicated in TR (Fig. 3). CyHV1 and CyHV2 ORFs that have counterparts in CyHV3 were given the same names as in CyHV3, and ORFs that lack counterparts were designated by alphabetically extended names (e.g., ORF10A in CyHV2 only, ORF10B in CyHV1 only, and ORF28A in both). The nomenclature systems used for other alloherpesviruses (including AngHV1), for which some ORFs are mentioned below, differ from that of the CyHVs. In general, CyHV ORFs are packed closely together. However, several sizeable regions were not predicted to encode functional proteins. The most-extensive regions in this category are present in CyHV1, specifically in TR near the right end of U, and are among the largest noted among the herpesviruses. It is possible that these regions have functions that are not concerned primarily with encoding proteins, for example, in specifying nontranslated transcripts.

The numbers of functional protein-coding ORFs predicted in the CyHV genomes (Table 1) should be viewed as best estimates, as there was uncertainty in some aspects of the interpretation. Examples include the ORF44-ORF45 region, where it is possible that ORF45 is extended at its 5′ end, perhaps replacing ORF44, and the ORF101 region, where an additional ORF might be incorporated into the antiparallel strand, extending into the region between ORF100 and ORF101.

FIG 4 Aligned sequences at the CyHV genome termini. Terminal nucleotides are highlighted in black, conserved residues are in gray, and gaps are indicated by hyphens. The corresponding unaligned AngHV1 sequences are shown below. For each sequence, the ellipsis indicates the remainder of the genome.

The original CyHV3 map was derived in part on the basis of comparisons with incomplete sequence data for CyHV1 (38). We made several refinements as a result of comparisons with the complete data for CyHV1 and CyHV2 (Fig. 3). These include the incorporation of a protein-coding region located between the original ORF131 and ORF132 that potentially specifies an amino acid sequence highly similar to that of *Danio rerio* hypothetical type 1 membrane protein LOC562542 (GenBank accession number XM_701036.5), of which the function is unknown. This region contains counterparts of the first two exons of the three-exon cellular gene, with the splice sites conserved. We propose tentatively that these exons are spliced to the 5′ end of ORF131, which is not related to LOC562542, thus adding a signal peptide and completing the encoding of a predicted type 1 membrane protein. The two LOC562542 exons are absent from CyHV1 and CyHV2, and in each case, the predicted ORF131 type 1 membrane protein has an integral signal peptide. Additional adjustments to the CyHV3 map included shortening the 5′ ends of ORF33, ORF56, ORF105, and ORF119, extending the 5′ end of ORF94, and extending the 3′ end of ORF6 by splicing. In addition, the marginal ORF58 was removed.

ORFs conserved among alloherpesviruses. Table 2 summarizes the relationships among, and features of, the CyHV genomes and also includes data for AngHV1 ORFs that have CyHV orthologs. Functional assignments were made in some instances on the basis of strong amino acid conservation with well-characterized proteins, and in other instances, they depended on conservation of motifs and were made less confidently. A full version of the information in Table 2 is provided in data set S1 in the supplemental material, which is sortable by ORF order in each genome and includes data on amino acid identity that support ortholog identification. The comparisons (not shown in data set S1 in the supplemental material) also involved the other fully sequenced alloherpesviruses, namely, ictalurid herpesvirus 1 (IcHV1; channel catfish virus; species *Ictalurid herpesvirus 1*, genus *Ictalurivirus*; GenBank accession number M75136.1) (54), ranid herpesvirus 1 (RaHV1; Lucké tumor herpesvirus; species *Ranid herpesvirus 1*, genus *Batrachovirus*; GenBank accession number DQ665917.1) (55), and ranid herpesvirus 2 (RaHV2; frog virus 4; species *Ranid herpesvirus 2*, genus *Batrachovirus*; GenBank accession number DQ665652.1) (55). Twelve CyHV ORFs (termed core ORFs) are conserved convincingly in all sequenced alloherpesviruses and were presumably inherited from a common ancestor (39). These ORFs encode proteins that are predicted to be involved in DNA replication (ORF79, DNA polymerase catalytic subunit; ORF71, helicase-primase helicase subunit; ORF46, helicase-primase primase subunit), DNA packaging (ORF33, DNA packaging terminase subunit 1), and capsid morphogenesis (ORF92, major capsid protein; ORF72, capsid triplex subunit 2; ORF78, capsid maturation protease), as well as proteins of unknown function (ORF47, protein Allo64, possible DNA packaging terminase subunit 2 [see the following paragraph]; ORF61, protein Allo54; ORF80, protein Allo60; ORF90, protein Allo37; ORF107, protein Allo56).

The ORF encoding DNA packaging terminase subunit 1 (ORF33 in the CyHVs) is well conserved among members of the order *Herpesvirales* (1). In contrast, the ORF encoding DNA packaging terminase subunit 2 is conserved among members of the family *Herpesviridae*, but its anticipated ortholog in alloherpesviruses is not readily apparent. However, there is weak similarity to the core CyHV ORF47 and its counterparts in other alloherpesviruses (Fig. 5a). The conserved region contains a cysteine-rich motif that is conserved throughout members of the family *Herpesviridae* and may represent a metal ion-binding domain (56).

The detection of only 12 core genes illustrates the substantially greater divergence that exists among members of the family *Alloherpesviridae* in comparison with members of the family *Herpesviridae*, for which about 40 genes form the core set (1). There is some evidence that two additional CyHV ORFs (ORF99 and ORF66) may also belong to the alloherpesvirus core set, although the cases for these assignments depend on weak sequence similarities and other discerned resemblances.

The predicted type 1 membrane protein encoded by ORF99 has clear orthologs in all sequenced alloherpesviruses except AngHV1 (39). The CyHV3 and IcHV1 proteins have been shown from proteomic studies to be located in virions, the latter fractionating specifically with the virion envelope (57, 58). They are weakly similar to the virion envelope protein encoded by AngHV1 ORF67 (59), which is also distantly related to the spike proteins of a member of the subfamily *Torovirinae* in the family *Coronaviridae* (39). This indicates that the ancestor of ORF99 and its counterparts in AngHV1 and other alloherpesviruses might have been derived from an ancient gene capture event involving a coronavirus (59), with similarity to the original gene having been retained only in AngHV1. On the other hand, it is possible that the ORF99 counterpart has been lost from the AngHV1 lineage and that any similarities to ORF67 are incidental.

The argument for inclusion of ORF66 in the core set is based on the fact that capsid structure is conserved in all characterized members of the order *Herpesvirales*, with the three main proteins making up the shell being the major capsid protein and the two subunits of the capsid triplex (60–62). Proteomic findings, based on the identification of IcHV1 and AngHV1 capsid proteins in the anticipated stoichiometry and the bioinformatic detection of their orthologs in CyHV3, have identified strong candidates for only two of these in CyHVs (ORF92, major capsid protein; ORF72, capsid triplex subunit 2). The observation that capsid triplex subunit 1 is the least conserved of the three proteins in members of the family *Herpesviridae* (58) suggests that the corresponding protein in alloherpesviruses may be too diverged for sequence similarities to be readily apparent across the family. The best candidates from proteomic studies of capsids for capsid triplex subunit 1 are ORF53 in IcHV1 and its ortholog ORF42 in AngHV1 (58, 59). An ortholog in CyHV3 is not readily apparent, but ORF66, which encodes a virion protein, is a possibility (57, 59). No sequence relationships have been demonstrated, but a short, conserved motif is evident (Fig. 5b).

It is also possible that another ORF should be added to the maps and counted among the core set. This is based on a recognized feature of the capsid maturation protease gene in members of the family *Herpesviridae*, specifically that an abundant, 3′-coterminal transcript initiates within the relevant ORF and produces an N-terminally truncated version of the capsid maturation protease that functions as the capsid scaffold protein, which is replaced by DNA during capsid maturation (62). There is evidence from transcriptomics that a similar situation exists in AngHV1 (40). If this arrangement also pertains to the CyHVs, the relevant core ORF would be named ORF78.5 because it amounts to a 5′ truncation of ORF78. However, given the lack of evidence for the expression of this ORF in CyHVs, it has not been added to the maps or to the totals of gene numbers cited above.

TABLE 2 Features of predicted CyHV functional protein-coding regions

ORF[a]	Family	Orientation in each CyHV genome[b]			AngHV1 ortholog	Protein	Feature(s)[c]
		1	2	3			
ORF1				**R**		Protein ORF1	TMD
ORF1A	ORF2	**L**				Protein ORF1A	
ORF2	ORF2	**L**		**L**		Protein ORF2	
ORF3	ORF2	**R**		**R**		Protein ORF3	
ORF4	TNFR			**R**		Protein ORF4	SP; *immune regulation*
ORF5		**R**	**R**	**R**		Protein ORF5	TMD
ORF6		**L**	**L**	**L**		Protein ORF6	
ORF7			**L**	**L**		Protein ORF7	
ORF8		**L**	**L**	**L**		Protein ORF8	
ORF9	ORF2			L		Protein ORF9	
ORF10		R	R	R	ORF74[e]	Protein ORF10	
ORF10A			L			Protein ORF10A	SP
ORF10B		L				Protein ORF10B	TMD
ORF11				L		Protein ORF11	
ORF12	TNFR			L		Protein ORF12	SP; *immune regulation*
ORF13				L		Protein ORF13	
ORF14				L		Protein ORF14	TMD
ORF15				L		Protein ORF15	TMD
ORF16			L	L		Membrane protein ORF16	Type 3 MP; SP; 7 TMDs; similar to GPCRs; *intracellular signaling*
ORF17		L	L	L		Protein ORF17	
ORF18		L	L	L		Protein ORF18	
ORF19		R	R	R	ORF123	Deoxyguanosine kinase	*Nucleotide metabolism*
ORF20			L	L		Protein ORF20	
ORF21		L	R	R		Protein ORF21	
ORF22	ORF22	L	R	R		Protein ORF22	
ORF23		L	R	R	ORF96	Ribonucleotide reductase subunit 2	*Nucleotide metabolism*
ORF24	ORF22		R	R		Protein ORF24	
ORF25A	ORF25	L				Membrane protein ORF25A	Type 1 MP; Igd
ORF25	ORF25	L	R	R		Membrane protein ORF25[f]	Type 1 MP; Igd
ORF25B	ORF25		[R]			Membrane protein ORF25B	Type 1 MP; Igd
ORF25C	ORF25		R			Membrane protein ORF25C	Type 1 MP
ORF25D	ORF25		R			Membrane protein ORF25D	Type 1 MP
ORF26	ORF25			[R]		Membrane protein ORF26	Type 1 MP
ORF27	ORF25			[R]		Membrane protein ORF27	Type 1 MP; Igd
ORF27A		L				Protein ORF27A	TMD
ORF28		L	L	L		Protein ORF28	NAD(P)-binding Rossmann-fold domain; similar to bacterial NAD-dependent epimerase/dehydratase
ORF28A		R	R			Protein ORF28A	SP; Igd; in CyHV1, possibly spliced to a downstream exon to specify a type 1 MP similar to neuroplastin
ORF29				R		Membrane protein ORF29	Type 3 MP; 8 TMDs
ORF30		R	R	[R]		Membrane protein ORF30	Type 1 MP
ORF31		L	L	L		Protein ORF31[f]	Similar to eukaryotic PLAC8 proteins
ORF32		R	R	R		Protein ORF32[f]	SP; similar to a family of Singapore grouper iridovirus proteins
ORF33[d]		L	L	L	ORF10	DNA packaging terminase subunit 1	ATPase domain; *DNA encapsidation*
ORF34		R	R	R		Protein ORF34[f]	SP
ORF35		R	R	R		Protein ORF35[f]	
ORF36		R	R	R		Protein ORF36[f]	
ORF37		L	L	L		Protein ORF37	TMD
ORF38				L		Protein ORF38	TMD
ORF39				L		Membrane protein ORF39	Type 3 MP; 2 TMDs; SP
ORF40				[L]		Membrane protein ORF40	Type 1 MP; Igd
ORF41	RING	L	L	L		Protein ORF41	
ORF42		L	L	L	ORF18	Protein ORF42[f]	
ORF43		R	R	R	ORF19[e]	Protein ORF43[f]	

(Continued on following page)

TABLE 2 (Continued)

ORF[a]	Family	Orientation in each CyHV genome[b]			AngHV1 ortholog	Protein	Feature(s)[c]
		1	2	3			
ORF44		R	R	R		Protein ORF44[f]	
ORF45		R	R	R	ORF20	Protein ORF45[f]	
ORF46[d]		R	R	R	ORF21	Helicase-primase primase subunit	Putative; *DNA replication*
ORF47[d]		R	R	R	ORF22	Protein Allo64	Weakly similar to herpesvirus DNA packaging terminase subunit 2; *DNA replication*
ORF48		R	R	R	ORF23[e]	Protein ORF48	Similar to protein kinases; *protein phosphorylation*
ORF49		R	R	R		Protein ORF49	
ORF50		L	L	L		Protein ORF50	
ORF51		L	L	L	ORF34	Protein ORF51[f]	
ORF52		R	R	R		Protein ORF52	SP
ORF53		L	L	L		Protein ORF53	
ORF54		L	L	L		Protein ORF54	Putative zinc-binding domain
ORF55		R	R	R		Thymidine kinase	*Nucleotide metabolism*
ORF56		L	L	L		Protein ORF56	
ORF57		R	R	R	ORF35	Protein ORF57[f]	Similar to crocodilepox virus protein CRV155
ORF59		L	L	L		Protein ORF59[f]	TMD
ORF60		L	L	L	ORF81	Protein ORF60[f]	
ORF61[d]		L	L	L	ORF82	Protein Allo54	
ORF62		R	R	R	ORF83[e]	Protein ORF62[f]	OTU-like cysteine protease domain
ORF63			R	R		Protein ORF63	
ORF64			R	R		Membrane protein ORF64	Type 3 MP; 12 TMDs; SP; similar to equilibrative nucleoside transporter ENT1
ORF65	ORF25	L		L		Membrane protein ORF65[f]	Type 1 MP; Igd
ORF66		L	L	L	ORF42	Capsid triplex subunit 1[f]	Putative; *capsid morphogenesis*
ORF67		L	L	L	ORF41	Protein ORF67	
ORF68		L	L	L	ORF40[e]	Protein ORF68[f]	Similar to myosin and related proteins
ORF69		L	L	L	ORF39[e]	Protein ORF69[f]	
ORF70		R	R	R	ORF38	Protein ORF70[f]	
ORF71[d]		R	R	R	ORF37	Helicase-primase helicase subunit	Putative; *DNA replication*
ORF72[d]		R	R	R	ORF36	Capsid triplex subunit 2[f]	Putative; *capsid morphogenesis*
ORF73		L	L	L		Protein ORF73	
ORF74		L	L	L	ORF63	Protein ORF74	
ORF75		R	R	R	ORF62	Protein ORF75	
ORF76		L	L	L	ORF59	Protein ORF76	
ORF77		R	R	R	ORF58[e]	Protein ORF77	
ORF78[d]		R	R	R	ORF57	Capsid maturation protease[f]	Putative; *capsid morphogenesis*
ORF79[d]		R	R	R	ORF55	DNA polymerase catalytic subunit	*DNA replication*
ORF80[d]		L	L	L	ORF52	Protein Allo60	
ORF81		R	R	R	ORF51	Membrane protein ORF81[f]	Type 3 MP; 4 TMDs
ORF82		L	L	L	ORF50	Membrane protein ORF82	Type 3 MP; 4 TMDs
ORF83		L	L	L	ORF49	Membrane protein ORF83	Type 3 MP; 4 TMDs
ORF84		L	L	L	ORF48[e]	Protein ORF84[f]	
ORF85		L	L	L	ORF47[e]	Protein ORF85	
ORF86		L	L	L	ORF46[e]	Protein ORF86	
ORF87		L	L	L	ORF45	Protein ORF87	
ORF88		R	R	R	ORF44	Protein ORF88	
ORF89		R	R	R	ORF43	Protein ORF89[f]	
ORF90[d]		L	L	L	ORF100	Protein Allo37[f]	
ORF91		R	R	R	ORF103	Protein ORF91	
ORF92[d]		R	R	R	ORF104	Major capsid protein[f]	*Capsid morphogenesis*
ORF93		R	R	R	ORF105	Protein ORF93	
ORF94		L	L	[L]		Protein ORF94	Similar to trypsin-like serine proteases; signal anchor

(Continued on following page)

TABLE 2 (Continued)

ORF[a]	Family	Orientation in each CyHV genome[b]			AngHV1 ortholog	Protein	Feature(s)[c]
		1	2	3			
ORF95		L	L	L		Protein ORF95[f]	
ORF96		L	L	L	ORF31	Protein ORF96	
ORF97		L	L	L	ORF30	Protein ORF97[f]	
ORF98		R	R	R	ORF29	Uracil-DNA glycosylase	*DNA repair*
ORF99		R	R	R	ORF67[e]	Membrane protein Allo46[f]	Type 1 MP
ORF100		L	L	L		Protein ORF100	
ORF101		R	R	R		Protein ORF101	
ORF102		L	L	L	ORF97	Protein ORF102	
ORF103		R	R	R		Protein ORF103	
ORF104		R	R	R	ORF87[e]	Protein ORF104	Similar to protein kinases; *protein phosphorylation*
ORF105				R		Protein ORF105	Marginal assignment
ORF106		R	R	R		Protein ORF106	TMD
ORF107[d]		R	R	R	ORF98	Protein Allo56	
ORF108		R	R	R		Protein ORF108[f]	TMD
ORF109		L	L	L		Protein ORF109	
ORF110			R	R		Protein ORF110	
ORF111		R	R	R		Protein ORF111	
ORF112		R	R	R		Protein ORF112[f]	Double-stranded nucleic acid-binding domain (helix-turn-helix)
ORF113		R	R	R		Protein ORF113	
ORF114		R	L	L		Membrane protein ORF114	Type 3 MP; 8 TMDs; in CyHV3, similar to *Danio rerio* LOC569866
ORF115		L	R	R		Membrane protein ORF115[f]	Type 1 MP
ORF116				R		Membrane protein ORF116	Type 1 MP
ORF117				L		Protein ORF117	TMD
ORF118			R	R		Protein ORF118	
ORF119			R	R		Protein ORF119	SP
ORF120		L	R	R		Protein ORF120	
ORF121		L	R	R		Protein ORF121	
ORF122			R	R		Protein ORF122	
ORF123		L	R	R	ORF5	Deoxyuridine triphosphatase[f]	*Nucleotide metabolism*
ORF124		L	R	R		Membrane protein ORF124	Type 1 MP
ORF125			R	R		Protein ORF125	SP
ORF126		L	R	R		Membrane protein ORF126	Type 1 MP
ORF127		L	R	R	ORF117	Protein ORF127	SP
ORF128	RING	L	R	L		Protein ORF128	Similar to SPRY and TRIM proteins
ORF129	ORF2	L	R	L		Protein ORF129	
ORF129A	TNFR	L				Protein ORF129A	SP; fibronectin type 3 domain; *immune regulation*
ORF130	ORF2	L	R	L		Protein ORF130	
ORF131		L	R	L		Membrane protein ORF131[f]	Type 1 MP; exons 1 and 2 in CyHV3 are similar to exons 1 and 2 of *Danio rerio* LOC562542
ORF132		L	R	L		Membrane protein ORF132[f]	Type 1 MP
ORF133		R	L	R		Protein ORF133	
ORF134				L		Interleukin-10	SP; *immune regulation*
ORF135A			R			Protein ORF135A	
ORF135	ORF2	R	L	R		Protein ORF135	
ORF136		R	L	R		Membrane protein ORF136[f]	Type 1 MP
ORF136A	TNFR	L				Protein ORF136A	SP; *immune regulation*
ORF136B	TNFR	R				Protein ORF136B	SP; *immune regulation*
ORF137	ORF22			L		Protein ORF137[f]	
ORF138			R	L		Membrane protein ORF138	Type 1 MP
ORF139			L	L		Membrane protein ORF139	Type 1 MP; similar to poxvirus B22R proteins
ORF140		L	L	R	ORF77	Thymidylate kinase	*Nucleotide metabolism*
ORF141		L	L	L	ORF116	Ribonucleotide reductase subunit 1	*Nucleotide metabolism*

(Continued on following page)

TABLE 2 (Continued)

ORF[a]	Family	Orientation in each CyHV genome[b]			AngHV1 ortholog	Protein	Feature(s)[c]
		1	2	3			
ORF142		L	L	L		Protein ORF142	
ORF142A			L			Membrane protein ORF142A	Type 1 MP
ORF143		L	L	L		Protein ORF143	
ORF144	RING	L	L	L		Protein ORF144	
ORF145		L	L	L		Protein ORF145	
ORF146		L	L	L		Membrane protein ORF146	Type 1 MP; Igd
ORF147		L	L	L		Protein ORF147	SP
ORF147A	TNFR	L			ORF65	Membrane protein ORF147A	Type 3 MP; SP; similar to TNFR and perforin; *probably immune regulation*
ORF147B		R	R		ORF66[e]	Membrane protein ORF147B	Type 1 MP
ORF147C			R			Protein ORF147C	TMD
ORF147D		R				Protein ORF147D	TMD
ORF147E		R				Protein ORF147E	TMD
ORF148	ORF25			L		Membrane protein ORF148[f]	Type 1 MP; Igd
ORF149	ORF25			L		Membrane protein ORF149[f]	Type 1 MP; Igd
ORF150	RING	R	R	R		Protein ORF150	
ORF150A	ORF2	R				Protein ORF150A	
ORF150B	ORF2	R				Protein ORF150B	
ORF150C		R				Membrane protein ORF150C	Type 1 MP; Igd
ORF150D	TNFR	R				Protein ORF150D	SP; *immune regulation*
ORF151		R	R	R		Protein ORF151	
ORF151A	TNFR		[R]			Protein ORF151A	
ORF151B	JUNB	L				Protein ORF151B	Leucine zipper domain; region immediately upstream also potentially encodes part of a JunB-like protein; *transcriptional regulation*
ORF151C	JUNB	L				Protein ORF151C	DNA-binding domain and partial leucine zipper domain; *transcriptional regulation*
ORF151D	JUNB	L				Protein ORF151D	DNA-binding and leucine zipper domains; *transcriptional regulation*
ORF152			R	R		Protein ORF152	
ORF152A			R			Membrane protein ORF152A	Type 3 MP; 8 TMDs
ORF153	ORF153		R	R		Membrane protein ORF153	Type 3 MP; 4 TMDs
ORF153A	ORF153		R			Membrane protein ORF153A	Type 3 MP; 4 TMDs
ORF153B	ORF153		R			Membrane protein ORF153B	Type 3 MP; 4 TMDs
ORF153C			L			Membrane protein ORF153C	Type 1 MP
ORF154			L	L		Protein ORF154	TMD
ORF155		L	L	L		Protein ORF155	
ORF156		L	L	L		Protein ORF156	
ORF1			R	**R**		Protein ORF1	TMD
ORF1A	ORF2	**L**				Protein ORF1A	
ORF2	ORF2	**L**	L	**L**		Protein ORF2	
ORF2A	ORF2		L			Protein ORF2A	
ORF3	ORF2	**R**	R	**R**		Protein ORF3	
ORF4	TNFR		L	**R**		Protein ORF4	SP; *immune regulation*
ORF5		**R**	**R**	**R**		Protein ORF5	TMD
ORF6		**L**	**L**	**L**		Protein ORF6	
ORF7			**L**	**L**		Protein ORF7	
ORF8		**L**	**L**	**L**		Protein ORF8	

[a] ORFs are listed in a default order, basically in relation to their arrangement in CyHV3.
[b] ORFs are oriented rightward (R) or leftward (L) in the CyHV1 (1), CyHV2 (2), and CyHV3 (3) genomes. Absence of notation indicates absence of the ORF. ORFs in square brackets are fragmented. ORFs located in TR are in bold. Regions that are translocated or inverted (or both) in CyHV1 or CyHV2 relative to CyHV3 are shaded.
[c] MP, predicted membrane protein; SP, predicted signal peptide; TMD, predicted transmembrane domain; Igd, predicted immunoglobulin domain; GPCR, G protein-coupled receptor; OTU, ovarian tumor. Proposed functional categories are in italic type.
[d] Conserved convincingly among all sequenced alloherpesviruses.
[e] Tentative assignment.
[f] Virion protein in CyHV3 (57).

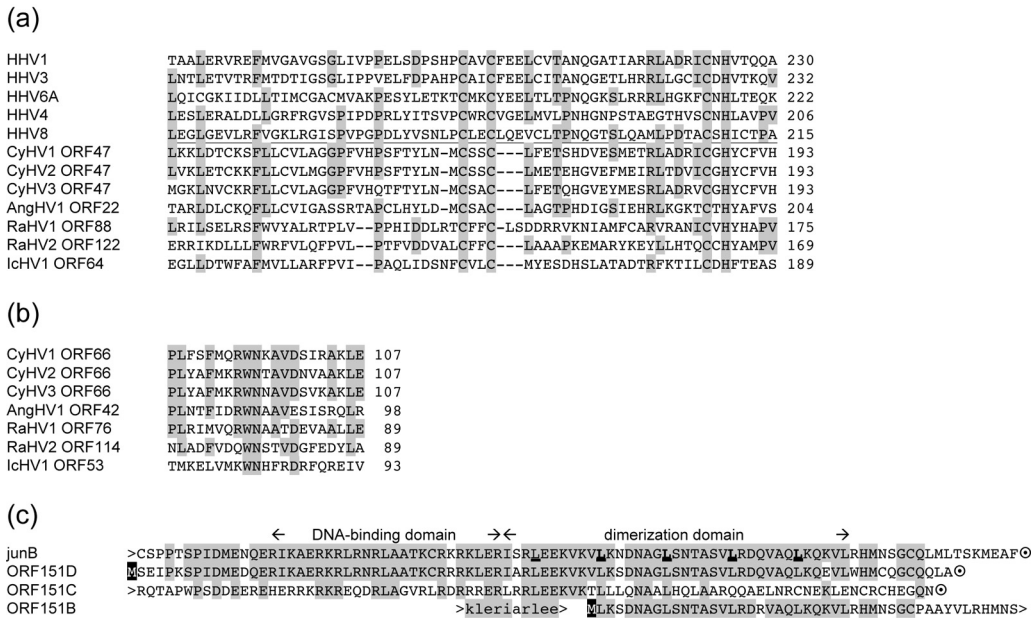

FIG 5 Amino acid sequence alignments of selected CyHV proteins. (a) A motif conserved among CyHV ORF47 and its counterparts in other alloherpesviruses and DNA packaging terminase subunit 2 in various human herpesviruses (HHVs), as representatives of the family *Herpesviridae* (herpes simplex virus type 1 [human herpesvirus 1 {HHV1}], GenBank accession number X14112; varicella-zoster virus [HHV3], GenBank accession number X04370; human herpesvirus 6A [HHV6A], GenBank accession number X83413; Epstein-Barr virus [HHV4], GenBank accession number V01555; Kaposi's sarcoma-associated herpesvirus [HHV8], GenBank accession number U75698). Residues shaded gray are conserved in at least two members of each group, which are separated by a line. (b) A motif conserved in CyHV ORF66 and its counterparts in other alloherpesviruses. Residues shaded gray are conserved among the three CyHVs and at least one other alloherpesvirus. (c) JunB-related sequences encoded by CyHV1. The ORF151B, ORF151C, and ORF151D sequences are shown aligned against the carp JunB sequence (GenBank accession number U81506), and conserved residues (shaded gray) and the DNA-binding and dimerization domains are indicated, with the residues forming the leucine zipper underlined. Initiating methionine residues are shaded in black, C termini are indicated by bull's-eyes, and N- and C-terminal regions that are not shown are denoted by arrowheads. Thus, the alignment includes the entire coding region of the ORF151D protein, the C-terminal regions of carp JunB and the ORF151C protein, and the N-terminal region of the ORF151B protein. An additional short sequence similar to part of JunB is encoded 1 bp upstream from ORF151B in an alternative reading frame and is shown in lowercase type. Residues shaded gray are conserved among at least three sequences.

In addition to the 12 ORFs that belong to the core set and the two or three others that may also belong to it, ORF147A is conserved between CyHV1 and AngHV1 and possibly ORF147B is conserved among CyHV1, CyHV2, and AngHV1, although neither is present in all of the CyHVs and AngHV1. It is possible that one or both of these ORFs were also inherited from a common ancestor of the CyHVs but has since been lost in specific lineages.

ORFs conserved among CyHVs. The CyHV genomes share a total of 120 orthologous ORFs (counting duplicates) arranged with a large degree of colinearity (Table 2). Thus, each contains a large, central section (ORF28-ORF113) in which conserved genes are ordered equivalently. However, various rearrangements to the left and right of the central section are evident in CyHV1 and CyHV2 compared with CyHV3 (Fig. 6). In the region to the left, the ORF21-ORF25 segment is inverted in CyHV1 and the ORF16-ORF19 segment is translocated to the right of ORF25D in CyHV2.

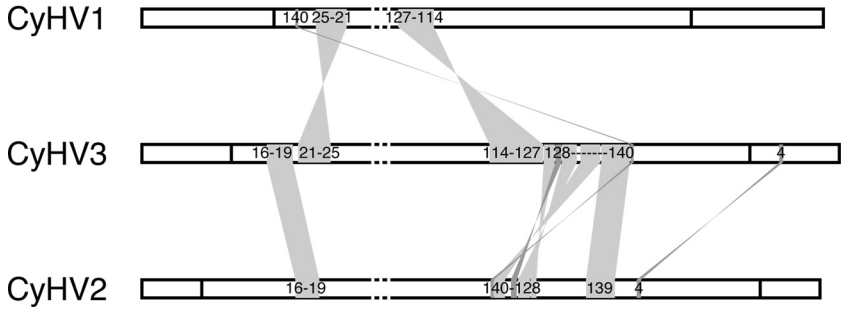

FIG 6 Relative gene layout in CyHVs. The genomes are shown to scale as horizontal bars with the boundaries of the terminal direct repeat (TR) marked by vertical black lines. Omission of the region from 60 to 180 kbp in each genome is implied by the dotted lines. The positions of ORFs that differ in location or orientation between CyHV1 and CyHV3 or CyHV2 and CyHV3 are indicated by gray shading, with the ORF prefix omitted from names. The situation in the region containing ORF128 to ORF140 is complex because ORF130 and ORF139 are transposed in CyHV2, the former remaining within this region but the latter not remaining (for details, see Fig. 2 and 3).

Also, ORF140 is inverted and translocated to the right of ORF10 in CyHV1. In the region to the right, the ORF114-ORF127 segment is inverted in CyHV1. In CyHV2, the ORF128-ORF140 segment is inverted, and, from this segment, ORF130 is translocated to the right of ORF133 and ORF139 is translocated to the right of ORF147. Also, ORF4 is inverted and translocated to the right of ORF150 in CyHV2, effectively transferring it from TR to U.

Of the 76 ORFs that are not conserved among all three CyHVs, 50 encode predicted membrane-associated proteins (i.e., they possess potential signal peptides, transmembrane domains, or both) and 34 are members of paralogous gene families (i.e., they share amino acid sequence similarity with other CyHV ORFs). These nonconserved ORFs presumably provide functions that fit the individual viruses to their biological niches. Eighteen ORFs are present only in CyHV1 (including ORF147A mentioned above and the JUN gene family described below), 13 are present only in CyHV2, and 21 are present only in CyHV3 (including ORF134, which encodes an interleukin-10-related protein) (63). ORFs conserved between two CyHVs but absent from the third number two in CyHV1 and CyHV2 (including ORF147B mentioned above), three in CyHV1 and CyHV3, and 19 in CyHV2 and CyHV3.

ORFs conserved in AngHV1. CyHV3 shares 40 clear orthologs with AngHV1 (39), and the analysis indicated that the total number of ORFs shared by the CyHVs and AngHV1 may be as many as 55 (counting duplicates) if tentative assignments are included (Table 2). The degree of genome colinearity among the CyHVs and AngHV1 is much less than that among the CyHVs, existing in the form of relatively short blocks that are rearranged in order and orientation with respect to each other (Table 2). These findings reinforce the observation that the CyHVs are more closely related to each other than they are to the other member of the genus *Cyprinivirus* (37). It is notable that the CyHVs may be perceived as exhibiting a pattern of recent coevolution (though perhaps not cospeciation) with their hosts, whereas this is not the case when the CyHVs are considered in relation to AngHV1 (37). Thus, in the context of the family *Alloherpesviridae*, the CyHVs and AngHV1 are too closely related to each other in comparison with their hosts, and it may be that one of these two lineages has evolved via an interspecies transfer event.

Fragmented ORFs. CyHV3 isolates have been reported to contain several fragmented ORFs (38). Since some of the mutations were shown to be common to independent isolates, it was concluded that they had occurred *in vivo*, and this led to the speculation that functional loss may be associated with increased pathogenicity in aquaculture settings. In reference strain KHV-U, five ORFs (ORF26, ORF27, ORF30, ORF40, and ORF94) are frameshifted and probably nonfunctional. Two ORFs (ORF25B and ORF151A) appear to be fragmented in CyHV2, but it is not known whether the presumed mutations occurred *in vivo* or *in vitro*. It is notable that CyHV2 ORF25B is a member of the ORF25 family, as are CyHV3 ORF26 and ORF27. No fragmented ORFs were identified in CyHV1.

Gene families. Five gene families have been identified in CyHV3 (38) (Fig. 3), and five and six gene families are present in CyHV1 and CyHV2, respectively (Fig. 1 and 2 and Table 2). These are named after a single member or common feature of the family. They are thought to have arisen via gene duplication events, although it is conceivable in some cases that a cellular gene may have been captured separately on more than one occasion, giving the appearance of an intrinsic gene family. All three CyHVs have multiple members of the ORF2, TNFR, ORF25, and RING families. ORF22 is present in all CyHVs and as a family of two in CyHV2 and CyHV3. ORF153 is present in CyHV2 and CyHV3 and as a family of three only in CyHV2. The JUNB family is confined to CyHV1 (see below).

ORF2 family members are numerous and highly divergent, with only a short region (approximately 17 residues containing the conserved motif QWXXG) shared by all except CyHV1 ORF129, which is truncated relative to CyHV2 and CyHV3 ORF129 but related to them in other regions. The ORF153 family specifies predicted type 3 membrane proteins. The ORF25 family encodes highly divergent predicted type 1 membrane glycoproteins, some of which exhibit weak similarities to immunoglobulin domains. All nonfragmented members of this family in CyHV3 (ORF25, ORF65, ORF148, and ORF149) are present in virions, presumably in the envelope (57). The evolutionarily dynamic nature of the ORF25 family is illustrated by the presence of remnants of additional members downstream from ORF25B and ORF25C in CyHV2 (Fig. 2). These vestiges lack ATG initiation codons and appear to comprise the 3′ ends of ORFs encoding the transmembrane domains of ORF25 family proteins. Thus, it seems that ORF25 family members may have come and gone frequently during CyHV evolution, in a complex fashion similar to that observed in some gene families in members of the family *Herpesviridae* (64). The TNFR and RING families are named for the similarity of their encoded proteins to tumor necrosis factor receptor or for the presence of a RING-type, C_3HC_4 zinc finger domain, respectively. These families also have members in AngHV1 (39). The JUNB family is named for the similarity of its predicted encoded proteins to the DNA-binding and dimerization domains of the transcription factor and oncoprotein JunB, particularly of fish, and is confined to CyHV1 in the form of three adjacent ORFs, namely, ORF151B, ORF151C, and ORF151D (Fig. 1). The situation is complex, in that ORF151D encodes both domains whereas similarity in ORF151C is restricted to the DNA-binding domain and similarity in ORF151B is restricted to the dimerization domain (Fig. 5c). Also, a sequence similar to the junction between these two domains is located close upstream from ORF151B in an alternative reading frame and may represent a nonfunctional relic of a fragmented gene. This is the first report of the presence of JUNB-related sequences in a herpesvirus, and investigations into the expression of these sequences and, more importantly, their involvement in tumor production should be included in future studies.

Concluding remarks. Our comparative analysis constitutes the first examination of a set of substantially related alloherpesvirus genomes and will help to build knowledge of the genomics of the family *Alloherpesviridae* to the level that exists for the much more extensively studied family *Herpesviridae*. As is proving to be the case for CyHV3, the CyHV1 and CyHV2 genome sequences will underpin future studies on the pathogenesis of carp herpesviruses, including those aimed at epidemiological, diagnostic, and therapeutic innovations.

ACKNOWLEDGMENTS

This work was supported by the Aquaculture and Fisheries Program, University of California, Davis, CA, and the United Kingdom Medical Research Council.

We are grateful to Wai Kwong Lee and Stewart Laing (BHF Glasgow Cardiovascular Research Centre, University of Glasgow) for providing capillary DNA sequencing services.

REFERENCES

1. **Pellett PE, Davison AJ, Eberle R, Ehlers B, Hayward GS, Lacoste V, Minson AC, Nicholas J, Roizman B, Studdert MJ, Wang F.** 2011. *Herpesvirales*, p 99–107. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy, 9th report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, London, United Kingdom.
2. **Gessner C.** 1558. Historiae Animalium, vol IV. C. Froschauer, Zurich, Switzerland.
3. **Hofer B.** 1904. Handbuch der Fischkrankheiten. Verlag der Allgemeinen Fischerei-Zeitung, B. Heller, Munich, Germany.
4. **Wolf K.** 1988. Fish viruses and fish viral diseases, p 253–263. Cornell University Press, Ithaca, NY.
5. **Lowenthal W.** 1907. Einschlussartige Zell- und Kernveränderungen in der Karpfenpocke. Z. Krebsforsch. **5**:197–204.
6. **Schubert G.** 1964. Elektronenmikroskopische Untersuchungen zur Pockenkrankheit des Karpfens. Z. Naturforsch. B **19**:675–682.
7. **Sano T, Fukuda H, Furukawa M, Hosoya H, Moriya Y.** 1985. A herpesvirus isolated from carp papilloma in Japan, p 307–311. *In* Ellis AE (ed), Fish and shellfish pathology. Academic Press, London, United Kingdom.
8. **Dixon PF.** 2008. Virus diseases of cyprinids, p 87–184. *In* Eiras JC, Segner H, Wahli T, Kapoor BG (ed), Fish diseases, vol 1. Science Publishers, Enfield, NH.
9. **Sano T, Fukuda H, Furukawa M.** 1985. *Herpesvirus cyprini*: biological and oncogenic properties. Fish Pathol. **20**:381–388.
10. **Sano T, Morita N, Shima N, Akimoto M.** 1990. A preliminary report on pathogenicity and oncogenicity of cyprinid herpesvirus. Bull. Eur. Ass. Fish Pathol. **10**:11–13.
11. **Sano T, Morita N, Shima N, Akimoto M.** 1991. *Herpesvirus cyprini*: lethality and oncogenicity. J. Fish Dis. **14**:533–543.
12. **Lu H, Zhu G, Fan L, Zhang L.** 2009. Etiology and pathology of epidermal papillomas in allogynogenetic crucian carp *Carassius auratus gibelio* (♀) × *Cyprinus carpio var. singuonensis* (♂). Dis. Aquat. Org. **83**:77–84.
13. **Sano N, Moriwake M, Hondo R, Sano T.** 1993. *Herpesvirus cyprini*: a search for viral genome in infected fish by *in situ* hybridization. J. Fish Dis. **16**:495–499.
14. **Goodwin AE, Sadler J, Merry GE, Marceaux EN.** 2009. Herpesviral haematopoietic necrosis virus (CyHV-2) infection: case studies from commercial goldfish farms. J. Fish Dis. **32**:271–278.
15. **Jung SJ, Miyazaki T.** 1995. Herpesviral haematopoietic necrosis of goldfish, *Carassius auratus* (L.). J. Fish Dis. **18**:211–220.
16. **Goodwin AE, Merry GE, Sadler J.** 2006. Detection of the herpesviral hematopoietic necrosis disease agent (*Cyprinid herpesvirus* 2) in moribund and healthy goldfish: validation of a quantitative PCR diagnostic method. Dis. Aquat. Org. **69**:137–143.
17. **Hedrick RP, Gilad O, Yun S, Spangenberg JV, Marty GD, Nordhausen RW, Kebus MJ, Bercovier H, Eldar A.** 2000. A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of a common carp. J. Aquat. Anim. Health **12**:44–57.
18. **Ronen A, Perelberg A, Abramowitz J, Hutoran M, Tinman S, Bejerano I, Steinitz M, Kotler M.** 2003. Efficient vaccine against the virus causing a lethal disease in cultured *Cyprinus carpio*. Vaccine **21**:4677–4684.
19. **Haenen OLM, Way K, Bergmann SM, Ariel E.** 2004. The emergence of koi herpesvirus and its significance to European aquaculture. Bull. Eur. Ass. Fish Pathol. **24**:293–307.
20. **Hedrick RP, Marty GD, Nordhausen RW, Kebus M, Bercovier H, Eldar A.** 1999. A herpesvirus associated with mass mortality of juvenile and adult koi *Cyprinus carpio*. Fish Health Newsl. Fish Health Sect. Am. Fish. Soc. **27**:7.
21. **Ariav R, Tinman S, Bejerano I.** 1999. First report of newly emerging viral disease of *Cyprinus carpio* species in Israel, p 151. Abstr. EAFP 9th Int. Conf.
22. **Bretzinger A, Fischer-Scherl T, Oumouma M, Hoffmann R, Truyen U.** 1999. Mass mortalities in koi, *Cyprinus carpio*, associated with gill and skin disease. Bull. Eur. Ass. Fish Pathol. **19**:182–185.
23. **Walster CI.** 1999. Clinical observation of severe mortalities in koi carp, *Cyprinus carpio*, with gill disease. Fish Vet. J. **3**:54–58.
24. **Perelberg A, Smirnov M, Huturan M, Diamant A, Bejerano Y, Kotler M.** 2003. Epidemiological description of a new viral disease afflicting cultured *Cyprinus carpio* in Israel. Isr. J. Aquaculture-Bamidgeh **55**:5–12.
25. **Sano M, Ito T, Kurita J, Yanai T, Watanabe N, Miwa S, Iida T.** 2004. First detection of koi herpesvirus in cultured common carp *Cyprinus carpio* in Japan. Fish Pathol. **39**:165–167.
26. **Hara H, Aikawa H, Usui K, Nakanishi T.** 2006. Outbreaks of koi herpesvirus disease in rivers of Kanagawa Prefecture. Fish Pathol. **41**:81–83.
27. **Costes B, Stalin Raj V, Michel B, Fournier G, Thirion M, Gillet L, Mast J, Lieffrig F, Bremont M, Vanderplasschen A.** 2009. The major portal of entry of koi herpesvirus in *Cyprinus carpio* is the skin. J. Virol. **83**:2819–2830.
28. **Walster CI.** 2003. A note on koi herpesvirus (KHV): current situation and issues arising. Fish Vet. J. **7**:77–81.
29. **Gilad O, Yun S, Adkison MA, Way K, Willits NH, Bercovier H, Hedrick RP.** 2003. Molecular comparison of isolates of an emerging fish pathogen, koi herpesvirus, and the effect of water temperature on mortality of experimentally infected koi. J. Gen. Virol. **84**:2661–2668.
30. **Gilad O, Yun S, Zagmutt-Vergara FJ, Leutenegger CM, Bercovier H, Hedrick RP.** 2004. Concentrations of a koi herpesvirus (KHV) in tissues of experimentally infected *Cyprinus carpio koi* assessed by real-time TaqMan PCR. Dis. Aquat. Org. **60**:179–187.
31. **Adkison MA, Gilad O, Hedrick RP.** 2005. An enzyme linked immunosorbent assay (ELISA) for detection of antibodies to the koi herpesvirus (KHV) in the serum of koi *Cyprinus carpio*. Fish Pathol. **40**:53–62.
32. **St-Hilaire S, Beevers N, Way K, Le Deuff RM, Martin P, Joiner C.** 2005. Reactivation of koi herpesvirus infections in common carp *Cyprinus carpio*. Dis. Aquat. Org. **67**:15–23.
33. **Dishon A, Davidovich M, Ilouze M, Kotler M.** 2007. Persistence of cyprinid herpesvirus 3 in infected cultured carp cells. J. Virol. **81**:4828–4836.
34. **St-Hilaire S, Beevers N, Joiner C, Hedrick RP, Way K.** 2009. Antibody response of two populations of common carp, *Cyprinus carpio* L., exposed to koi herpesvirus. J. Fish Dis. **32**:311–320.
35. **Eide KE, Miller-Morgan T, Heidel JR, Kent ML, Bildfell RJ, Lapatra S, Watson G, Jin L.** 2011. Investigation of koi herpesvirus latency in koi. J. Virol. **85**:4954–4962.
36. **Waltzek TB, Kelley GO, Stone DM, Way K, Hanson L, Fukuda H, Hirono I, Aoki T, Davison AJ, Hedrick RP.** 2005. Koi herpesvirus represents a third cyprinid herpesvirus (CyHV-3) in the family *Herpesviridae*. J. Gen. Virol. **86**:1659–1667.
37. **Waltzek TB, Kelley GO, Alfaro ME, Kurobe T, Davison AJ, Hedrick RP.** 2009. Phylogenetic relationships in the family *Alloherpesviridae*. Dis. Aquat. Org. **84**:179–194.
38. **Aoki T, Hirono I, Kurokawa K, Fukuda H, Nahary R, Eldar A, Davison AJ, Waltzek TB, Bercovier H, Hedrick RP.** 2007. Genome sequences of three koi herpesvirus isolates representing the expanding distribution of an emerging disease threatening koi and common carp worldwide. J. Virol. **81**:5058–5065.
39. **van Beurden SJ, Bossers A, Voorbergen-Laarman MHA, Haenen OLM, Peters S, Abma-Henkens MHC, Peeters BPH, Rottier PJM, Engelsma MY.** 2010. Complete genome sequence and taxonomic position of anguillid herpesvirus 1. J. Gen. Virol. **91**:880–887.
40. **van Beurden SJ, Gatherer D, Kerr K, Herzyk P, Galbraith G, Peeters BPH, Rottier PJM, Engelsma MY, Davison AJ.** 2012. Anguillid herpesvirus 1 transcriptome. J. Virol. **86**:10150–10161.
41. **Doszpoly A, Kovács ER, Bovo G, LaPatra SE, Harrach B, Benkő M.** 2008. Molecular confirmation of a new herpesvirus from catfish (*Ameiurus melas*) by testing the performance of a novel PCR method, designed to target the DNA polymerase gene of alloherpesviruses. Arch. Virol. **153**:2123–2127.
42. **Sano N, Hondo R, Fukuda H, Sano T.** 1991. *Herpesvirus cyprini*: restriction endonuclease cleavage profiles of the viral DNA. Fish Pathol. **26**:207–208.
43. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
44. **Cunningham C, Gatherer D, Hilfrich B, Baluchova K, Dargan DJ, Thomson M, Griffiths PD, Wilkinson GWG, Schulz TF, Davison AJ.** 2010. Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens. J. Gen. Virol. **91**:605–615.
45. **Li H, Ruan J, Durbin R.** 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. **18**:1851–1858.
46. **Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D.** 2010. Tablet—next generation sequence assembly visualization. Bioinformatics **26**:401–402.
47. **Davison AJ, Dolan A, Akter P, Addison C, Dargan DJ, Alcendor DJ, McGeoch DJ, Hayward GS.** 2003. The human cytomegalovirus genome

revisited: comparison with the chimpanzee cytomegalovirus genome. J. Gen. Virol. **84**:17–28.
48. **Honess RW, Gompels UA, Barrell BG, Craxton M, Cameron KR, Staden R, Chang Y-N, Hayward GS.** 1989. Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes. J. Gen. Virol. **70**:837–855.
49. **Arbuckle JH, Medveczky MM, Luka J, Hadley SH, Luegmayr A, Ablashi D, Lund TC, Tolar J, De Meirleir K, Montoya JG, Komaroff AL, Ambros PF, Medveczky PG.** 2010. The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. Proc. Natl. Acad. Sci. U. S. A. **107**:5563–5568.
50. **Kaufer BB, Jarosinski KW, Osterrieder N.** 2011. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. J. Exp. Med. **208**:605–615.
51. **Stern-Ginossar N, Weisburd B, Michalski A, Le VTK, Hein MY, Huang S-X, Ma M, Shen B, Qian S-B, Hengel H, Mann M, Ingolia NT, Weissman JS.** 2012. Decoding human cytomegalovirus. Science **338**:1088–1093.
52. **Zhang G, Raghavan B, Kotur M, Cheatham J, Sedmak D, Cook C, Waldman J, Trgovcich J.** 2007. Antisense transcription in the human cytomegalovirus transcriptome. J. Virol. **81**:11267–11281.
53. **Gatherer D, Seirafian S, Cunningham C, Holton M, Dargan DJ, Baluchova K, Hector RD, Galbraith J, Herzyk P, Wilkinson GWG, Davison AJ.** 2011. High-resolution human cytomegalovirus transcriptome. Proc. Natl. Acad. Sci. U. S. A. **108**:19755–19760.
54. **Davison AJ.** 1992. Channel catfish virus: a new type of herpesvirus. Virology **186**:9–14.
55. **Davison AJ, Cunningham C, Sauerbier W, McKinnell RG.** 2006. Genome sequences of two frog herpesviruses. J. Gen. Virol. **87**:3509–3514.
56. **Abbotts AP, Preston VG, Hughes M, Patel AH, Stow ND.** 2000. Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. J. Gen. Virol. **81**:2999–3009.
57. **Michel B, Leroy B, Stalin Raj V, Lieffrig F, Mast J, Wattiez R, Vanderplasschen AF, Costes B.** 2010. The genome of cyprinid herpesvirus 3 encodes 40 proteins incorporated in mature virions. J. Gen. Virol. **91**:452–462.
58. **Davison AJ, Davison MD.** 1995. Identification of structural proteins of channel catfish virus by mass spectrometry. Virology **206**:1035–1043.
59. **van Beurden SJ, Leroy B, Wattiez R, Haenen OLM, Boeren S, Vervoort JJM, Peeters BPH, Rottier PJM, Engelsma MY, Vanderplasschen AF.** 2011. Identification and localization of the structural proteins of anguillid herpesvirus 1. Vet. Res. **42**:105.
60. **Booy FP, Trus BL, Davison AJ, Steven AC.** 1996. The capsid architecture of channel catfish virus, an evolutionarily distant herpesvirus, is largely conserved in the absence of discernible sequence homology with herpes simplex virus. Virology **215**:134–141.
61. **Davison AJ, Trus BL, Cheng N, Steven AC, Watson MS, Cunningham C, Le Deuff RM, Renault T.** 2005. A novel class of herpesvirus with bivalve hosts. J. Gen. Virol. **86**:41–53.
62. **Rixon FJ.** 1993. Structure and assembly of herpesviruses. Semin. Virol. **4**:135–144.
63. **van Beurden SJ, Forlenza M, Westphal AH, Wiegertjes GF, Haenen OLM, Engelsma MY.** 2011. The alloherpesviral counterparts of interleukin 10 in European eel and common carp. Fish Shellfish Immunol. **31**:1211–1217.
64. **Alcendor DJ, Zong J, Dolan A, Gatherer D, Davison AJ, Hayward GS.** 2009. Patterns of divergence in the vCXCL and vGPCR gene clusters in primate cytomegalovirus genomes. Virology **395**:21–32.