Arch Virol (2012) 157:1559–1564
DOI 10.1007/s00705-012-1316-9

BRIEF REPORT

Sequencing and analysis of the complete genome of Rana grylio virus (RGV)

Xiao-Ying Lei · Tong Ou · Ruo-Lin Zhu · Qi-Ya Zhang

Received: 13 February 2012 / Accepted: 19 March 2012 / Published online: 28 April 2012

Abstract Infection with Rana grylio virus (RGV), an iridovirus isolated in China in 1995, resulted in a high mortality rate in frogs. The complete genome sequence of RGV was determined and analyzed. The genomic DNA was 105,791 bp long, with 106 open reading frames (ORFs). Dot plot analysis showed that the gene order of RGV shared colinearity with three completely sequenced ranaviruses. A phylogenetic tree was constructed based on concatenated sequences of iridovirus 26 core-gene-encoded proteins, and the result showed high bootstrap support for RGV being a member of the genus *Ranavirus* and that iridoviruses of other genera also clustered closely. A microRNA (miRNA) prediction revealed that RGV could encode 18 mature miRNAs, many of which were located near genes associated with virus replication. Thirty-three repeated sequences were found in the RGV genome. These results provide insight into the genetic nature of RGV and are useful for laboratory diagnosis for vertebrate iridoviruses.

Iridoviruses are nucleo-cytoplasmic large DNA viruses (NCLDVs) and contain circularly permutated and terminally redundant double-stranded genomes [1–3]. The family *Iridoviridae* currently contains five genera, including the genera *Iridovirus*, *Chloriridovirus*, *Lymphocystivirus*, *Megalocytivirus* and *Ranavirus* [4]. So far, the entire genomes of 18 iridoviruses have been sequenced completely (Table S1). Three of these are invertebrate-infecting iridoviruses, and the other 15 are associated with aquatic vertebrates. These include lymphocystis disease virus 1 (LCDV-1) and lymphocystis disease virus-China (LCDV-C), belonging to the genus *Lymphocystivirus*, and infectious spleen and kidney necrosis virus (ISKNV), rock bream iridovirus (RBIV), turbot reddish body iridovirus (TRBIV), and other viruses, belonging to the genus *Megalocytivirus* that can infect fish species. Iridoviruses of the genus *Ranavirus*, including frog virus 3 (FV3), soft-shelled turtle iridovirus (STIV), epizootic hematopoietic necrosis virus (EHNV), and other viruses can infect amphibians, fish and reptiles [5–12].

Members of the genus *Ranavirus* have been recognized as major pathogens of economically and ecologically important cold-blooded vertebrates and have become serious problems in modern aquaculture, fish farming, and wildlife conservation because of their epidemic morbidity and ability to cause mortality [13]. Rana grylio virus (RGV) is a pathogenic agent that results in greater than 90 % mortality in cultured pig frog (*Rana grylio*), which was the first iridovirus isolated in China in 1995 [14, 15]. RGV has been recognized as a member of the family *Iridoviridae* and is closely related to FV3, based on previous studies on morphogenesis, cellular interaction, antigenicity, restriction fragment length polymorphism (RFLP) and major capsid protein (MCP) sequence similarity [15–18]. To date, some genes of RGV have been identified and characterized, such as 3β-hydroxysteroid dehydrogenase (3β-HSD), deoxyuridine triphosphatase (dUTPase), an envelope protein gene (53R), thymidine kinase (TK) and a gene belonging to the "essential for respiration and

Electronic supplementary material The online version of this article (doi:10.1007/s00705-012-1316-9) contains supplementary material, which is available to authorized users.

X.-Y. Lei · T. Ou · R.-L. Zhu · Q.-Y. Zhang (✉)
State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Graduate School of the Chinese Academy of Sciences, Wuhan 430072, China
e-mail: zhangqy@ihb.ac.cn

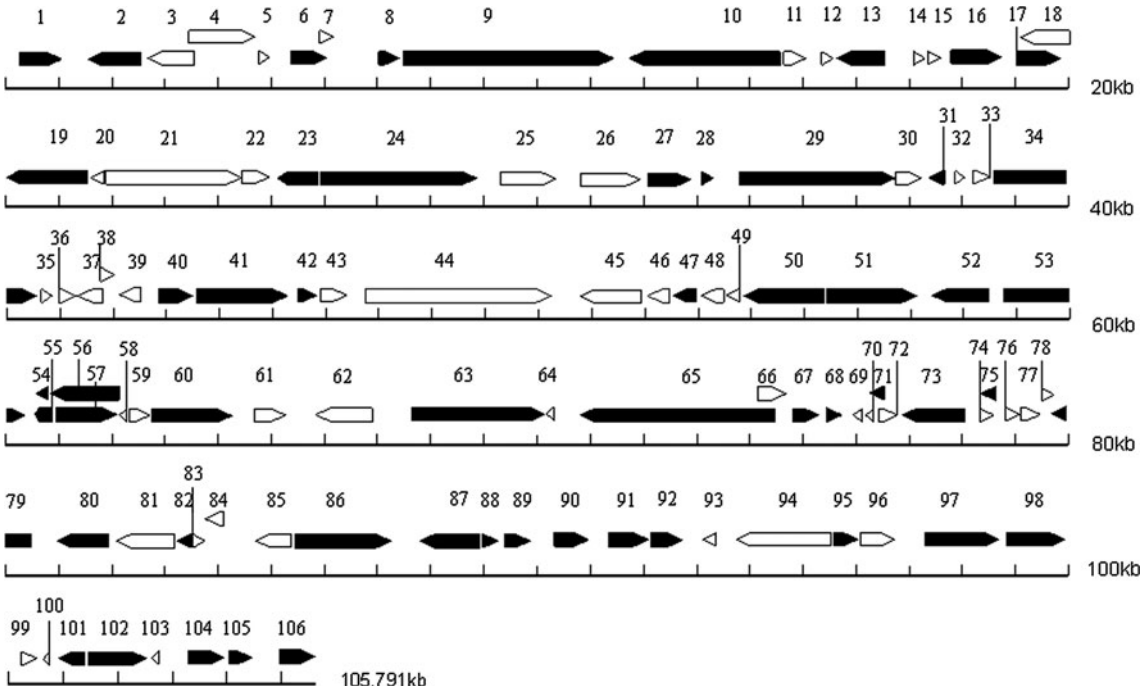

Fig. 1 Organization of the RGV genome. Predicted ORFs are numbered from left to right and represented by arrows indicating their approximate size, location and orientation based on the positions of methionine start and stop codons. Black arrows represent the ORFs with predicted functions, and the white ones represent those with unknown function

viability" family (ERV1) [19–24]. We found that RGV could induce apoptosis mediated by mitochondria [25]. RGV has the potential to be used as a viral vector for expression of foreign genes in fish cells [26].

So far, the complete sequence of RGV has not been available, and the relationship of RGV to other iridoviruses has not been well understood. To aid in the understanding of the molecular mechanisms of amphibian iridovirus pathogenesis and its evolutionary status, the complete sequence of RGV genomic DNA was determined and analyzed.

RGV was isolated from tissues of diseased young pig frogs with lethal syndrome in Hubei Province, China, in 1995, and *Epithelioma papulosum cyprinid* (EPC) cells were used for virus propagation, grown in TC 199 medium supplemented with 10 % fetal bovine serum (FBS) at 25 °C [14]. Virus propagation and purification were performed as we described previously [17]. RGV genomic DNA was prepared from purified virus particles. The purified virus was incubated with TES (10 mM Tris, 1 mM EDTA, 1 % SDS, pH 8.0, 100 μg/ml proteinase K) at 37 °C for 2 h. Then, the lysate was subjected to phenol-chloroform extraction and ethanol precipitation as described in our previous work [27].

Primers were designed based on complete DNA sequence alignments of FV3, tiger frog virus (TFV) and STIV, which are shown in Table S2. The amplified PCR products were about 1500 bp in length, and the genes studied previously were not amplified again. All of the fragments were purified using a TIANgel Mini Purification Kit (Tiangen Biotech), cloned into vector pMD18-T (TaKaRa), and sequenced in both directions in an ABI PRISM 3700 automated DNA sequencer using M13 primers. Genomic DNA assembly, structure analysis, and amino acid sequence analysis were carried out using the DNASTAR software package (Lasergene, Madison, WI, USA). The ORFs were predicted using Gene Finding in the virus genome program at the website http://www.softberry.com and NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Comparisons of homologous protein-encoding regions of RGV with those of other viruses were conducted using the BLASTP program at NCBI website (http://www.ncbi.nlm.nih.gov). Transmembrane domains (TMs) were predicted using TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0) and the DAS-TM filter server (http://mendel.imp.univie.ac.at/sat/DAS) [28]. Repetitive DNA sequences were detected using REPuter and the Tandem Repeats Finder [29, 30]. The complete genome was scanned for miRNA coding regions using VMir [31], and possible miRNA coding sequences were further analyzed using MiPred [32] and RNAfold server (http://rna.tbi.univie.ac.at). DNA dot matrix plot analysis was performed using DNAMAN version 6 (Lynnon Corp.).

Sequences of 26 core-gene-encoded proteins from 17 other completely sequenced iridoviruses were concatenated as a continuous amino acid sequence with the same order

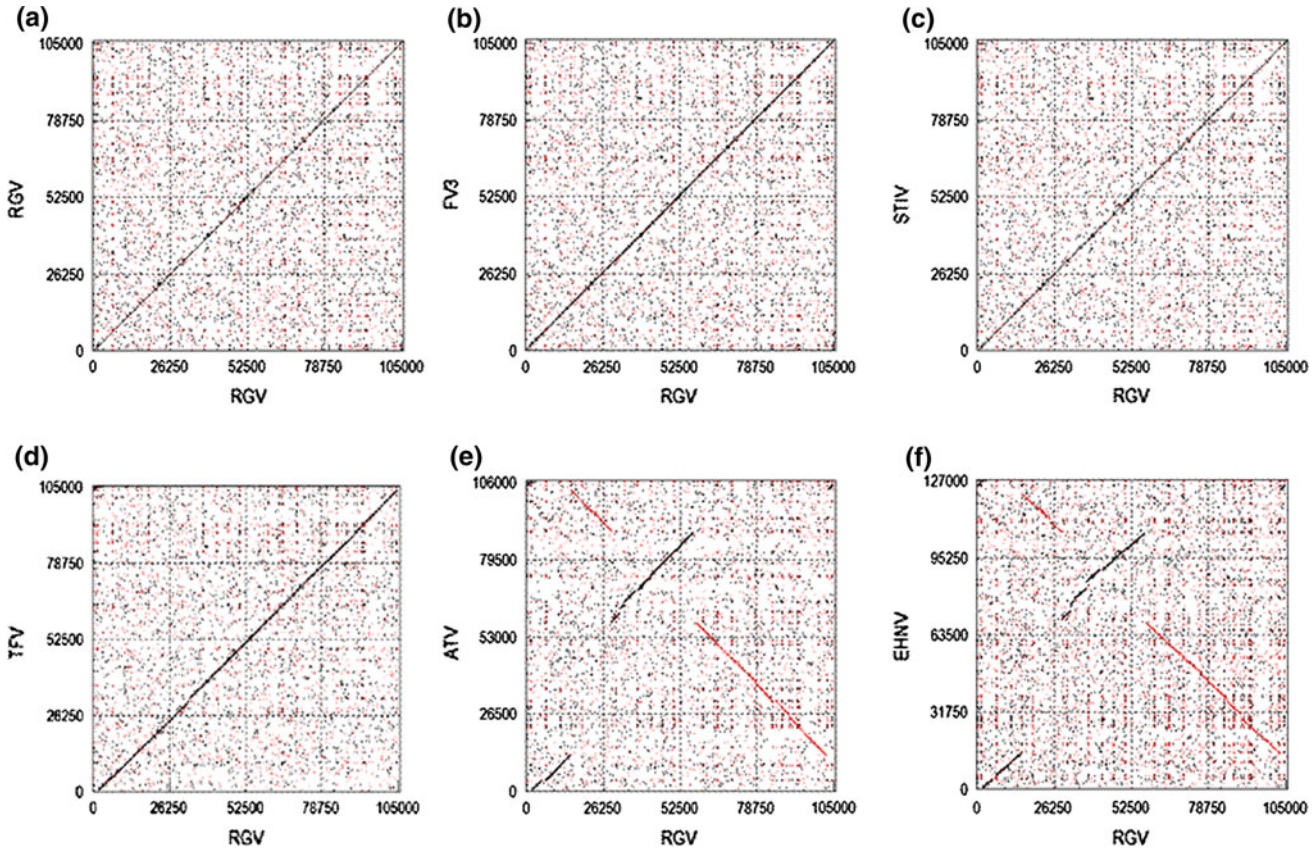

Fig. 2 DNA dot plot analysis of the RGV genome in comparison with itself and other ranaviruses. The horizontal axis represents the RGV genome. The vertical axes represent (**a**) the RGV genome, (**b**) the FV3 genome, (**c**) the STIV genome, (**d**) the TFV genome, (**e**) the ATV genome and (**f**) the EHNV genome. The complete genomic sequences were aligned using DNAMAN version 6, and both strands of the genome DNA sequences were aligned for the dot matrix plot. Solid lines show the high level of sequence similarity. Black dots show the results of comparisons of RGV with the positive-sense strand of the virus indicated in the vertical axis, and red dots show those with the corresponding plus-sense strand. Solid lines show the high level of sequence similarity (color figure online)

as RGV. LYCIV was not included, as its complete sequence has not been annotated. The phylogenetic tree was constructed using MrBayes 3.2 following multi-sequence alignment using Clustal X 1.83 and format conversion using Mega 4 [33]. A mixed amino acid model analysis was applied with 100,000 generations and a sampling frequency of 100.

ORFs were identified by the following criteria [9, 12, 34]: (1) they were at least 120 bp long, (2) they could be detected by the two annotation methods, and (3) they were not located within larger ORFs or have homologs to at least two known virus ORFs.

The genome size, G+C content, and potential ORFs of RGV were compared to those of other iridoviruses, and these data are shown in Table S1. The RGV genome (GenBank accession no. JQ654586) contained a double-stranded DNA consisting of 105,791 bp, with a G+C content of 55 %. One hundred and six putative ORFs could be identified in the RGV genome by computer-assisted analysis, the length of which ranged from 126 to 3885 bp. ORFs that completely overlapped with others and had no homology to those of other viruses were not annotated in this work. The locations, sizes, predicted functions and motifs of each putative ORF are listed in Table S3, together with homologous proteins in other iridoviruses (FV3, STIV, LCDV-C and ISKNV). Eighty-nine percent of the ORFs have orthologous genes in FV3, 84 % in STIV, and only 25 % in ISKNV.

A diagrammatic representation of the RGV genome is shown in Fig. 1. Fifty-five ORFs were predicted to have functions or functional domains (black arrows) related to DNA replication, transcription, nucleotide metabolism, protein synthesis and modification, viral structure, or host-virus interactions (Table S3). Additionally, of the 22 partially overlapping ORFs, one (72R) had a corresponding ortholog only in the ATV genome, two (38R, 70L) had orthologs only in the STIV genome, others had orthologs in at least two viruses, and 57R was completely within 56L.

DNA dot plot analysis comparing the RGV genomic DNA with itself and other iridovirus genomes revealed that RGV shares high colinearity and sequence similarity with two frog ranaviruses (FV3 and TFV) and one reptile

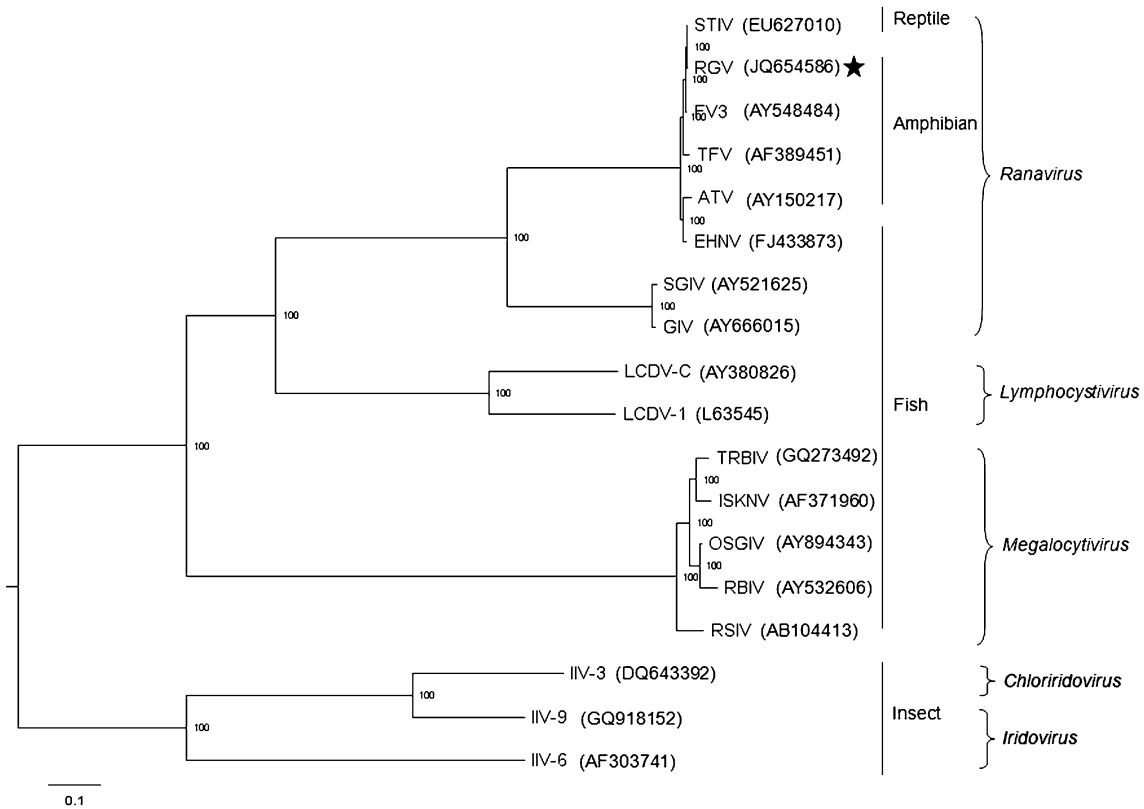

Fig. 3 Concatenated phylogenetic tree of iridoviruses. Twenty-six core-gene-encoded proteins from 17 other completely sequenced iridoviruses were rearranged as continuous amino acid sequences with the same order as RGV. Multi-sequence alignment was carried out using Clustal X 1.83, and the format of the output file was converted to nex format using Mega 4. The phylogenetic tree was constructed using MrBayes 3.2. The program uses a mixed amino acid model with 100,000 generations and a sampling frequency of 100. The host of each virus and virus genus are listed at the right

ranavirus (STIV) and that ATV and EHNV share partial colinearity with RGV (Fig. 2). GIV, SGIV, LCDV-C, ISKNV and CIV have little colinearity or sequence similarity to RGV (data not shown).

Iridoviruses are believed to contain 26 core genes, which have been identified in 12 iridovirus genomes by Eaton et al. [34]. All of the core genes were found in the RGV genome (the corresponding ORFs are marked with asterisks in Table S3), and the homologous ORFs of other sequenced iridoviruses are listed in Table S4. A phylogenetic tree was constructed based on the Bayesian inference algorithm using the concatenated sequences of 26 core-gene-encoded proteins, and the hosts and genera of the viruses are indicated (Fig. 3). The tree shows that the iridoviruses infecting vertebrates are clearly separated from those infecting invertebrates. As was showed by dot plot analysis, RGV is most closely related to STIV and clusters with ranaviruses with high support, confirming that RGV is a member of the genus *Ranavirus*. It was also observed that ranaviruses of other subgroups clustered together, i.e., FV3-like viruses (RGV, STIV, FV3, TFV), ATV-like viruses (ATV and EHNV), and GIV-like viruses (SGIV and GIV), which is consistent with previous studies [2, 11, 35].

Ranaviruses have a wide range of hosts, including fish, amphibians and reptiles, and they can be translocated across large distances with ecological and economic consequences [36]. There is evidence that the ancestral ranavirus was a fish virus and that several recent host shifts have taken place, with subsequent speciation of viruses in their new hosts [11]. Indeed, there is also evidence that ranaviruses switched hosts from tiger salamanders to freshwater sport fish during bait trade [37]. Based on this information and the phylogenetic relationships shown in Fig. 3, we speculate that RGV was imported to China with the commercial activities of pig frogs and that host switching from amphibians to reptiles has likely occurred.

Thirteen precursor sequences encoding 18 mature miRNA were identified (Table S5). All but one of these miRNAs were located within or near ORFs with known functions or predicted motifs associated with the genes that play important roles in virus DNA replication, transcription and assembly, such as 63R (DNA polymerase), 24R (D5 family NTPase), 9R (DNA-dependent RNA polymerase II

largest subunit), 65L (DNA-dependent RNA polymerase II second-largest subunit), 98R (ICP46), and 53R (lipid membrane protein). Indeed, most of the miRNAs were identified in the opposite orientation to their associated ORFs. Recently, a number of miRNAs have been identified for DNA viruses such as herpesviruses, polyomaviruses, ascoviruses and adenoviruses, which have evolved to exploit RNA silencing for regulating the expression of their own genes, host genes, or both [38–40]. For iridoviruses, a computational method has been used to predict IIV-9-encoded miRNAs, and deep-sequencing technology has been applied to identify miRNAs encoded by SGIV [35, 41]. Thus, it is possible that the potential miRNAs encoded by RGV regulate the expression of their own genes. However, how these miRNAs function needs further investigation.

Thirty-three repeated sequences were found in the RGV genome. The size of the repeated sequences ranged from 2 to 222 nt, and the copy numbers from 2 to 41 (Table S6). Nineteen repeat units were located within the putative ORFs: 2L, 39L, 40R, 43R, 47L, 50L, 81L, 85L, 93L, 94L and 96R contained one; 21R, 35R and 45L contained two or more; and the rest were located in the noncoding regions. Interestingly, a unique microsatellite consisting of 41 tandemly repeated CA dinucleotides (also known as simple sequence repeat, SSR) was located in the noncoding region between ORFs 79L and 80L, and 34 such repeats were also found in FV3 and STIV. Previous studies revealed that repeated sequences are common in iridoviruses, poxviruses, herpesvirus, baculoviruses, adenoviruses and retroviruses and may serve as regulatory elements involved in transcription, gene regulation and protein function [42, 43], suggesting that the repeat sequences in the RGV genome could play a regulatory role in viral replication. Furthermore, microsatellites such as CA and GCAGGA, as mutational hotspots, could produce genetic polymorphisms that are useful for studies of quantitative genetic variation and evolutionary adaptation.

In conclusion, RGV was isolated in Hubei Province, China, in 1995 [14]. The complete genome of RGV was sequenced and analyzed. The RGV genome contains all of the 26 core iridovirus genes. A Bayesian phylogenetic tree showed that RGV is closely related to STIV and FV3, which belong to the genus *Ranavirus* of the family *Iridoviridae*. Further analysis of the host relationships revealed that host shifting probably took place. Other characteristics of NCLDV, such as virus-encoded miRNAs and repeated sequences were also found in the RGV genome, and these might play an important regulatory role in gene transcription, protein expression or virus replication. The information we have provided may be useful for further studies on the mechanisms of amphibian iridovirus pathogenesis and the evolutionary relationships of the iridoviruses.

Acknowledgments This work was funded by National Major Basic Research Program (2010CB126303, 2009CB118704), Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-EW-Z-3), National Natural Science Foundation of China (31072239), and an FEBL research grant (Y15B121F01).

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Boyer M, Yutin N, Pagnier I, Barrassi L, Fournous G, Espinosa L, Robert C, Azza S, Sun S, Rossmann MG, Suzan-Monti M, La Scola B, Koonin EV, Raoult D (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc Natl Acad Sci USA 106(51): 21848–21853
2. Chinchar VG, Yu KH, Jancovich JK (2011) The molecular biology of frog virus 3 and other iridoviruses infecting cold-blooded vertebrates. Viruses 3(10):1959–1985
3. Chinchar VG, Hyatt A, Miyazaki T, Williams T (2009) Family *Iridoviridae*: poor viral relations no longer. Curr Top Microbiol Immunol 328:123–170
4. Jancovich JK, Chinchar VG, Hyatt A, Miyazaki T, Williams T, Zhang QY (2011) Family *Iridoviridae* virus taxonomy. In: 9th report of the ICTV. Elsevier, San Diego, pp 193–210
5. Tidona CA, Darai G (1997) The complete DNA sequence of lymphocystis disease virus. Virology 230:207–216
6. Zhang QY, Xiao F, Xie J, Li ZQ, Gui JF (2004) Complete genome sequence of lymphocystis disease virus isolated from China. J Virol 78:6982–6994
7. He JG, Deng M, Weng SP, Li Z, Zhou SY, Long QX, Wang XZ, Chan SM (2001) Complete genome analysis of the mandarin fish infectious spleen and kidney necrosis iridovirus. Virology 291(1):126–139
8. Do JW, Moon CH, Kim HJ, Ko MS, Kim SB, Son JH, Kim JS, An EJ, Kim MK, Lee SK, Han MS, Cha SJ, Park MS, Park MA, Kim YC, Kim JW, Park JW (2004) Complete genomic DNA sequence of rock bream iridovirus. Virology 325(2):351–363
9. Shi CY, Jia KT, Yang B, Huang J (2010) Complete genome sequence of a Megalocytivirus (family *Iridoviridae*) associated with turbot mortality in China. Virol J 7:159
10. Tan WG, Barkman TJ, Gregory Chinchar V, Essani K (2004) Comparative genomic analyses of frog virus 3, type species of the genus *Ranavirus* (family *Iridoviridae*). Virology 323:70–84
11. Huang YH, Huang XH, Liu H, Gong J, Ouyang ZL, Cui HC, Cao JH, Zhao YT, Wang XJ, Jiang YL, Qin QW (2009) Complete sequence determination of a novel reptile iridovirus isolated from soft-shelled turtle and evolutionary analysis of *Iridoviridae*. BMC Genomics 10:224
12. Jancovich JK, Bremont M, Touchman JW, Jacobs BL (2010) Evidence for multiple recent host species shifts among the Ranaviruses (Family *Iridoviridae*). J Virol 3:2636–2647
13. Tsai CT, Ting JW, Wu MH, Wu MF, Guo IC, Chang CY (2005) Complete genome sequence of the grouper iridovirus and comparison of genomic organization with those of other iridoviruses. J Virol 79(4):2010–2023
14. Zhang QY, Li ZQ, Jiang YL, Liang SC, Gui JF (1996) Preliminary studies on virus isolation and cell infection from disease frog *Rana grylio*. Acta Hydrobioliogica Sinica 20:390–392 (Chinese with English Abstract)
15. Zhang QY, Xiao F, Li ZQ, Gui JF, Mao JH, Chinchar GV (2001) Characterization of an iridovirus form the cultured pig frog (*Rana grylio*) with lethal syndrome. Dis Aquat Org 48:27–36

16. Zhang QY, Li ZQ, Gui JF (1999) Studies on morphogenesis and cellular interactions of *Rana grylio* virus in an infected fish cell line. Aquaculture 175:185–197
17. Zhang QY, Zhao Z, Xiao F, Li ZQ, Gui JF (2006) Molecular characterization of three *Rana grylio* virus (RGV) isolates and *Paralichthys olivaceus* lymphocystis disease virus (LCDV-C) in iridoviruses. Aquaculture 251:1–10
18. Huang XH, Huang YH, Yuan XP, Zhang QY (2006) Electron microscopic examination of the viromatrix of *Rana grylio* virus in a fish cell line. J Virol Methods 133:117–123
19. Sun W, Huang YH, Zhao Z, Gui JF, Zhang QY (2006) Characterization of the *Rana grylio* virus 3β-hydroxysteroid dehydrogenase and its novel role in suppressing virus-induced cytopathic effect. Biochem Biophys Res Commun 351:44–50
20. Zhao Z, Ke F, Gui JF, Zhang QY (2007) Characterization of an early gene encoding for dUTPase in *Rana grylio* virus. Virus Res 123:128–137
21. Zhao Z, Ke F, Huang YH, Zhao JG, Gui JF, Zhang QY (2008) Identification and characterization of a novel envelope protein in *Rana grylio* virus. J Gen Virol 89:1866–1872
22. Zhao Z, Ke F, Shi Y, Zhou GZ, Gui JF, Zhang QY (2009) *Rana grylio* virus thymidine kinase gene: an early gene of iridovirus encoding for a cytoplasmic protein. Virus Genes 38(2):345–352
23. Ke F, Zhao Z, Zhang QY (2009) Cloning, expression and subcellular distribution of a *Rana grylio* virus late gene encoding ERV1 homologue. Mol Biol Rep 36:1651–1659
24. Kim YS, Ke F, Lei XY, Zhu R, Zhang QY (2010) Viral envelope protein 53R gene highly specific silencing and iridovirus resistance in fish cells by a miRNA. PLoS ONE 5(4):e10308
25. Huang YH, Huang XH, Gui JF, Zhang QY (2007) Mitochondrion-mediated apoptosis induced by *Rana grylio* virus infection in fish cells. Apoptosis 12:1569–1577
26. He LB, Ke F, Zhang QY (2012) *Rana grylio* virus as a vector for foreign gene expression in fish cells. Virus Res 163(1):66–73
27. Du C, Zhang Q, Li C, Miao D, Gui J (2004) Induction of apoptosis in a carp leucocyte cell line infected with turbot (*Scophthalmus maximus* L.) rhabdovirus. Virus Res 101(2):119–126
28. Cserzö M, Eisenhaber F, Eisenhaber B, Simon I (2002) On filtering false positive transmembrane protein predictions. Protein Eng 15(9):745–752
29. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27(2):573–580
30. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R (2001) REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res 29(22):4633–4642
31. Grundhoff A, Sullivan CS, Ganem D (2006) A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. RNA 12:733–750
32. Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 35:W339–W344
33. Ronquist F, Huelsenbeck J, Teslenko M (2011) Draft MrBayes version 3.2 Manual: tutorials and model summaries. MrBayes: Bayesian Inference of Phylogeny. http://mrbayes.sourceforge.net/manual.php. Accessed November 15, 2011
34. Eaton HE, Metcalf J, Penny E, Tcherepanov V, Upton C, Brunetti CR (2007) Comparative genomic analysis of the family *Iridoviridae*: reannotating and defining the core set of iridovirus genes. Virol J 4:11
35. Wong CK, Young VL, Kleffmann T, Ward VK (2011) Genomic and proteomic analysis of invertebrate iridovirus type 9. J Virol 85(15):7900–7911
36. Miller D, Gray M, Storfer A (2011) Ecopathology of ranaviruses infecting amphibians. Viruses 3(11):2351–2373
37. Picco AM, Karam AP, Collins JP (2010) Pathogen host switching in commercial trade with management recommendations. EcoHealth 7(2):252–256
38. Pfeffer S, Zavolan M, Grässer FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded microRNAs. Science 304:734–736
39. Skalsky RL, Cullen BR (2010) Viruses, microRNAs, and host interactions. Annu Rev Microbiol 64:123–141
40. Jurak I, Kramer MF, Mellor JC, van Lint AL, Roth FP, Knipe DM, Coen DM (2010) Numerous conserved and divergent microRNAs expressed by herpes simplex viruses 1 and 2. J Virol 84(9):4659–4672
41. Yan Y, Cui H, Jiang S, Huang Y, Huang X, Wei S, Xu W, Qin Q (2011) Identification of a novel marine fish virus, Singapore grouper iridovirus-encoded microRNAs expressed in grouper cells by Solexa sequencing. PLoS ONE 6(4):e19148
42. Jancovich JK, Mao J, Chinchar VG, Wyatt C, Case ST, Kumar S, Valente G, Subramanian S, Davidson EW, Collins JP, Jacobs BL (2003) Genomic sequence of a ranavirus (family *Iridoviridae*) associated with salamander mortalities in North America. Virology 316(1):90–103
43. Kashi Y, King DG (2006) Simple sequence repeats as advantageous mutators in evolution. Trends Genet 22(5):253–259