

Manuscript Number: MEEGID-D-16-00062R3

Title: Whole genome sequencing and phylogenetic characterization of brown bullhead (*Ameiurus nebulosus*) origin ranavirus strains from independent disease outbreaks

Article Type: Short communication

Keywords: Iridoviridae; molecular characterization; Hungary

Corresponding Author: Dr. Krisztian Banyai, Ph.D.

Corresponding Author's Institution: MTA-ATK-AOTI

First Author: Enikő Fehér

Order of Authors: Enikő Fehér; Andor Doszpoly; Balázs Horváth; Szilvia Marton; Barbara Forró; Szilvia Farkas; Krisztian Banyai, Ph.D.; Tamás Juhász

Abstract: Ranaviruses are emerging pathogens associated with high mortality diseases in fish, amphibians and reptiles. Here we describe the whole genome sequence of two ranavirus isolates from brown bullhead (*Ameiurus nebulosus*) specimens collected in 2012 at two different locations in Hungary during independent mass mortality events. The two Hungarian isolates were highly similar to each other at the genome sequence level (99.9% nucleotide identity) and to a European sheatfish (*Silurus glanis*) origin ranavirus (ESV, 99.7%-99.9% nucleotide identity). The coding potential of the genomes of both Hungarian isolates, with 136 putative proteins, were shared with that of the ESV. The core genes commonly used in phylogenetic analysis of ranaviruses were not useful to differentiate the two brown bullhead ESV strains. However genome-wide distribution of point mutations and structural variations observed mainly in the non-coding regions of the genome suggested that the ranavirus disease outbreaks in Hungary were caused by different virus strains. At this moment, due to limited whole genome sequence data of ESV it is unclear whether these genomic changes are useful in molecular epidemiological monitoring of ranavirus disease outbreaks. Therefore, complete genome sequencing of further isolates will be needed to identify adequate genetic markers, if any, and demonstrate their utility in disease control and prevention.

Highlights

- The Hungarian ranavirus strains shared great genome sequence identity.
- Genome organization of the Hungarian ranavirus isolates was similar to that of the ESV.
- Non-coding regions of ranaviruses could be useful in molecular epidemiological studies.

**Whole genome sequencing and phylogenetic characterization of brown bullhead
(*Ameiurus nebulosus*) origin ranavirus strains from independent disease outbreaks**

Enikő Fehér^{a*}, Andor Doszpoly^{a*}, Balázs Horváth^b, Szilvia Marton^a, Barbara Forró^a, Szilvia L. Farkas^a, Krisztián Bányai^{a✉}, Tamás Juhász^c

^a Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, Hungária krt. 21, 1143 Budapest, Hungary

^b Biological Research Center, Hungarian Academy of Sciences, Temesvári krt. 62, Szeged, 6726, Hungary

^c Veterinary Diagnostic Directorate, National Food Chain Safety Office, Tábornok u. 2, 1143 Budapest, Hungary

* These authors have contributed equally to this work

✉ Corresponding author: Krisztián Bányai

bkrota@hotmail.com

Hungária krt 21., Budapest, Hungary

Abstract

Ranaviruses are emerging pathogens associated with high mortality diseases in fish, amphibians and reptiles. Here we describe the whole genome sequence of two ranavirus isolates from brown bullhead (*Ameiurus nebulosus*) specimens collected in 2012 at two different locations in Hungary during independent mass mortality events. The two Hungarian isolates were highly similar to each other at the genome sequence level (99.9% nucleotide identity) and to a European sheatfish (*Silurus glanis*) origin ranavirus (ESV, 99.7%-99.9% nucleotide identity). The coding potential of the genomes of both Hungarian isolates, with 136 putative proteins, were shared with that of the ESV. The core genes commonly used in phylogenetic analysis of ranaviruses were not useful to differentiate the two brown bullhead ESV strains. However genome-wide distribution of point mutations and structural variations observed mainly in the non-coding regions of the genome suggested that the ranavirus disease outbreaks in Hungary were caused by different virus strains. At this moment, due to limited whole genome sequence data of ESV it is unclear whether these genomic changes are useful in molecular epidemiological monitoring of ranavirus disease outbreaks. Therefore, complete genome sequencing of further isolates will be needed to identify adequate genetic markers, if any, and demonstrate their utility in disease control and prevention.

Keywords: *Iridoviridae*; molecular characterization; Hungary

Text

Ranaviruses are known to cause acute, systemic diseases and are often associated with mass mortality events in fish, amphibians and reptiles (Holopainen et al., 2009; Hyett et al., 2000; Stöhr et al., 2015). Ranaviruses (family *Iridoviridae*) are large, double stranded DNA viruses with approximate genome lengths of 100 to 140 kilobase pair and an estimated 72 to 162 open reading frames (ORFs). Although functions of the majority of these genes remain to be determined, the genes of major capsid proteins and those associated with replication, transcription and metabolism have been assigned (Holopainen et al., 2009; Jancovich et al., 2010; Stöhr et al., 2015). Conserved genomic regions have been used to develop molecular diagnostic tools and to delineate phylogeny and evolution of heterologous virus strains (Abrams et al., 2013; Holopainen et al., 2009; Jancovich et al., 2010; Stöhr et al., 2015).

Ranaviruses with available sequence information can be clustered into the amphibian-like ranaviruses (ALRV) and the grouper iridovirus-like ranaviruses; within the ALRV group three major lineages can be distinguished (Abrams et al., 2013; Holopainen et al., 2009; Jancovich et al., 2010; Stöhr et al., 2015).

Brown bullhead (*Ameiurus nebulosus*), a 20-30 cm long (max. length 55 cm) siluriform fish species of the Ictaluridae family, is indigenous and widely distributed in North America. This fish was introduced to Europe in 1880 and as a highly invasive fish colonized the Hungarian freshwaters at the beginning of the 20th century. Brown bullheads live in a variety of habitats, mostly lakes and ponds rich in aquatic vegetation (www.fishbase.org). Although brown bullheads are considered to pose some threat to indigenous European freshwater fish species, line-fishing fans appreciate their gastronomic value.

During the spring of 2012, diseased (n=7) and dead (n=38) brown bullhead specimens were collected at two different locations in Hungary, Esztergom and Mártély (Fig 1), where mass mortality events affected the respective local brown bullhead populations. The gross pathological examination that recorded hemorrhages in the skin, paddles, gills and internal organs (including liver and kidney) was consistent with typical ranavirus infection of brown bullhead (Juhász et al., 2013). In the present study one fish specimen from each affected population was further examined in order to obtain more information about the causative agents. Parasitic and bacterial infections were not diagnosed. EPC (Epithelioma Papulosum Cyprini) and BF-2 (Bluegill fibroblast-2) cell lines were inoculated with the supernatant of the homogenized and centrifuged pools of organ samples (Juhász et al., 2013). After incubation at 20 °C for 72 hours rounded cells and basophil cytoplasmic inclusion bodies were visualized by hematoxylin and eosin staining. Whole genome sequences of two virus isolates, designated 13051/2012 and 14612/2012, obtained from the 2nd passage on tissue culture were determined using previously described protocols (Bányai et al., 2014, 2015). In brief, after freezing and thawing, that was followed by concentration of the viral particles with PEG 6000 solution, DNA was purified from pelleted virus particles. The Ion Torrent compatible libraries were prepared from extracted genomic DNA samples applying the NEBNext® Fast DNA Fragmentation & Library Prep Set for Ion Torrent (New England Biolabs). The Ion Torrent Xpress barcode adapters (Life Technologies) were used to individually label different libraries. To obtain clonally amplified fragments emulsion PCR was carried out using the Ion OneTouch™ 200 Template Kit (Life Technologies) on a OneTouch v2 equipment (Life Technologies) as recommended by the manufacturer. Templated beads were enriched using an Ion OneTouch™ ES pipetting robot (Life

Technologies) and then the 200 bp sequencing protocol was performed on a 316 chip (Life Technologies) using the Ion Torrent PGM (Life Technologies) semiconductor sequencing equipment. Sequence reads were assembled and analyzed using the CLC Genomics Workbench v8.5 (www.clcbio.com) (using the following parameters: mismatch cost, 2; insertion cost, 3; deletion cost, 3; insertion open cost, 6; insertion extend cost, 1; deletion open cost, 6; deletion extend cost, 1; length fraction, 0.6; similarity fraction, 0.8) and the results were confirmed by the Geneious 8.1.7. software (www.geneious.com). As a result, the genomes of strains 13051/2012 and 14612/2012 were determined at an average coverage of 186X and 124X, respectively. At this coverage it was not possible to find minority sequence variants, if any, thus we consider the obtained consensus sequences to represent the genomes of the field strains.

Analysis of the whole genome sequences (GenBank accession numbers KT989884 and KT989885) revealed that the two brown bullheads were infected with ranaviruses closely related to the European sheatfish virus (ESV; GenBank accession no., JQ724856) isolated from moribund sheatfish (*Silurus glanis*) (Mavian et al., 2012). By using the MEGA6 software (choosing the p-distance and pairwise deletion options; www.megasoftware.net), we found nearly complete genome sequence identity (~99.9%) between the European sheatfish virus strain and the brown bullhead ranaviruses, 13051/2012 and 14612/2012. The genome lengths of the isolates 13051/2012 and 14612/2012 were 127,751 and 127,549 bp, respectively (Fig 2). In total, 136 ORFs were predicted using FgenesV (www.softberry.com) and DNA Javascript Translator 1.1 (Perry et al., 2002); thus, the coding capacity of ESV and the brown bullhead ranavirus genomes were shared. Among the 136 ORFs, 31 could be linked to a putative function, whereas the functions of 105 hypothetical gene products are currently unknown (Mavian et al., 2012). Sequence length differences were associated with major deletions in the intergenic regions (strain 13051/2012, position 4042 to 4075; strain 14612/2012, position 25512 to 25715). Single nucleotide substitutions were seen across the entire genome and many of these were shared in both Hungarian brown bullhead ranavirus strains (Fig 2). In the coding regions relatively few nucleotide substitutions were accompanied with amino acid change (Table 1). Collectively, the shared genomic structure and the high genome wide sequence similarity between brown bullhead and European sheatfish origin ranaviruses indicated that all three strains belong to the same ranavirus species (hereafter referred to as ESV).

When comparing various host origin ranavirus genomes using the Mauve program (darlinglab.org/mauve) we observed that organization of the ESV, Epizootic haematopoietic

necrosis virus (EHNV) and *Ambystoma tigrinum* virus (ATV) genomes was nearly identical. This confirmed their close genetic relationship (Fig 3), although the genome of ATV, which originates from an amphibian host (tiger salamander, *Ambystoma tigrinum*), was at least 20 kbp shorter than its counterparts in fish hosts. Analysis of the gene locations in the ranavirus genomes revealed that some homologous blocks of ORFs, e.g. those that encode the putative myristylated protein, the DNA dependent RNA polymerase a subunit, the NTPase/helicase and hypothetical proteins (on the left of the genomic map; Fig 3) and several hypothetical proteins (on the right of the genomic map) were present in all of the investigated genomes and had the same orientation. The remaining blocks of genes, covering the great majority of the genomes, showed various arrangements or differed in their orientation. In particular, gene blocks in the central genomic regions were highly variable with respect to their location and orientation, greatly increasing the variability of the genomic architecture and pointing out that during their evolution, ranaviruses have undergone multiple genomic rearrangement events. Analyzing the genome sequences of ranaviruses, Jancovich et al. (2010) concluded that amphibian ranaviruses may have originated from fish ranavirus ancestor(s). Our new sequence data on ESV strains showing the conserved arrangement of genes in four fish ranavirus genomes and one amphibian ranavirus genome is consistent with this hypothesis.

To help uncover the background of the independent disease outbreaks detected in brown bullhead in Hungary, phylogenetic analysis (using the software MEGA6, www.megasoftware.net) was performed on genes of the conserved major capsid protein, the DNA polymerase, and the neurofilament triplet H1-like protein, three genomic regions commonly used in molecular diagnostics and phylogenetic analysis (Holopainen et al., 2009) (Fig 4A). The Hungarian ESV isolates formed a common cluster with the prototype ESV (97.5-100% nt identity for the three genes; using MEGA6, www.megasoftware.net) and other ranaviruses detected in European siluriform fishes. Furthermore the EHNV was found to be the closest relative of this cluster (91.9-99.6% nt identity for the three genes), which together with the ATV, were typical members of the main groups of ALRV (Jancovich et al., 2010; Stöhr et al., 2015). Of interest, no marked sequence variability was observed in the genomic region encoding the neurofilament triplet H1-like protein, when the Hungarian ESV strains were compared. In this gene the Hungarian ESV sequences showed merely two nt differences (99.7% nt identity) without amino acid sequence change, whereas both isolates differed by 18 nt (97.5% identity) and 6 aa (97.4% aa identity) when compared with the homologous region of the ESV genome (Table 1). In addition, the sequence variability between brown bullhead and European sheatfish ESV strains was not better resolved in the phylogenetic tree generated

from the concatenated nucleotide sequence alignment containing 17 core genes (Fig 4B; Supplementary file) (Jancovich et al., 2010; Stöhr et al., 2015).

The geographic origin of ESVs detected in Europe is currently unclear. Additional ranavirus genome sequences would be required from the Americas where bullhead is indigenous to understand whether the detection of ESV strains in Hungarian brown bullhead specimens could have been a result of host species adaptation of ESV in imported siluriform fishes or the virus had been introduced with specimens imported from North America and this event had been followed by interspecies transmission event to the autochthonous European sheatfish. Nonetheless, partial or full length genome sequences from European indigenous and imported siluriform fishes strongly suggest that European ranavirus strains affecting populations of these fish species probably have a common evolutionary origin and interspecies transmission events among these fish species may contribute to the geographic dispersal of ESV or closely related ranaviruses.

In conclusion, the genomes of the Hungarian ESV ranavirus isolates were highly similar to each other, yet the slight differences we observed in the genome may help tracing the source and the spread of ranavirus infections in epidemiological follow-up. In this respect the genomic regions commonly used to delineate phylogeny of ranaviruses (i.e. MCP, DNA polymerase, the H1 region or other core proteins) are probably not ideal genetic markers due to the minimal sequence variation in these genomic regions. Instead, the authors of this report feel that whole genome sequencing and analysis should be the ultimate alternative of the genetic characterization of circulating ranavirus strains. This approach may help uncover relevant changes, if any, in the genome structure (incl. recombination or rearrangement events or major indel mutations including copy number variation in the repetitive regions) suitable for molecular epidemiological investigations. The ability of ESV ranaviruses to infect various (indigenous and imported) siluriform fish species (Ahne et al., 1989; Bovo et al., 1993; Juhász et al., 2013; Pozet et al., 1992) warrants closer monitoring of ranaviruses in order to aid prevention and control the decline of local fish populations.

Acknowledgements

This study was supported by the Momentum program awarded by the Hungarian Academy of Sciences (K. Bányai), the KTIA-AIK-12-1-2013-0017 Grant (S. Farkas), and the OTKA PD104315 (A. Doszpoly). E. Fehér, A. Doszpoly, S. L. Farkas, and S. Marton were supported by the Bolyai János Research Scholarship.

Compliance with ethical standards

Conflict of Interest: The authors report no conflict of interest.

Ethical approval: No animal experiments were done in the study. For routine laboratory investigation diseased animals were euthanized humanly according to current regulations (law 1998/XXVIII).

References

- Abrams, A.J., Cannatella, D.C., Hillis, D.M., Sawyer, S.L., 2013. Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. *J. Gen. Virol.* 94, 2082–2093.
- Ahne, W., Schlotfeldt, H.J., Thomsen, I., 1989. Fish viruses: Isolation of an icosahedral cytoplasmic deoxyribovirus from sheatfish (*Siluris glanis*). *J. Vet. Med. B* 36, 333–336.
- Bányai, K., Borzák, R., Ihász, K., Fehér, E., Dán, Á., Jakab, F., Papp, T., Hetzel, U., Marschang, R.E., Farkas, S.L., 2014. Whole-genome sequencing of a green bush viper reovirus reveals a shared evolutionary history between reptilian and unusual mammalian orthoreoviruses. *Arch. Virol.* 159, 153–158.
- Bányai, K., Palya, V., Dénes, B., Glávits, R., Ivanics, É., Horváth, B., Farkas, S.L., Marton, S., Bálint, Á., Gyuranecz, M., Erdélyi, K., Dán, Á., 2015. Unique genomic organization of a novel Avipoxvirus detected in turkey (*Meleagris gallopavo*). *Infect. Genet. Evol.* 35, 221–229.
- Bovo, G., Comuzi, M., De Mas, S., Ceschia, G., Giorgetti, G., Giacometti, P., Cappellozza, E., 1993. Isolamento di un agente virale irido-like da pesce gatto (*Ictalurus melas*) dallelevamento. *Bollettino Societa Italiana di Patologia Ittica* 11, 3–10.
- Holopainen, R., Ohlemeyer, S., Schütze, H., Bergmann, S.M., Tapiovaara, H., 2009. Ranavirus phylogeny and differentiation based on major capsid protein, DNA polymerase and neurofilament triplet H1-like protein genes. *Dis. Aquat. Organ.* 85, 81–91.
- Hyatt, A.D., Gould, A.R., Zupanovic, Z., Cunningham, A.A., Hengstberger, S., Whittington, R.J., Kattenbelt, J., Coupar, B.E., 2000. Comparative studies of piscine and amphibian iridoviruses. *Arch. Virol.* 145, 301–331.

Jancovich, J.K., Bremont, M., Touchman, J.W., Jacobs, B.L., 2010. Evidence for multiple recent host species shifts among the Ranaviruses (family *Iridoviridae*). *J. Virol.* 84, 2636–2647.

Juhász, T., Woynárovichné, L.M., Csaba G., Farkas, S.L., Dán, Á., 2013. Isolation of a ranavirus causing mass mortality in brown bullheads (*Ameiurus nebulosus*) in Hungary. *Magy. Állatorv. Lapja* 12, 763–768.

Mavian, C., López-Bueno, A., Fernández Somalo, M.P., Alcamí, A., Alejo, A., 2012. Complete genome sequence of the European sheatfish virus. *J. Virol.* 86, 6365–6366.

Perry, W.L., 2002. JavaScript DNA translator: DNA-aligned protein translations. *Biotechniques* 33, 1318–1320.

Pozet, F., Morand, M., Moussa, A., Torhy, C., de Kinkelin, P. 1992. Isolation and preliminary characterisation of a pathogenic icosahedral deoxyribovirus from the catfish *Ictalurus melas*. *Dis. Aquat. Organ.* 14, 35–42.

Stöhr, A.C., López-Bueno, A., Blahak, S., Caeiro, M.F., Rosa, G.M., Alves de Matos, A.P., Martel, A., Alejo, A., Marschang, R.E., 2015. Phylogeny and differentiation of reptilian and amphibian ranaviruses detected in Europe. *PLoS One* 10:e0118633.

Figure legends

Fig. 1. The geographic origin of ESV ranavirus strains examined in this study.

Fig. 2. Comparison of the Hungarian brown bullhead ranavirus genomes to the reference European sheatfish virus genome. The “right” and “left” reading direction of the predicted ORFs, respectively, are indicated above and below the line representing the viral genome. SNV, single nucleotide variant; SV, structural variant; Complex, includes various extent of rearrangement and substitutions.

Fig. 3. Similarity plot of the aligned ranavirus genomes, using the default parameters of the progressive Mauve alignment tool. Homologous boxes of ORFs are labeled with identical colors. Annotated genes located on the plus and minus strands are shown as white boxes.

Fig. 4. *Panel A*. Unrooted maximum-likelihood phylogenetic trees of the genes encoding the major capsid protein (MCP), the DNA polymerase (DP) and the neurofilament triplet H1-like

protein (H1). Nucleotide based phylogenetic trees were generated using 500 bootstrap replications with the Tamura 3-parameter model (MCP and DP) or the Jukes-Cantor model (H1) (MEGA6, www.megasoftware.net). Bootstrap values greater than 60 are shown at the branch nodes. Hungarian ranavirus isolates are labeled by dots. Abbreviations in taxon names: ADRV, Andrias davidianus ranavirus; ATV, Ambystoma tigrinum stebbensi virus; CGSIV, Chinese giant salamander iridovirus; CIV, Cod iridovirus; CMTV, Common midwife toad ranavirus; ECV, European catfish virus (A.m., Ameiurus melas; A.n., Ameiurus nebulosus); EHNV, Epizootic haematopoietic necrosis virus; ESV, European sheatfish virus (S.g., Silurus glanis); FV3, Frog virus 3; GGRV, German gecko ranavirus; PPIV, Pike-perch iridovirus; RGIV, Rana grylio iridovirus; SERV, Short-finned eel ranavirus; STIV, Soft-shelled turtle iridovirus; TFV, Tiger frog virus; TIV, Turbot iridovirus. *Panel B.* Unrooted maximum-likelihood phylogenetic tree of 17 concatenated ranavirus core genes using the General Time Reversible model with 500 bootstrap replicates (MEGA6, www.megasoftware.net). The position of the ORFs are represented according to the genome of the European sheatfish virus (JQ724856). The concatenated genes are as follows: major capsid protein (22L), DNA polymerase (61L), DNA-dependent RNA polymerase subunit a (11R), NTPase/helicase (12L), DNA repair enzyme RAD2 (17L), hypothetical protein (18R), thiol oxidoreductase (24L), deoxynucleoside kinase (27L), RNase III (34R), ribonucleotide reductase small subunit (53R), DNA-dependent RNA polymerase subunit b (58R), myristylated membrane protein (73L), hypothetical protein (97R), D5 family NTPase (112L), hypothetical protein (113R), protein kinase (117L), A32-like virion packaging ATPase (120L). Abbreviations in taxon names: ADRV, Andrias davidianus ranavirus; ATV, Ambystoma tigrinum stebbensi virus; CGSIV, Chinese giant salamander iridovirus; CIV, Cod iridovirus; CMTV, Common midwife toad ranavirus; ECV, European catfish virus (A.m., Ameiurus melas; A.n., Ameiurus nebulosus); EHNV, Epizootic haematopoietic necrosis virus; ESV, European sheatfish virus (S.g., Silurus glanis); FV3, Frog virus 3; GGRV, German gecko ranavirus; PPIV, Pike-perch iridovirus; RGIV, Rana grylio iridovirus; SERV, Short-finned eel ranavirus; STIV, Soft-shelled turtle iridovirus; TFV, Tiger frog virus; TIV, Turbot iridovirus.

Table 1. Synonymous and non-synonymous mutations detected in the ORFs of ESV strains 13051/2012 and 14612/2012 compared to the reference ESV genome (JQ724856)

ORF	Product	13051/2012		14612/2012	
		nt	aa	nt	aa
12L	NTPase/helicase	-	-	G1398A	-
27L	deoxynucleoside kinase	G94A	V32M	G94A	V32M
43R	NTPase/helicase	G634A	D212N	G634A	D212N
56L	dUTPase	A404C	D135A	A404C	D135A
58R	DNA-dependent RNA polymerase b subunit	G1805A	G602D	-	-
61L	DNA polymerase	C38A	S13Y	C38A	S13Y
71R	helicase	C489T	-	C489T	-
82R	p31K protein	nt301-306 AAGACC insertion	aa101-102 KT insertion	nt301-306 AAGACC insertion	aa101-102 KT insertion
105L	neurofilament triplet H1-like protein	C279A	S93R	C279A	S93R
		T286G	L96V	T286G	L96V
		G291A	-	G291A	-
		G319C	V107L	G319C	V107L
		T324G	N108K	T324G	N108K
		C325A	-	C325A	-
		C327A	-	C327A	-
		T333C	-	T333C	-
		G336C	-	G336C	-
		C343A	-	C343A	-
		T345A	-	T345A	-
		T351A	-	T351A	-
		T354G	-	T354G	-
		C355G	L119V	C355G	L119V
		A361C	-	A361C	-
		A363C	-	A363C	-
		C372G	-	C372G	-
		C373G	L125V	C373G	L125V
114L	dihydrofolate reductase	-	-	C102T	-

Note: Mutations in the putative ORFs with unknown function: 2L (14612/2012), 35L, 42R, 59L, 63L (13051/2012 and 14612/2012), 77R (13051/2012), 81L, 85R, 90R (13051/2012 and 14612/2012), 93R (14612/2012), 100R (13051/2012 and 14612/2012), 105R (13051/2012 and 14612/2012), 108L, 109R, 110L (13051/2012 and 14612/2012), 125R (13051/2012).

Figure 1



Figure 2

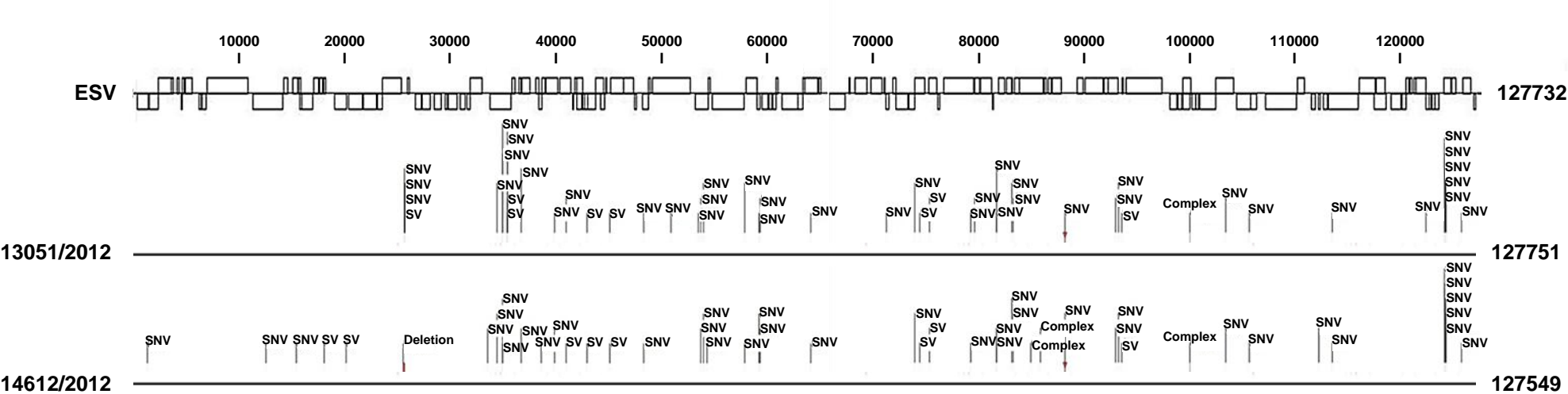


Figure 3

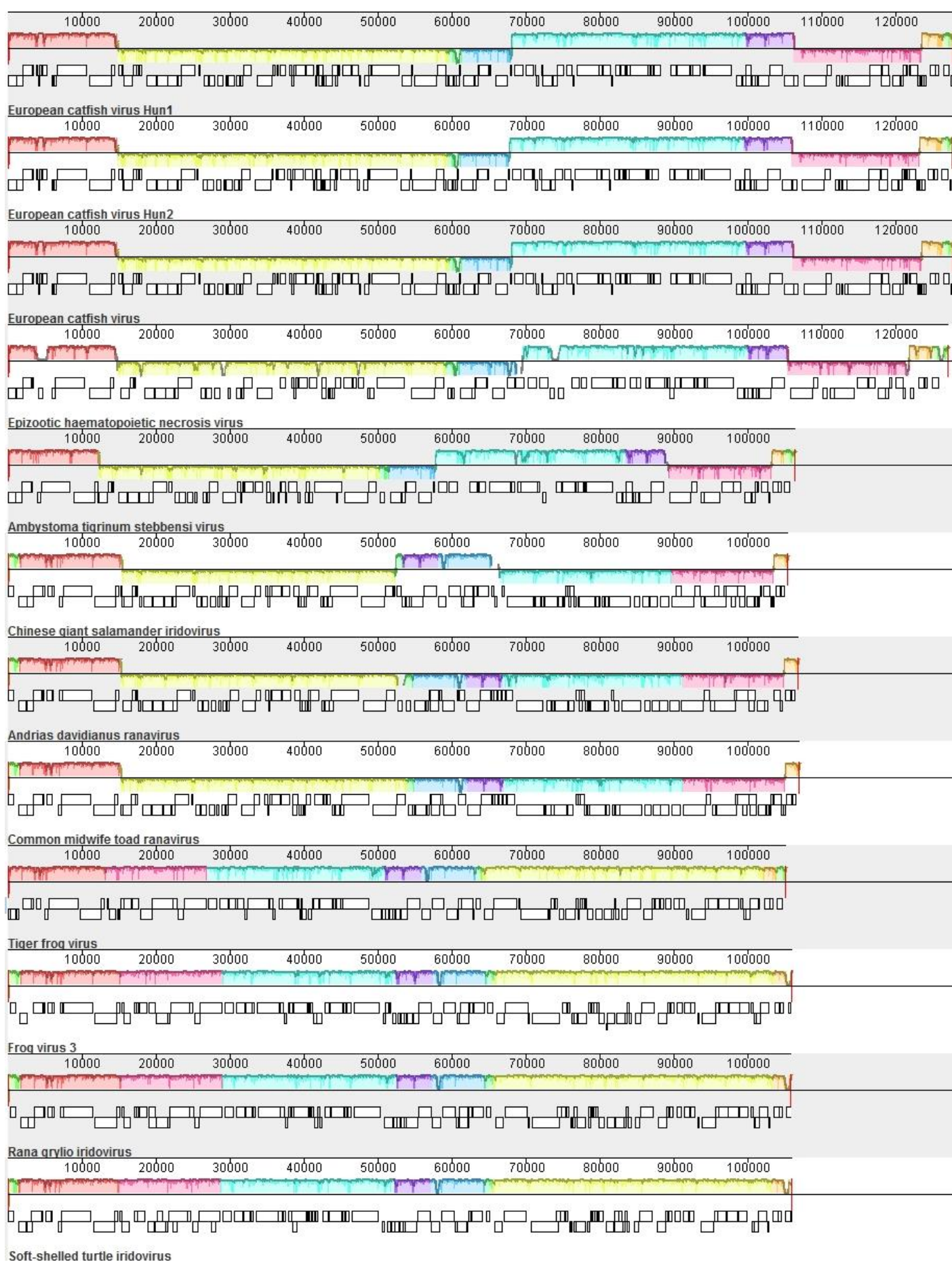


Figure A



B



Supplementary Material

[Click here to download Supplementary Material: Feher_ranavirus_IGE_supplfile_20160923_revision.doc](#)