Genetic analysis of a novel nidovirus from fathead minnows

William N. Batts,1 Andrew E. Goodwin2 and James R. Winton1

1US Geological Survey, Western Fisheries Research Center, 6505 Northeast 65th Street, Seattle, Washington 98115, USA
2University of Arkansas at Pine Bluff, 1200 N. University Drive, Mail Slot 4912, Pine Bluff, Arkansas 71601, USA

A bacilliform virus was isolated from diseased fathead minnows (Pimephales promelas). Analysis of the complete genome coding for the polyprotein (pp1ab), spike (S), membrane (M) and nucleocapsid (N) proteins revealed that the virus was most like white bream virus (WBV), another bacilliform virus isolated from white bream (Blicca bjoerkna L.) and the type species of the genus Bafinivirus within the order Nidovirales. In addition to similar gene order and size, alignment of deduced amino acid sequences of the pp1ab, M, N and S proteins of the fathead minnow nidovirus (FHMNV) with those of WBV showed 46, 44, 39 and 15 % identities, respectively. Phylogenetic analysis using the conserved helicase domain of the replicase showed FHMNV was distinct from WBV, yet the closest relative identified to date. Thus, FHMNV appears to represent a second species in the genus Bafinivirus. A PCR assay was developed for the identification of future FHMNV-like isolates.

Here, we report another fish nidovirus isolated in January 1997 from diseased fathead minnows (Pimephales promelas) being reared at a baitfish farm in central Arkansas, USA (Iwanowicz & Goodwin, 2002). The ponds were stocked with minnows that had been distributed throughout the USA, potentially mixed with farm-raised or wild fish from other regions of the USA and returned to the farm (a practice no longer followed). Affected fish suffered from chronic mortality with clinical signs that included haemorrhage of the eyes and skin. Healthy fathead minnows challenged with the virus developed similar signs and the agent was reisolated from moribund fish. The isolate was initially characterized as a bacilliform virus, possibly belonging to the family Rhabdoviridae, although the virus had the uncharacteristic property of inducing syncytia in infected cell cultures (Iwanowicz & Goodwin, 2002).

To further characterize the isolate, cultures of fathead minnow cells (Gravell & Malsberger, 1965) were propagated at 25 °C in Eagle’s minimum essential medium (MEM) supplemented with 10 % FBS, inoculated with virus and incubated at 15 °C. Cytopathic effects, which included formation of syncytia, occurred by days 2 or 3 with complete monolayer involvement by days 6 or 7. Virus-containing supernatant was clarified by centrifugation at 1000 g for 10 min and aliquots frozen at −80 °C.

Because the virus was initially thought to be a rhabdovirus, polymerase gene sequences from representative fish rhabdoviruses were aligned to identify conserved regions for amplification by RT-PCR. Viral genomic RNA was released from a 1:50 dilution of clarified supernatant by heating for 2 min at 95 °C and subjected to RT-PCR according to published procedures (Huang et al., 1996; Ahne et al., 1999). The PCR products were purified with a StrataPrep kit (Stratagene) and sequenced by using a 310 genetic analyser (Applied Biosystems). An ORF of approximately 100 nt was initially obtained and the
authentic sequence was used to design primers for 5’ and 3’ RACE (Promega). Eventually, sufficient sequence was obtained in the 3’ direction to show similarity (44% aa identity) with the membrane (M) protein of the fish nidovirus, WBV (GenBank accession no. DQ898157).

The full genome sequence of the fathead minnow nidovirus (FHMNV) was determined by using BigDye chemistry and a 3130 genetic analyser (Applied Biosystems), edited using Sequencher 4.1 (Gene Codes) and aligned with the CLUSTAL W algorithm in MacVector 6.0 software (Accelrys). The FHMNV genome contained 27 291 nt (GenBank accession no. GU002364) including four major ORFs and terminal regions. The genome organization for FHMNV appeared to be the same as described by Schütze et al. (2006) for WBV: a 5’ UTR (759 nt), the 1a–1b polyprotein (pp1ab; 7174 aa), spike glycoprotein (S; 1190 aa), M protein (225 aa), N protein (165 aa) and a 3’ UTR (195 nt) ending in a long poly(A) tract. The ORFs encoding the pp1ab, S, M and N proteins of FHMNV were 21 522, 3570, 675 and 495 nt in length, respectively, similar to those of WBV. The relative simplicity of the genome organization of both FHMNV and WBV was evident as there were no accessory proteins encoded among the structural protein genes and no haemagglutinin-esterase (HE) gene was found as present in the toroviruses and some coronaviruses (Gorbalenya et al., 2006; de Groot et al., 2012a).

The same ‘slippery sequence’ was evident in both FHMNV and WBV and consisted of the heptanucleotide ‘UUUAACG’ upstream of the ORF1a stop codon, where a -1 frameshift would generate the predicted full-length pp1ab of 7174 aa as observed for other nidoviruses having a large replicase gene region (Britton & Cavanagh, 2008). A nidovirus-specific ‘SDD’ signature sequence in the C-terminal portion of pp1ab of WBV was also evident in the genome of FHMNV encoded by nt 17 361–17 369.

The long (759 nt) 5’ UTR of the FHMNV genome shared a 46% identity with that of WBV. The entire 3’ UTR of FHMNV showed a 51% sequence identity with WBV, while a 34 nt region of the 3’ UTR adjacent to the poly(A) tract was 73% identical. Genomic FHMNV UTR sequences upstream of ORFs 2, 3 and 4 were nearly identical. Conserved non-anucleotide sequences predicted to be the transcription-regulating sequence elements of WBV (Schütze et al., 2006) were found to be nearly identical in FHMNV. Nidoviruses, including WBV, produce subgenomic RNAs that share a 5’ leader sequence identical to the 5’ end of the genome (Siddell & Snijder, 2008). To determine whether FHMNV-infected cell cultures contained subgenomic RNAs similar to those of WBV, we used RT-PCR primers for 5’ leader and 3’ body sequences of ORFs 2, 3 and 4 that generated DNA fragments of the predicted sizes, in support of the subgenomic RNA synthesis model (data not shown).

The large replicase polyproteins of the nidoviruses are processed by viral proteases to yield more than a dozen mature proteins, including a chymotrypsin-like (3C-like or ‘main’) protease and a 5’-to-3’ helicase (de Groot et al., 2012b). To compare the helicase and main protease domains within the replicase polyprotein of WBV (Schütze et al., 2006; Ulferts et al., 2011) with orthologous sequences of FHMNV, pairwise alignments of predicted amino acid sequences were performed using CLUSTAL W. The helicase domain of FHMNV (aa 5644–5924) shared a 70% identity with that of WBV. The main protease domain of FHMNV (aa 3785–3910) was 62% identical to that of WBV and contained the highly conserved Ser–His–Asp catalytic triad (aa 3894, 3798 and 3824, respectively) recently described for WBV by Ulferts et al. (2011).

The FHMNV S glycoprotein had a putative furin cleavage site (RKKR↓ residues 803–806) that gives rise to an active fusion peptide frequently associated with a syncytial form of cytopathic effect and best known among viruses in the family Paramyxoviridae (Franke et al., 2006). Characteristic of many fusion peptides, the deduced amino acid sequence of the FHMNV fusion peptide contained an abundance of alanines (24% of the amino acids) in this region. A furin cleavage site was also identified for the glycoprotein of WBV by Bosch & Rottier (2008); however, the furin cleavage site and fusion peptide of FHMNV aligned more closely with those of representative paramyxoviruses than to WBV (Fig. 1a). The fusion peptide sequence of FHMNV aligned poorly with homologous regions of coronaviruses or toroviruses.

For WBV, three heptad repeats are predicted in the S glycoprotein by the program Multicoil (Bosch & Rottier, 2008) and by the updated program Paircoil2 (Online); however, the S glycoprotein of FHMNV was predicted to have only two heptad repeats by analysis with Paircoil2 (data not shown). Even though there was apparent functional similarity, the amino acid sequences were too dissimilar in these regions to align with confidence.

The calculated pI of the FHMNV S protein was 6.0, similar to that of WBV, gill-associated virus (GAV) and equine torovirus (EToV) at pI 6.2–6.6, which may aid in the fusion of these viruses with host cells. M and N proteins of FHMNV had calculated pIs of 8.8 and 10.0, respectively, similar to many members within the order Nidovirales.

The FHMNV S glycoprotein contained a 223 aa region with relatively high identity (27%) and significant E-value (3 × 10^-7) with the chitinase of Agrotis segetum nucleopolyhedrovirus (Fig. 1b). Chitinases hydrolyse the natural biopolymer of chitin and have been identified in baculoviruses as late gene products involved, along with cathepsin, in terminal liquefaction of the host and facilitating dispersal into the environment (Hawtin et al., 1997). The presence of this chitinase domain may provide evidence for an insect host or reservoir in the biology of FHMNV or, alternatively, to an evolutionary link with invertebrates, such as crustaceans, which are infected by nidoviruses in the family Roniviridae.

In order to compare the genetic similarities among FHMNV, WBV and representative members of the
established families of nidoviruses, pairwise alignments of both nucleotide and predicted amino acid sequences were performed using CLUSTAL W. In addition to the predicted amino acid sequences of the ORFs, primers 5'-CACATT-CTTAACACACTCAAGAA-3' and 5'-GTTTTCTTTTTCCACGTTCTGGGGCCA-3' were designed to amplify the relatively conserved 281 aa region of the helicase domain of FHMNV (residues Ala5644 to Cys5924). The FHMNV helicase sequence was deposited separately in GenBank under the accession number GU002365. The predicted amino acid identities varied considerably when pairwise alignments were performed on individual ORFs or the helicase domain (Table 1). The highest amino acid identities between FHMNV and WBV were obtained for the helicase domain (70%) followed by the entire 1ab polyprotein itself (46%). The predicted amino acid sequence of the S glycoprotein of FHMNV was only 15% identical to that of WBV, while the M and N proteins showed 44 and 39% identity, respectively. Relatively low amino acid identities were observed with members of other genera of the order Nidovirales with the best alignments obtained for the helicase domain and polyprotein of toroviruses (Table 1).

The deduced amino acid sequences of the full-length S, M, N genes, the helicase domain and 3' portions of the 1ab gene of FHMNV were used to infer phylogenetic relationships with other representative viruses belonging to the order Nidovirales, including WBV of the genus Bafinivirus. Predicted amino acid sequences were aligned with homologues from selected members of the order Nidovirales using porcine reproductive and respiratory syndrome virus (PRRSV; GenBank accession no. U87392) as the outgroup. Alignments were done with CLUSTAL W followed by phylogenetic analyses using the neighbour-joining and parsimony programs in the PAUP* version 4.0b software package (Swofford, 1998). Consensus trees were derived using 1000 bootstrap replicates of the datasets, and bootstrap values above 70 were considered significant (Hillis & Bull, 1993). These relationships are shown in Fig. 2 as a phylogenetic tree generated using sequences of the helicase domain. In this unrooted tree, FHMNV and WBV are shown to be most closely related, clearly branching together and separated from other groups of nidoviruses. The closest neighbouring group was the torovirus cluster lending added support to the placement of the genus Bafinivirus within the subfamily Torovirinae (de Groot et al., 2012a).

### Table 1

<table>
<thead>
<tr>
<th>Virus</th>
<th>ORF</th>
<th>Predicted Amino Acid Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHMNV</td>
<td>1ab</td>
<td>46%</td>
</tr>
<tr>
<td>WBV</td>
<td>1ab</td>
<td>15%</td>
</tr>
<tr>
<td>FHMNV</td>
<td>S</td>
<td>44%</td>
</tr>
<tr>
<td>WBV</td>
<td>S</td>
<td>15%</td>
</tr>
<tr>
<td>FHMNV</td>
<td>M</td>
<td>39%</td>
</tr>
<tr>
<td>WBV</td>
<td>M</td>
<td>15%</td>
</tr>
<tr>
<td>FHMNV</td>
<td>N</td>
<td>39%</td>
</tr>
<tr>
<td>WBV</td>
<td>N</td>
<td>15%</td>
</tr>
</tbody>
</table>

### Fig. 1

(a) Alignment of the fusion peptide region of FHMNV S protein with furin cleavage sites (CS) and fusion peptides of selected paramyxoviruses. Amino acids identical to the consensus are shown in black, similar amino acids in grey. Numbers are locations within the ORF for the respective proteins. Abbreviations and accession numbers used are: AMPV, avian metapneumovirus (GenBank accession no. EAW43430); HPIV3, human parainfluenzavirus 3 (GenBank accession no. ABZ85672); J virus (GenBank accession no. AAX86031); FDLV, Fer-de-Lance virus (GenBank accession no. AAN18264); WBV, white bream virus (GenBank accession no. ABI97395). (b) Comparison of the FHMNV S protein with ASNV, *Agrotis segetum* nucleopolyhedrovirus chitinase (GenBank accession no. AAZ38189) domain. Identical residues are shown in black, similar amino acids in grey.
Most of the other phylogenetic trees using structural protein sequences provided a less informative topology, due to low genetic similarity. Phylogenetic analysis using S glycoprotein sequences produced the poorest quality trees, probably due to the very low degree of identity (Table 1).

The evolutionary history and taxonomy of the family *Nidovirales* is an area of active interest. Within members of the family, large differences in morphology, genome size and the numbers or types of accessory proteins suggest these viruses have evolved through the processes of both mutation and recombination; although the selective forces driving virus evolution remain poorly understood (Gorbalenya, 2008). Members of the genus *Bafinivirus* group together by genome size with the toro-, corona- and roniviruses in having an unusually large RNA genome of 26–32 kb, compared with the smaller arteriviruses at 13–26 kb (Gorbalenya et al., 2006). The PCR-based detection method described by Ahne et al. (1999), except that reverse transcription was performed for 15 min at 50 °C, primer annealing during the 30 cycles was at 50 °C for 30 s, and a final extension step at 72 °C for 7 min was included. Amplified products were purified and sequenced to confirm their identity. This RT-PCR assay was used to confirm the identity of several additional syncytia-producing viruses isolated from fathead minnows reared in the states of Minnesota, Wisconsin and Illinois as well as muskellunge (*Esox masquinongy*) from Nebraska. Nearly identical sequences were obtained for these new isolates. Interestingly, upon follow-up investigation, the captive muskellunge were being fed a diet of live fathead minnows.

The impact of FHMNV on wild and cultured fathead minnows or other fish species is not known. Mortality in the original case in Arkansas was significant and laboratory challenges demonstrated that the virus produced clinical signs and mortality similar to the original case (Iwanowicz & Goodwin, 2002). Extirpation of the infected population in Arkansas seems to have been effective in eradicating the virus from farms and wild fish in that state. Eradication may have been made more effective by the hot summer temperatures of central Arkansas. The origins of more recent isolates of FHMNV seem to indicate that there is a reservoir of the virus in fathead minnows in more northerly states. Many cases of FHMNV infection may go unconfirmed because the syncytial form of cytopathic effect caused by FHMNV is similar to that produced by members of the genus *Aquareovirus* that are often isolated from healthy fish, including fathead minnows (Goodwin et al., 2006). The PCR-based detection method described in this paper should aid in rapidly determining the viral status of fathead minnows and easily identify cytopathic

## Table 1. Comparison of the pairwise per cent identities in deduced amino acid sequences of FHMNV proteins to those of other representative nidoviruses

The numbers in parentheses are the numbers of amino acids aligned for each domain or ORF. NA, Not applicable; WBV, white bream virus; EToV, equine torovirus; GAV, gill-associated virus; IBV, infectious bronchitis virus; EAV, equine arteritis virus.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Helicase (281) l ab domain</th>
<th>ORF 1ab (7174) polyprotein</th>
<th>ORF 2 (1190) S</th>
<th>ORF 3 (225) M</th>
<th>ORF 4 (165) N</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBV</td>
<td>70</td>
<td>46</td>
<td>15</td>
<td>44</td>
<td>39</td>
</tr>
<tr>
<td>EToV</td>
<td>35</td>
<td>18</td>
<td>13</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>GAV</td>
<td>25</td>
<td>14</td>
<td>14</td>
<td>NA</td>
<td>14</td>
</tr>
<tr>
<td>IBV</td>
<td>23</td>
<td>13</td>
<td>15</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>EAV</td>
<td>16</td>
<td>11</td>
<td>NA</td>
<td>12</td>
<td>NA</td>
</tr>
</tbody>
</table>

The family *Roniviridae* was named for *Ron*, a sigla for rod-shaped nidovirus referring to the virion morphology of viruses in the order. The family currently contains the single genus *Okavirus* consisting of several genotypes with GAV of shrimp as the type species. Both FHMNV and WBV also have this bacilliform or rod-shaped morphology (Granzow et al., 2001; Iwanowicz & Goodwin, 2002), but are not closely related genetically to GAV, although they similarly occur in aquatic hosts.
effects due to FHMNV in cell cultures. Further study using this method will provide information useful to fish culturists, regulators and managers of natural fisheries.

Acknowledgements

We thank Raoul de Groot of the Faculty of Veterinary Medicine, University of Utrecht for providing the latest information on nidovirus taxonomy, Gael Kurath of the USGS Western Fisheries Research Center for her insights into phylogenetic analyses, Janet Warg of the USDA National Veterinary Services Laboratories for helpful discussions, and Journal of General Virology’s Editor John Ziebuhr of The Queen’s University of Belfast and two anonymous reviewers for important suggestions that significantly improved the manuscript. Additional isolates of FHMNV were kindly provided by David Thompson of the Utah Division of Wildlife Resources and Kathy Kurth of the Wisconsin Veterinary Diagnostic Laboratory. This work was funded by the US Geological Survey. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government.

References


