Distribution and variation of NV genes in fish rhabdoviruses

Gael Kurath, Keith H. Higman and Harry V. Björklund†

Northwest Biological Science Center, National Biological Service, 6505 NE 65th St, Seattle, Washington 98115, USA

The fish rhabdovirus infectious haematopoietic necrosis virus (IHNV) contains a non-virion (NV) gene between the glycoprotein (G) and polymerase (L) genes on its RNA genome. The present study investigated three other fish rhabdovirus genomes and found that the NV gene of hirame rhabdovirus is closely related to the NV of IHNV, whereas the viral haemorrhagic septicemia NV gene showed evidence of significant divergence. Most importantly, spring viraemia of carp virus, the only vesiculovirus-like fish rhabdovirus examined, did not have an NV gene at its genomic RNA G–L junction. These results suggest that the presence of an NV gene is characteristic of the unassigned fish rhabdovirus subgroup previously classified as lyssaviruses, and that the NV gene is not essential for replication in fish cells per se, since it is absent in a vesiculovirus-like fish rhabdovirus.

Rhabdoviruses are among the most widely distributed viruses in nature, infecting many plant, invertebrate and vertebrate hosts (Wunner & Peters, 1991). Those that infect fish are particularly interesting because their hosts live in a wide variety of habitats and include such diverse fish as salmon, trout, cod, eel, carp,pike, perch, flatfish and snakehead (Wolf, 1988; Hetrick & Hedrick, 1993). When these rhabdoviruses were first described they were designated as members of either the Lyssavirus or Vesiculovirus mammalian rhabdovirus genera on the basis of the electrophoretic migration of their proteins (Hill et al., 1975; McAllister & Wagner, 1975; Wunner & Peters, 1991). Since it is currently recognized that these classifications require modification we will use the terms ‘unassigned’ and ‘vesiculo-like’ to distinguish the two subgroups of fish rhabdoviruses previously classified as lyssaviruses and vesiculoviruses respectively. This is done in accordance with the most recent ICTV report (Wunner et al., 1995), with the understanding that the unassigned subgroup previously classified as lyssaviruses may be designated as an independent genus in the future (Morzunov et al., 1995; Björklund et al., 1996).

Infectious haematopoietic necrosis virus (IHNV) is a fish rhabdovirus in the unassigned subgroup which lethally infects many salmonid fish species. The genome organization of IHNV is known (Kurath et al., 1985), and there is complete or nearly complete sequence information available for the genomic RNA of three IHNV strains (Koener et al., 1987; Gilmore & Leong, 1988; Morzunov et al., 1995; Schuetze et al., 1995; our unpublished data). In addition to the five structural genes common to other rhabdoviruses, the IHNV genome contains a unique sixth gene, NV, which encodes a non-virion protein that is expressed in infected cells but is not present in purified virions (Kurath & Leong, 1985; Schuetze et al., 1996). The NV protein is approximately 12000 Da (Kurath & Leong, 1985), and the NV gene location is between the glycoprotein (G) and polymerase (L) genes, making the IHNV genomic order 3′ N-M1-M2-G-NV-L 5′ (Kurath et al., 1985). The sequences of the NV genes of several IHNV strains have been determined (Morzunov et al., 1995; Nichol et al., 1995; Schuetze et al., 1995). The function of the NV gene and/or protein is unknown, and there are no reports of genes analogous to NV in the genomes of rhabdoviruses of non-fish hosts.

In this study the G–L junctions were characterized for three additional fish rhabdoviruses chosen to represent various degrees of relatedness to IHNV. Hirame rhabdovirus (HIRRV) is a cold water virus in the unassigned subgroup, isolated from flounder (Paralichthys olivaceus) and ayu (Plecoglossus altivelis) in Japan (Kimura et al., 1986). HIRRV is thought to be closely related to IHNV due to serological cross-reactivity (Nishizawa et al., 1991) and similarity of the G gene sequences (Björklund et al., 1996). Viral haemorrhagic septicemia virus (VHSV), a cold water unassigned virus of salmonids, is more distantly related to IHNV but within the same subgroup (Wolf, 1988). Spring viraemia of carp virus (SVCV) is most distantly related to IHNV because it is a warm water vesiculo-like virus isolated from common carp (Cyprinus carpio) in Europe (Wolf, 1988).
Table 1. Features of (a) the NV genes and (b) predicted NV proteins, of IHNV, HIRRV and VHSV

Values for (a) are in nucleotides.

<table>
<thead>
<tr>
<th>Virus</th>
<th>NV gene length*</th>
<th>Upstream UTR length</th>
<th>ORF location</th>
<th>ORF length</th>
<th>Downstream UTR length</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHNV</td>
<td>371</td>
<td>28</td>
<td>29–361</td>
<td>333</td>
<td>0</td>
</tr>
<tr>
<td>HIRRV</td>
<td>378</td>
<td>35</td>
<td>36–368†</td>
<td>333</td>
<td>0</td>
</tr>
<tr>
<td>VHSV</td>
<td>423</td>
<td>23</td>
<td>24–389</td>
<td>360</td>
<td>24</td>
</tr>
</tbody>
</table>

(b) Predicted NV proteins

<table>
<thead>
<tr>
<th>Virus</th>
<th>No. of amino acids</th>
<th>Mol. mass (calculated)</th>
<th>Estimated pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHNV</td>
<td>111</td>
<td>13314 Da</td>
<td>6.5</td>
</tr>
<tr>
<td>HIRRV</td>
<td>111</td>
<td>12720 Da</td>
<td>8.4</td>
</tr>
<tr>
<td>VHSV</td>
<td>122</td>
<td>13795 Da</td>
<td>6.2</td>
</tr>
</tbody>
</table>

* Gene length is defined here as beginning immediately after the (A) of the putative transcription termination signal of the G gene, and ending with the (A) putative transcription termination signal of the NV gene.
† The HIRRV gene also has an upstream in-frame ATG at nt 9–11 which is in a poor Kozak consensus context and is not considered likely to initiate translation for reasons described in the text.

The virus strains used in this work were IHNV RB1, in which the NV gene was first discovered, (J. C. Leong, Oregon State University, Corvallis, Oreg., USA), HIRRV 3401-H (T. Nishizawa, University of Hiroshima, Higashi-Hiroshima, Japan), the Danish F1 strain of VHSV (P. de Kinkelin, INRA, Jouy-en-Josas, France), and the reference strain of SVCV (N. Fijan, University of Arkansas, Pine Bluff, Ark., USA). Viruses were propagated in epithelioma papulosum cyprini cells at 15 °C (SVCV) or 25 °C (IHNV, HIRRV, VHSV) as described (Bjorklund et al., 1996), followed by virus purification and RNA extraction (Kurath & Leong, 1985). Libraries of clones containing cDNA to the purified genomic RNAs of HIRRV, VHSV and SVCV were generated (cDNA Synthesis System Plus kit, Amersham) and screened to identify clones representing the region between the G and L genes of each virus. At the same time a new library of cDNA clones to the IHNV RB1 strain was prepared and screened to re-confirm the presence and sequence of the known NV gene.

Sequence analyses of these clones revealed that the IHNV, HIRRV and VHSV genomes each had a small gene, similar in size to the known IHNV NV gene, at their G–L junctions. RT–PCR (Jorgensen et al., 1995) was then used to obtain clones representing exact full-length NV genes of IHNV, HIRRV and VHSV. In contrast, the G–L junction of SVCV consisted of only four nucleotides, CTAT, which were identical in three independent cDNA clones. While this is slightly different from the dinucleotide CT intergenic regions at every other SVCV gene junction (Bjorklund et al., 1996), it was clear that SVCV has no NV gene equivalent. Thus, all three unassigned fish rhabdoviruses had NV genes, and the only vesiculo-like fish rhabdovirus analysed had no NV gene.

For each unassigned virus a consensus NV sequence from at least four independent clones was generated (GenBank accessions U47846–U47848). Since the transcription initiation and termination signals for these genomes have not been experimentally determined, we will consider the NV genes to begin immediately following the (A) of the putative translation termination signal of the G gene. All three NV genes began with the conserved CTTGGCAC sequence (mRNA-sense), described as a putative transcription initiation signal unique to the unassigned fish rhabdoviruses, and ended with the conserved AGATA(gAA)G (mRNA-sense) putative transcription termination–polyadenylation signal (Morzunov et al., 1995; Bjorklund et al., 1996). Using this definition the lengths of the NV genes of IHNV, HIRRV and VHSV were 371, 378 and 423 nt respectively.

The NV sequences were analysed using MacVector version 4.1.4 software (International Biotechnologies). Pairwise nucleotide and amino acid sequence comparisons and alignments were done using the Genetics Computer Group (GCG) Wisconsin version 8.0 software. Features of the three NV gene consensus sequences are summarized in Table 1. For each sequence there was a single major open reading frame (ORF) extending for most of the length of the gene. The IHNV sequence also had a second smaller ORF in a different plus-sense reading frame extending from nt 160 to 336, which was noted previously (P. W. Chiou & J. C. Leong, unpublished data). The significance of this second ORF, if any, is unknown.
Fish rhabdovirus NV genes

Fig. 1. Alignment of the predicted NV proteins encoded by the IHNV RB1, HIRRV 3401-H and VHSV F1 NV gene sequences. Amino acid sequence alignments were generated by the Pileup program in GCG, using settings gap weight 3±00 and length weight 0±10. Identical (*) and functionally similar (C) amino acids are indicated for HIRRV and VHSV relative to the IHNV sequence. Functional similarity was assigned using the following groups (I, L, V, M), (F, W, Y), (P, G), (S, T, A), (D, E), (N, Q), (K, R, H) and (C), as described by Poch et al. (1990).

Table 2. Comparison of percentage identity/percentage similarity values for different proteins* of IHNV, HIRRV and VHSV

<table>
<thead>
<tr>
<th>Gene</th>
<th>IHNV cf. HIRRV</th>
<th>IHNV cf. VHSV</th>
<th>HIRRV cf. VHSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (partial†)</td>
<td>793/89±1</td>
<td>52/3/74±1</td>
<td>50/0/71±8</td>
</tr>
<tr>
<td>M1</td>
<td>63/9/78±4</td>
<td>34/7/49±1</td>
<td>33/8/50±2</td>
</tr>
<tr>
<td>M2</td>
<td>72/0/79±8</td>
<td>36/1/55±2</td>
<td>40/0/56±8</td>
</tr>
<tr>
<td>G</td>
<td>74/3/83±3</td>
<td>38/8/54±2</td>
<td>38/9/54±6</td>
</tr>
<tr>
<td>NV</td>
<td>54/1/72±1</td>
<td>23/3/47±6</td>
<td>16/5/40±4</td>
</tr>
</tbody>
</table>

* Amino acid sequences used were predicted from known nucleotide sequences of IHNV RB1 (N, E. Emmenegger & G. Kurath, unpublished, U50402; M1 and M2, P. Ormonde, G. Kurath & J. C. Leong, unpublished; G, E. Emmenegger & G. Kurath, unpublished, U50401; NV, this work), HIRRV 3401-H (N, M1 and M2, T. Nishizawa & G. Kurath, unpublished, D45422; G, Bjorklund et al., 1996, U24073; NV, this work), VHSV 0771 (N, Bernard et al., 1990, D00687; M1 and M2, Benmansour et al., 1994, U02629 and U03502; G, Thiry et al., 1991, X59148) and VHSV F1 (NV, this work).
† Only a partial N gene sequence is available for HIRRV, so only a 175 amino acid region, corresponding to amino acids 144–317 of the 412 amino acid IHNV RB1 N gene, has been compared between the three viruses for this table.

The HIRRV major ORF starts at the first ATG at nt 9–11, but this ATG is in a poor context for initiation of translation, with a C in the -3 position and a T at +4 (Kozak, 1986). The next in-frame ATG, at nt 36–38, is in a very strong Kozak consensus context, and would encode an NV protein that aligns closely with the IHNV NV protein (described below). Thus, it is probable that translation of the HIRRV NV protein initiates at the second ATG, and we have assumed this in Table 1 and for our subsequent analyses. It should be noted, however, that this assumption is not proven.

Computer comparisons showed that the IHNV and HIRRV NV genes are closely related. Alignment of these two sequences showed 63±0% nucleotide sequence identity, with regions of identity distributed quite evenly throughout the genes (data not shown). Similarly, the predicted NV protein sequences encoded by the IHNV and HIRRV genes show significant alignment (54±1% identity, 72±1% similarity, Fig. 1 and Table 2), with the similarities distributed evenly throughout the protein. In contrast, the VHSV NV nucleotide sequence showed at best 38±9% and 37±9% identities with the HIRRV and IHNV sequences respectively. The VHSV predicted NV amino acid sequence alignment with the IHNV and HIRRV NV proteins is shown in Fig. 1. Although the NV amino acid identity values between VHSV and IHNV or HIRRV are difficult to distinguish conclusively from background levels, the amino acid similarity values are higher (Table 2), indicating a distant but discernible relationship between these genes.

Clear support of the relationships described above is evident upon comparison of the hydrophilicity plots for these three NV proteins (data not shown). The NVs of IHNV and HIRRV are highly similar, while the VHSV NV protein bears very little resemblance to the other two plots. The only region showing...
any possible structural conservation is from approximately amino acid 80 to the C terminus, while the first 80 amino acids indicate a striking lack of conservation of overall structure.

The NV sequences described here have high levels of homology with other NV gene sequences available in the literature. For IHNV, the NV sequences are now available for 13 strains (Nichol et al., 1995; Schuetze et al., 1995), in addition to the extant RB1 strain sequence given here. These genes are all identical in length, have ORFs of identical lengths, and have greater than 97% overall nucleotide sequence identity. For VHSV, the NV gene sequences of two other European isolates are now available (Basurco & Benmansour, 1995; Schuetze et al., 1996) in addition to the VHSV F1 NV sequence given here. These genes are all 422–423 nt long and they have ORFs of identical lengths. The VHSV F1 sequence shows greater than 98% sequence identity with the European VHSV isolate NVs. There are no other HIRRV NV gene sequences or SVCV G–L junction sequences available for comparison.

The three NV genes described here reveal that the predicted NV proteins of IHNV and HIRRV have distinct similarities in both amino acid sequence and overall structure, while the NV of VHSV is very different, as suggested previously (Basurco & Benmansour, 1995). This suggests that there is a relatively low level of evolutionary constraint on the NV gene. Comparisons of amino acid sequences confirm that the NV genes of these three viruses are significantly more diverse than the other structural genes of the same three viruses (Table 2), as has been noted previously (Morzuov et al., 1995; Basurco & Benmansour, 1995). If the NV gene itself has been retained but the NV protein products bear little resemblance, then perhaps the role of the NV protein has diverged between these viruses. Whatever this diverged role is, it is unlikely to involve a host-specific phenomenon, since the natural hosts for IHNV and VHSV are both salmon and trout, while the natural hosts for HIRRV are flounder and ayu.

The viruses characterized differ in that the temperature range for SVCV is 4–31 °C, with optimum replication at 20–22 °C, while IHNV, HIRRV and VHSV all have colder temperature ranges of 4–20 °C, with replication optima of approximately 10–15 °C (Wolf, 1988). Any possible significance of these temperature characteristics to the presence or function of the NV gene must await future studies of warm water unassigned and/or cold water vesiculo-like fish rhabdoviruses.

The finding that SVCV has no NV gene allows us to propose that the presence of an NV gene is a genetic hallmark of the unassigned fish rhabdoviruses. It also means that, whatever the role of NV, it is not required for replication in fish cells per se. SVCV, and possibly the other fish vesiculo-like rhabdoviruses, either have no requirement for the function carried out by the NV protein in unassigned fish rhabdoviruses, or they use some other means to carry out the same function, such as incorporation into polymerase functions.

Within the general similarity of gene organization of rhabdovirus RNA genomes, the G–L gene junction has been found to contain a high level of flexibility and is a major point of distinction between the different established genera of rhabdoviruses. The presence of NV genes at the G–L junctions of all three fish rhabdoviruses in the unassigned subgroup is strong support for the suggestion made previously (Morzuov et al., 1995; Bjorklund et al., 1996) that a new genus should be established for these viruses within the family Rhabdoviridae.

We would like to thank Pinwen P. Chiou and JoAnn Leong of Oregon State University for access to unpublished sequence data and for many fruitful discussions throughout this work.

References


Received 29 August 1996; Accepted 17 September 1996