Sequence analysis of the genome of piscine orthoreovirus (PRV) associated with heart and skeletal muscle inflammation (HSMI) in Atlantic salmon (Salmo salar)

Torstein Tengs, Dr.
Sequence Analysis of the Genome of Piscine Orthoreovirus (PRV) Associated with Heart and Skeletal Muscle Inflammation (HSMI) in Atlantic Salmon (Salmo salar)

Turhan Markussen1, Maria K. Dahle1, Torstein Tengs1, Marie Lavoll1, Øystein W. Finstad3, Christer R. Wiik-Nielsen1, Søren Grove1, Silje Lauksund2, Børre Robertsen2, Espen Rimstad3*

1 Department of Laboratory Services, National Veterinary Institute, Oslo, Norway, 2 Norwegian College of Fishery Science, University of Tromsø, Tromsø, Norway, 3 Department of Food Safety and Infection Biology, Norwegian School of Veterinary Science, Oslo, Norway

Abstract

Piscine orthoreovirus (PRV) is associated with heart- and skeletal muscle inflammation (HSMI) of farmed Atlantic salmon (Salmo salar). We have performed detailed sequence analysis of the PRV genome with focus on putative encoded proteins, compared with prototype strains from mammalian (MRV T3D)- and avian orthoreoviruses (ARV-138), and aquareovirus (GCRV-873). Amino acid identities were low for most gene segments but detailed sequence analysis showed that many protein motifs or key amino acid residues known to be central to protein function are conserved for most PRV proteins. For M-class proteins this included a proline residue in μ2 which, for MRV, has been shown to play a key role in both the formation and structural organization of virus inclusion bodies, and affect interferon-β signaling and induction of myocarditis. Predicted structural similarities in the inner core-forming proteins λ1 and σ2 suggest a conserved core structure. In contrast, low amino acid identities in the predicted PRV surface proteins μ1, σ1 and σ3 suggested differences regarding cellular interactions between the reovirus genera. However, for σ1, amino acid residues central for MRV binding to sialic acids- and cleavage- and myristoylation sites in μ1 required for endosomal membrane penetration during infection are partially or wholly conserved in the homologous PRV proteins. In PRV σ3 the only conserved element found was a zinc finger motif. We provide evidence that the S1 segment encoding σ3 also encodes a 124 aa (p13) protein, which appears to be localized to intracellular Golgi-like structures. The S2 and L2 gene segments are also potentially polycistronic, predicted to encode a 71 aa- (p8) and a 98 aa (p11) protein, respectively. It is concluded that PRV has more properties in common with orthoreoviruses than with aquareoviruses.

Introduction

Piscine orthoreovirus (PRV) is associated with heart and skeletal muscle inflammation (HSMI) of farmed Atlantic salmon (Salmo salar) [1]. Infectious viral diseases are prevalent in farmed fish, and HSMI is an example of an emerging disease in the intensive farming of Atlantic salmon. HSMI is usually observed 5–9 months after transfer of the fish from the freshwater stage to seawater grow out areas [2] and is characterized by inflammation of the epi-, endo- and myocard and of the red skeletal muscle. The majority of the fish in an affected cage will show lesions in the heart and the cumulative mortality may reach 20% [2]. The virus has not successfully been cultivated continuously in cell cultures, although PRV harvested from two week cultures in GF-1 cells have been used for challenge experiments [3]. The PRV genome was mapped by high-throughput pyrosequencing of material from diseased fish and found to consist of 10 dsRNA segments [1]. By the use of real time RT-PCR it has been shown that PRV is widely distributed among both farmed and wild Atlantic salmon [1,4]. The sizes of the genomic segments are distributed in the classical orthoreoviral groups L1–3, M1–3 and S1–4. The S1 and S2 segments are possibly bicistronic having accessory small putative open reading frames.

The 3′-terminal nucleotide sequence (UCAUC-3′) in the PRV genome is conserved and identical to both orthoreoviruses and aquareoviruses [1,5,6]. On the other hand, the 5′-terminal nucleotide sequence (5′-GAUAAA/U) of PRV is unique, as are the analogue sequences for each of the individual species within the Orthoreovirus genus (MRV, ARV, Nelson Bay, Baboon and Reptilian orthoreoviruses) and of those species in the Aquareovirus genus (Aquareovirus A and C) for which the 5′-end sequences are known [5,7].

Based upon phylogenetic analysis a common evolutionary origin of the genera Aquareovirus and Orthoreovirus has been revealed [5]. Phylogenetic analyses performed separately for each PRV genus (Aquareovirus A and C) for which the 5′-end sequences are known [5,7].
segment has shown that this virus branches off the common root of the Aquareovirus and Orthoreovirus genera which could indicate that PRV may represent a genetic new lineage, divergent from other reovirus genera [1]. To our knowledge, aquareoviruses and orthoreoviruses, with an amino acid identity of 42% in the dsRNA-dependent RNA polymerase (RdRp), are the only reovirus genera with identity of more than 30% in the RdRp that are placed in separate genera [3]. The prototype strain mammalian reovirus type 3 Dearing (MRV T3D) was chosen for comparison as this strain is the best studied within the genus. Low amino acid homologies to the MRV proteins l1, l2, l3, µ1, µ2, µ3, σ2 and σ2NS are found in PRV as well as in the aquareoviruses (AqRV) [1]. AqRV have been isolated from a wide variety of aquatic animals, including molluscs, fish and crustaceans, while the orthoreoviruses have been found in reptiles, birds and mammals. There is low sequence homology of genes and proteins between the species in genus Orthoreovirus indicating a long time divergence [6]. AqRV have 11 genomic segments while both the orthoreoviruses and PRV have 10.

Viral particles of MRV and ARV consist of a double layered protein capsid with inner and outer layers. Studies of MRV indicate that the σ1 protein attaches to cell surface receptors and thus is important for the cell and tissue tropisms [8–10]. MRV particles enter the cell through receptor-mediated endocytosis. MAbs directed against the σ1 protein, the σ3 and µ1C outer capsid proteins, and the core spike protein λ2 can neutralize MRV [11].

Following endocytosis, the outer capsid undergoes proteolysis within the acidic compartment of the endosomes, resulting in the removal of σ3 and cleavage of µ1 to µ1C and µ1N [12,13]. The resulting intermediate subviral particles (ISVPs) penetrate the endosomal lipid bilayer, probably through the action of exposed hydrophobic parts of the cleaved µ1 protein [14–17]. This makes endosomal membrane penetration possible and is followed by cytoplasmic release of transcriptionally active viral cores [18,19]. Inside the cores, full-length capped but non-polyadenylated viral mRNAs are made. Cap formation requires the sequential activity of polynucleotide phosphorylase, guanylytransferase and methyltransferase [20]. The λ1 protein functions as helicase and triphosphatase, λ2 as the guanylytransferase and λ3 is the RNA polymerase [21–25]. S-adenosyl-L-methionine (SAM) is the substrate (methyl donor) for the methylation of the type 1 cap mediated by λ2 [24,26–29]. The transcripts act as templates for both translation and replication of viral genomic dsRNA [30].

The present study was performed to compare the properties of the putative PRV proteins to analogues of the orthoreoviruses MRV, ARV and the aquareovirus grass carp reovirus (GCRV). Through in silico analysis, 10 deduced amino acid sequences of PRV proteins were assigned. It was concluded that PRV is more related to the genus Orthoreovirus than to the Aquareovirus.

Results

Genome Organization

The PRV genome consists of 10 segments containing at least 10, but has possibly 13 ORFs or more. The genome has a length of 23320 nt and a GC content of 47%. For both orthoreoviruses and aquareoviruses the ultimate nucleotides from each end are inverted complements [5,31–34]. The length of the 5'-UTRs was shortest for the L segments, 7–18 nucleotides, while M3 and S4 had the longest with 84 and 38 nucleotides, respectively (Figure 1). The length of the 3'-UTRs varied between 44–89 nucleotides, the longest were found in the M segments (Figure 1). Segment comparison showed that PRV 5'-UTRs were on average significantly shorter than the 3'-UTRs. In PRV, the three ultimate nucleotides in 5’- and 3’-end of each segment are inverted complements. RNA secondary structure predictions using mFold version 2.3 [35] performed at 15°C of the 5’- and 3’-UTRs of mRNA from each genomic segment was assayed using energy minimization criteria. The predictions were panhandle structures, but the last 3’-end nucleotides were not a part of the stem structures. For segment S2, however, the prediction was not a panhandle structure (data not shown).

The PRV genome segments except for S1 and S4 were assigned according to the assignment used for MRV (Table 1). Hence, for example, the PRV segment that encodes the core protein λ2 is called L2, although it is slightly longer than the other L segments. Based upon sequence homology to MRV and ARV in particular, eight of the deduced translation products are assumed structural proteins. For PRV, the S1, S2 and possibly L2 gene segments have additional internal open reading frames (ORFs) in addition to the σ3, σ2 and λ2-encoding ORFs (Table 1, Figure 1).

L-class Gene Segments

PRV gene segment L1 is predicted to encode the λ3 protein (Table 1, Figure 1). This is the virus’ RNA-dependent RNA polymerase (RdRp) responsible for viral transcription and replication. It displays the highest amino acid sequence similarity to MRV, ARV and GCRV (Table 2). Multiple sequence alignment of PRV λ3 with the corresponding protein sequences from MRV, ARV and GCRV revealed that important polymerase motifs are also conserved in the PRV protein (Figure S1). Comparing with detailed structure-function analysis made of MRV T3D λ3 [36], we found high conservation particularly in the catalytic core (aa 350–900), and of specific amino acids predicted to be responsible for interaction with the RNA template, NTP and the 5’-cap. The previously described catalytic motifs I, II, III and F1–F3 could easily be identified [37–40], including the universal RdRp GDD motif (within motif III) involved in transcriptional initiation [41–43]. The N- and C-terminal domains of the polymerase, predicted to be involved in interaction with the viral protein λ2 and RNA helicase λ1 for MRV, are somewhat less conserved. This is line with a lower sequence conservation of the other PRV λ proteins.

PRV gene segment L2 encodes the λ2 protein (Table 1, Figure 1). In MRV and ARV (λC) this is the capping enzyme, the main contributor when generating the 5'-terminal cap, on virally encoded mRNAs [24,26–29]. The MRV- and ARV-proteins contain both the guanylytransferase and methyltransferase activities necessary for the generation of this type 1 cap structure [24,27,29,44,45]. For PRV, multiple sequence alignment reveals relatively low amino acid identities to the corresponding proteins in MRV, ARV and GCRV, although key amino acid residues and important functional domains are highly conserved (Table 2, Figure S2). These include residues K190 which is essential for autoguanylation, H223 and H292 which are essential for guanylytransferase activity and the S-adenosyl-L-methionine (SAM) binding pocket [44–47]. K171, a significant but not essential contributor for autoguanylation [46], is not conserved in PRV at this position although a KY motif sits just two positions downstream. A region containing an ATP/GTP binding site in the MRV protein [45] is also highly conserved in all four homologues proteins. Also, a hypersensitive cleavage site has been identified when using recombinant MRV λ2 and ARV λC [44,46]. In the PRV protein, the two amino acids on either side of this cleavage site are identical to those in ARV. Secondary structure predictions using PSIPRED v3.0 [48,49] suggests that this site resides in a predicted random coil region shortly after a predicted strand region, perhaps constituting an exposed region.
lysines are conserved in all four homologous proteins (been shown to be essential for NTPase activity [63]. Both these binding region was predicted using BindN [59] (data not shown).

...residues or domains are conserved in PRV but lower than for MRV, the homologous protein in MRV, ARV and GCRV proteins and shown to bind zinc in core crystal structure analysis [51,53,54], is also conserved in PRV. A CCHH zinc-finger motif in particular, conserved in the MRV, ARV and GCRV proteins and shown to bind zinc in core crystal structure analysis [51,53,54], is also conserved in PRV (Figure S3). For PRV, as for the homologous proteins from the selected reovirus prototypes, a predominantly N-terminal RNA/DNA binding region was predicted using BindN [59] (data not shown).

M-class Gene Segments

PRV gene segment L3 is predicted to encode the λ1 protein (Table 1, Figure 1). This is the core capsid shell homologue of MRV λ1, ARV λA and GCRV VP3 [50–52]. Multiple sequence alignment showed higher amino acid identities between PRV and the three prototype strains than observed for λ2 but lower than for λ3 (Table 2, Figure S3). Key functional residues or domains are conserved in PRV λ1. A CCHH zinc-finger motif in particular, conserved in the MRV, ARV and GCRV proteins and shown to bind zinc in core crystal structure analysis [51,53,54], is also conserved in PRV λ1 (Figure S3). For PRV, the ~500 residue N-terminal region has been shown to exhibit triphosphate phosphohydrolase (NTPase), helicase, and RNA triphosphatase activities while the ~200 residue N-terminal parts, a region predominantly hydrophilic, is assumed to be responsible for the dsRNA binding property of λ1 [21,23,52,55–58]. For PRV, as for the homologous proteins from the selected reovirus prototypes, a predominantly N-terminal RNA/DNA binding region was predicted using BindN [59] (data not shown).

A second putative nucleotide binding motif, D446EVG449 [65], is only partially conserved. The P208 in MRV T3D has also recently been shown to be a determinant of type I IFN antagonism and a modulator of induction of myocarditis in neonatal mice [69]. P208 is conserved in all four proteins.

PRV gene segment M2 encodes the μ1 protein (Table 1, Figure 1). This is the homologue of the major outer capsid protein of MRV [54,70,71]. In ARV and GCRV the homologous proteins are called μB and VP4, respectively [5,50]. The highest amino acid sequence identity was found towards the homologous proteins in MRV and ARV, somewhat lower for GCRV (Table 2). Multiple sequence alignment showed that the N-terminal sequence parts of the four proteins display higher conservation compared to the rest of the protein (Figure S5). A post-translational autolytic cleavage site in the MRV protein between N42 and P31, which produces a small N-terminal- (μ1N) and a larger C-terminal fragment (μ1C), required for MRV endosomal membrane penetration and infection, is conserved in all four proteins [70,71]. The cleavage of MRV μ1 during infection is dependent upon N-myristoylation at G2, as well as binding to the S1 gene product σ3 (see below) [72]. Similarly, ARV μB is also myristoylated at G2 and post-translationally cleaved, with μB and its cleavage product μBC associating with eB [68,73]. The G2-residue is conserved in all four proteins (Figure S5). These two processes, myristoylation and cleavage, are believed to be crucial for membrane penetration with μ1N as the principal mediator [17,74,75] (see Figure S5 legends for more details).

PRV gene segment M3 encodes the non-structural μNS protein (Table 1, Figure 1). This is the homologous counterpart of MRV and ARV μNS [50,76,77]. The ARV μNS protein is significantly shorter compared to the MRV protein, a difference which has been attributed to deletions [Figure S6] [78]. In GCRV, the homologous protein is NS80, the product of gene segment 4 [5]. A very low amino acid sequence identity is observed.
<table>
<thead>
<tr>
<th>Segment</th>
<th>Protein name</th>
<th>Length (aa)</th>
<th>Theoretical weight (kDa)</th>
<th>Predicted location in virion and functional properties</th>
<th>PRV encoded proteins</th>
<th>MRV T3D&lt;sup&gt;b&lt;/sup&gt;</th>
<th>ARV-138</th>
<th>GCRV-873&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>λ3</td>
<td>1286</td>
<td>144.2</td>
<td>RNA-dependent RNA polymerase</td>
<td>λ3 (1267)</td>
<td>λB (L2) (1259)</td>
<td>VP2 (52) (1274)</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>λ2, p11?</td>
<td>1290</td>
<td>143.7</td>
<td>Guanylyltransferase, methyltransferase</td>
<td>λ2 (1289)</td>
<td>λC (L3) (1285)</td>
<td>VP1 (51) (1299)</td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>λ1</td>
<td>1282</td>
<td>141.5</td>
<td>Helicase, NTPase, RNA triphosphatase</td>
<td>λ1 (1275)</td>
<td>λA (L1) (1293)</td>
<td>VP3 (53) (1214)</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>μ2</td>
<td>760</td>
<td>86.0</td>
<td>NTPase, RNA triphosphatase, RNA binding</td>
<td>μ2 (736)</td>
<td>μA (M1) (712)</td>
<td>VP5 (55) (728)</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>μ1</td>
<td>687</td>
<td>74.2</td>
<td>Outer capsid protein, membrane penetration</td>
<td>μ1 (708)</td>
<td>μB (M2) (676)</td>
<td>VP4 (56) (648)</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>μNS</td>
<td>752</td>
<td>83.5</td>
<td>Non-structural protein, central in virus inclusion formation</td>
<td>μNS (721)</td>
<td>μNS (M3) (635)</td>
<td>NS80 (54) (742)</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>σ3, p13</td>
<td>330, 124</td>
<td>370, 130</td>
<td>σ3: outer capsid protein, zinc metalloprotein</td>
<td>σ3 (54) (365)</td>
<td>σ8 (53) (367)</td>
<td>VP7 (510) (276),</td>
<td>NS26 (511) (244)</td>
</tr>
<tr>
<td>S2</td>
<td>σ2, p8?</td>
<td>420, 71</td>
<td>459, 8.1</td>
<td>σ2: Inner capsid protein, RNA binding</td>
<td>σ2 (52) (418)</td>
<td>σA (52) (416)</td>
<td>VP6 (58) (412)</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>μNS</td>
<td>354</td>
<td>39.1</td>
<td>Non-structural protein, involved in virus inclusion formation</td>
<td>μNS (53) (366)</td>
<td>μNS (54) (367)</td>
<td>NS38 (59) (352)</td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>σ1</td>
<td>315</td>
<td>346</td>
<td>Cell attachment protein</td>
<td>σ1 (51) (455)</td>
<td>C (51) (326)</td>
<td>NS31 (57) (274)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>σ1s</td>
<td>315</td>
<td>346</td>
<td>Cell attachment protein</td>
<td>σ1s (51) (120)</td>
<td>σ10 (51) (98)</td>
<td>NS16 (57) (146)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Proteins encoded by the PRV genome and functional properties as predicted from comparative studies with selected reovirus prototype strains.

<sup>a</sup>L1-M3 PRV gene segments are annotated according to mammalian reoviruses (MRV). PRV L1 has been changed to L3, and vice versa, compared to that suggested by [1]. PRV S-class gene segments are annotated according to [1].

For mammalian reovirus (MRV), avian orthoreovirus (ARV) and grass carp reovirus (GCRV) several proteins are produced from alternative reading frames or by post-translational proteolytic cleavage. In the latter case, if the exact cleavage site is known, the lengths of both proteolytic fragments are included in the table.

<sup>b</sup>T3D = Type 3 Dearing strain.

<sup>c</sup>GCRV contains an eleventh genomic segment which encodes a non-structural protein, NS26. VP7 is homologues to σ3/σB.

<sup>d</sup>Cytotoxic, nonfusogenic integral membrane protein [96].

doi:10.1371/journal.pone.0070075.t001
S-class Gene Segments

**PRV gene segment S1 encodes the major outer capsid protein σ3** (Table 1, Figure 1). In MRV, ARV and GCRV, the major outer capsid proteins are encoded by gene segment S4 (σ3), S3 (σB) and segment 10 (VP7), respectively [5,76,77,94,95]. PRV σ3 was recently determined to be encoded by the second smallest S-class gene segment [96]. Although amino acid identities between the PRV protein and that of the reovirus prototype strains used are very low and in the range of non-related proteins (Table 2), a Zn-finger motif is evolutionary conserved in all four proteins (Figure S7) [96,97]. Expression of σ3 in mammalian VERO cells at 37°C and salmonid CHSE cells at 20°C demonstrated that the protein was primarily cytoplasmic. In CHSE cells, green fluorescence was observed diffusely throughout the cytoplasm while in VERO cells the protein seemed to form large inclusions (Figures 2a, b).

**PRV gene segment S2 encodes the core clamp protein σ2** (Table 1, Figure 1). In MRV, ARV and GCRV, the homologues proteins are σ2, σA and VP6, respectively [5,54,76,77,98,99]. PRV S2 is the largest S-class gene segment, and possibly bicistronic (Table 1, Figure 1). Multiple sequence alignment of the four σ2/σA/VP6 proteins show overall low amino acid identities (Table 2, Figure S8). Still, between MRV cores and ARV σA (sharing 29% amino acid identity), comparisons of crystal structure data have shown a highly similar overall topology, including higher α-helical content in their C-terminal regions [54,100–102]. For PRV σ2, PSIPRED [48,49] also predicts a high α-helical content in its C-terminal region (Figure 3). In fact, comparing the predicted secondary structure profiles between all four reovirus proteins reveals a remarkable conservation of secondary structure (Figure 3), providing strong support that this gene segment encodes σ2. In further support for the correct annotation of PRV σ2, the predicted pI of the PRV protein is close to that of the MRV- and ARV proteins [96].

**PRV gene segment S3 encodes the non-structural σNS protein (Table 1, Figure 1).** In MRV and ARV, σNS is encoded by gene segment S3 and S4 respectively, while in GCRV the homologues protein, NS38, is encoded by segment 9 [5,50,103]. Similar to the three prototype strains, PRV S3 also contains a single ORF, encoding a protein with size close to that of σNS/NS38 (Table 1, Figure 1). Multiple sequence alignment with these four protein sequences reveals overall very low amino acid sequence identities, not only to the PRV protein but between the three prototype strains as well (Table 2, Figure S9). In addition, PSIPRED suggests some level of protein structure conservation between PRV σNS and the proteins from the three prototype strains (not shown). Furthermore, the predicted pI of the putative PRV σNS protein shows generally more closeness to the homologues proteins in MRV, ARV and GCRV compared to the other three major PRV S-class proteins [96]. Taken together this suggests that a correct assignment of PRV σNS has been made.

MRV σNS has been detected in both the nucleus and cytoplasm of infected and transfected cells, with the former being linked to its nucleic acid binding capability [83,103,104]. PSORTII does not predict the presence of nuclear localization signals in the PRV protein (not shown). But, within regions of the alignment displaying somewhat higher level of conservation, NetNES 1.1 [105] does predict the presence of NESs in all four proteins (Figure S9).

The major gene product of PRV gene segment S4 is the σ1 cell-attachment protein (Table 1, Figure 1). In MRV and ARV, the cell attachment proteins σ1 and σC are the major gene products from the bicistronic and tricistronic S1 gene segments, respectively [10,106–109] (Table 1). The aquareovirus GCRV does not seem to encode a homologue of the orthoreovirus cell attachment proteins [99,110]. Rather, GCRV gene segment 7 encodes two non-structural proteins (Table 1) [5].

Multiple sequence alignment with PRV σ1 and MRV σ1/ARV σC shows amino acid identities in the range of 14–21%,

---

### Table 2. Percentage amino acid identity among all ungapped positions between pairs; predicted PRV proteins and the homologues proteins from three reovirus prototype strains.

<table>
<thead>
<tr>
<th></th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
<th>PRV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ι3</td>
<td>ι2</td>
<td>ι1</td>
<td>μ2</td>
<td>μ1</td>
<td>μNS</td>
<td>ι3</td>
<td>ι2</td>
<td>ι1</td>
</tr>
<tr>
<td>MRV T3D*</td>
<td>43</td>
<td>23</td>
<td>32</td>
<td>20</td>
<td>28</td>
<td>17</td>
<td>13</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>ARV-138a</td>
<td>44</td>
<td>22</td>
<td>31</td>
<td>21</td>
<td>27</td>
<td>17</td>
<td>15</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>GCRV-873*</td>
<td>38</td>
<td>25</td>
<td>31</td>
<td>21</td>
<td>24</td>
<td>13</td>
<td>15</td>
<td>17</td>
<td>13</td>
</tr>
</tbody>
</table>

*Ref. Table 1 for gene segment annotations and names of homologues proteins in MRV, ARV and GCRV. Identity values are from separate pairwise alignments of the protein sequences.

*Value from a manually adjusted pairwise alignment of the two proteins.

GCRV does not appear to have a cell attachment protein homologue to σ1/σC.

---

* doi:10.1371/journal.pone.0070075.t002

---

Table 2. Percentage amino acid identity among all ungapped positions between pairs; predicted PRV proteins and the homologues proteins from three reovirus prototype strains.
depending on the degree of manual adjustment of the alignment (Table 2, Figure 4; only PRV and MRV aligned). Annotation of PRV σ1, based on the high probability presence of predicted α-helical coiled-coil structure(s) in the N-terminal region of the protein, was recently determined to be similar to that of MRV σ1 and ARV σC (Figure S10) [96,111–114]. Secondary structure predictions using PSIPRED provides additional support for a PRV σ1 N-terminal region dominated by α-helices (not shown). For the GCRV NS31 protein, on the other hand, predicted secondary structure profiles are very different from that of the three other proteins (not shown), with no predicted coil structures (Figure S10).

The MRV σ1 protein can broadly be divided into three distinct domains, the N-terminal tail which partially inserts into the virion, the body which contains the region that binds to sialic acids, and the C-terminal head domain, which binds the receptor junctional adhesion-molecule-A (JAM-A) [115–119]. For the MRV T3D strain, σ1 binds to α-linked sialic acids [120,121]. Sequence- and structure analyses of MRV variants have indicated that amino acids in positions 198, 202–205 are involved in binding to sialic acids [122–124]. The alignment in Figure 4 suggests that several of these MRV σ1 residues may be conserved in PRV, but not in the ARV protein (not shown). In fact, these residues are less conserved in the MRV T1L and T2J serotypes compared to PRV σ1 (not shown). Finally, the predicted isoelectric point (pI) for this putative PRV protein lies in the acidic range, as do MRV σ1 and ARV σC [96]. All together, we provide additional support that this protein, encoded by the smallest PRV S-class gene segment, is the cell attachment protein homologous to those of other orthoreoviruses.

Accessory ORFs

PRV gene segment S1 is bicistronic and has an internal ORF encoding a 124 aa protein (p13) (Table 1, Figure 1). This protein is expressed from the σ3 ORF when transfected in VERO and CHSE cells, as determined by immunofluorescent staining using anti-p13 serum, where it colocalizes with the trans-Golgi marker WGA (Figures 2a, b). Colocalization of p13 with this marker was also seen following transfection with an expression plasmid containing only the p13 ORF (Figures 2a, b).
protein was originally assumed to be a fusion-associated small transmembrane (FAST) protein capable of inducing cell-cell fusion and syncytium formation, but was recently determined to encode a nonfusogenic integral membrane protein with cytotoxic properties [1,96]. Scanning the PRV genome for potential accessory ORFs revealed several potential candidates. Both S2 and L2 are putative polycistronic gene segments, where the former has the potential to encode a second 71 aa hypothetical protein (p8) and the latter contains several smaller ORFs with AUG- or non-canonical GUG/CUG start codons that could encode hypothetical proteins of sizes ranging from 55 to 135 aa. The possibility of translation initiation from non-canonical start codons should not be excluded, as has been shown for the AtSRV p22 protein which is produced from a noncanonical CUG translation start codon in segment 7 [125]. Besides the L2-encoding ORF, the largest L2 ORF containing an AUG translation initiation codon could encode a small protein of 98 aa (p11) (Table 1, Figure 1). None of these putative gene products display any apparent sequence similarities.

Figure 3. PSIPRED secondary structure predictions of PRV α2 (green) and the homologous proteins from the reovirus prototype strains MRV T3D (α2, pink), ARV-138 (αA, blue) and GCRV-873 (VP6, brown).

doi:10.1371/journal.pone.0070075.g003
towards the additional gene products from MRV S1 (σ1S), ARV S1 (p17) or GCRV segment 7 (NS16) (not shown). ARV p10 and GCRV NS16 have been shown to be FAST proteins capable of inducing cell-cell fusion and syncytium formation [126,127]. No sequence similarities to other known reovirus FAST proteins, such as Atlantic salmon reovirus (ASRV p22), Turbot reovirus (SMRV p22), Baboon reovirus (BRV p15), Reptilian reovirus (RRV p14), Broome virus (BroV p13) and Nelson Bay reovirus (NBV p10) was observed either (not shown) [7,125,127–130]. Low sequence similarities to other known reovirus FAST proteins, such as Atlantic salmon reovirus (ASRV p22), Turbot reovirus (SMRV p22), Baboon reovirus (BRV p15), Reptilian reovirus (RRV p14), Broome virus (BroV p13) and Nelson Bay reovirus (NBV p10) was observed either (not shown) [7,125,127–130]. Low sequence similarities to other known reovirus FAST proteins, such as Atlantic salmon reovirus (ASRV p22), Turbot reovirus (SMRV p22), Baboon reovirus (BRV p15), Reptilian reovirus (RRV p14), Broome virus (BroV p13) and Nelson Bay reovirus (NBV p10) was observed either (not shown) [7,125,127–130].

Discussion

In the present study we have performed detailed comparative sequence analysis of the non-translated (UTR) regions and the putative proteins encoded by the PRV genome to prototype strains from mammalian and avian orthoreoviruses, and one aquareovirus. The results suggest that the PRV genome encodes at least 10 proteins, but it may also contain up to 13 or more ORFs.

In general, amino acid identities between PRV and the three prototype strains were low for most gene segments, highest for the L-class gene segment encoding the RdRp. Functional constraints often cause viral core proteins to remain more conserved than the outer capsid proteins, as illustrated by the conserved structural motifs predicted for PRV proteins σ1 and σ3. In the latter case, a remarkable strong conservation of protein secondary structure across genus lines of orthoreoviruses and aquareoviruses, was predicted. This illustrates conservation of secondary structure over primary sequence, reflecting the importance of structural features for function(s), where even minor alterations would be deleterious. Furthermore, residues important for the enzymatic actions of reovirus proteins are to a larger extent more conserved than other residues, as exemplified by the PRV σ3 and σ2 proteins. Examples of conserved amino acid residues include the two lysines essential for the NTPase activity of µ2 [63] and the post-translational autolytic cleavage- and myristoylation sites in µ1 [68,70,71,73]. The latter strongly suggests that analogous mechanisms are involved in membrane penetration by this PRV protein as in orthoreo- and aquareoviruses. The amino acid identities between PRV and the three prototype strains were low for most gene segments, highest for the L-class gene segment encoding the RdRp. Functional constraints often cause viral core proteins to remain more conserved than the outer capsid proteins, as illustrated by the conserved structural motifs predicted for PRV proteins σ1 and σ3. In the latter case, a remarkable strong conservation of protein secondary structure across genus lines of orthoreoviruses and aquareoviruses, was predicted. This illustrates conservation of secondary structure over primary sequence, reflecting the importance of structural features for function(s), where even minor alterations would be deleterious. Furthermore, residues important for the enzymatic actions of reovirus proteins are to a larger extent more conserved than other residues, as exemplified by the PRV σ3 and σ2 proteins. Examples of conserved amino acid residues include the two lysines essential for the NTPase activity of µ2 [63] and the post-translational autolytic cleavage- and myristoylation sites in µ1 [68,70,71,73]. The latter strongly suggests that analogous mechanisms are involved in membrane penetration by this PRV protein as in orthoreo- and aquareoviruses. The amino acid identities between PRV and the three prototype strains for the µNS, σ1 and σ3 proteins were very low or similar to that of non-related proteins (<20%) [31]. Sequence analysis did, though, reveal protein motifs or key amino acid residues linked to important protein functions are conserved also in PRV, particularly to those of MRV and ARV.

Figure 4. Multiple sequence alignment of PRV σ1 with MRV T3D σ1. Black lines represent putative nuclear export signals (NEP) in MRV and PRV, respectively, as predicted by NetNes 1.1. ■ = L149 in the MRV protein involved in a second predicted NES. ▲ = residues in the MRV protein involved in binding to sialic acid residues. The alignment has been manually adjusted.

doi:10.1371/journal.pone.0070075.g004
For the μNS proteins (NS80 in GCRV), the limited sequence similarities may be linked to its multifunctionality. Together with σNS (NS36 in GCRV) they are central in recruitment of core proteins and forming viral factories, sequence independent binding of ssRNA, and viral in RNA packaging and replication [66,67,68,80–87,103,132]. Virus proteins that make up the core have evolved over time due to selection pressures and thus the resulting rate of change in multifunctional proteins would be higher by comparison, resulting in only rudimentary similarities being left at the primary sequence level, although important structural features may be more conserved. Whether the PRV μNS gene also produces two gene products as the MRV protein does, or whether post-translational cleavage is involved as in ARV μNS, could not be determined from sequence information alone.

Several proteins encoded by reoviruses have been shown, or are suggested, to exhibit type I IFN antagonistic properties, including MRV μ2, ARV σA and MRV σ3 [133–137]. A previously established reporter gene system [138] was used to investigate whether expression of PRV μ2, σ2 and σ3 could have an antagonistic effect on the type I IFN response in a salmonid cell line (not shown). The reporter constructs use a salmon minimal type I IFN (IFNα1) promoter [139] and an interferon-stimulated response element (ISRE)-reporter (Agilent technologies). Transfection method and activation of type I IFN response was performed as previously described [138,139]. No effect could be detected by any of the three viral proteins, neither on IFNα1- nor ISRE induction.

The monocistronic S4 gene segment was determined to encode the PRV cell attachment protein σ1, the finding supported by predicted p1’s but more importantly, by structural motifs common to both MRV σ1 and ARV σC, as was recently also described by others [96]. Several amino acid residues in MRV σ1 shown to bind to sialic acids may also be conserved in the PRV protein, which might suggest that PRV utilizes a receptor mediated uptake mechanism involving similar but not identical sialic acid structures.

PRV σ3 expression in mammalian VERO cells at 37°C and in salmonid CHSE cells at 20°C showed a predominantly cytoplasmic localization, although the staining pattern differed in the two cell lines. The punctuated staining pattern observed in the VERO cells may be linked to aberrant folding of the protein at the higher temperature.

FAST proteins have been reported from the Orthoreovirus and Aquareovirus genera [131]. In contrast to fusogenic reoviruses, for which syncytia are commonly registered, syncytia are not common histopathological findings in HSMM diseased fish [2]. Sequence data alone was not sufficient to determine whether the putative p8, p11 or p13 could be a FAST protein. All three proteins contained predicted properties or motifs only partially consistent with FAST proteins. Recently, using an avian cell line cultivated at 37°C, it was shown that p13 is a cytotoxic protein binding to intracellular membranes, and not a FAST protein [96]. Our findings supported this, and we show that p13 was produced from the internal ORF in the σ3 coding sequence, in both VERO- and the CHSE cells, where it colocalizes with a marker for the trans-Golgi network. The colocalization of p13 and this marker was further supported by the transfection assays using an expression plasmid construct containing the p13 ORF only. Transfection studies with expression plasmid constructs containing the p13 ORF performed in both VERO and CHSE cells have failed to induce cell-cell fusion (data not shown). Transfection with expression plasmid constructs containing the putative p8 and p11 ORFs were not performed. The absence of observed fusogenic activity of PRV and lack of induced syncytia after cellular expression of p13 may indicate that PRV is not fusogenic, as also recently described for p13 [96]. It should not be excluded though that efficient cell-cell fusion is dependent upon a second coexpressed PRV protein, as seen for the GCRV NS16 whose activity is enhanced following coexpression with NS26 [126]. Also, it should be considered whether p13 is a functional equivalent to the MRV σ1s or ARV p17 proteins.

For members of the Reoviridae more than 30% amino acid sequence identity of the RdRp is used as indicative of genus affiliation. But there are exceptions to this, the Rotavirus B polymerase is only 22% identical to other rotaviruses [5]. Sequence alignment of the PRV L3 protein showed several conserved polymerase motifs, and the identities to the RdRps of the orthoreoviruses MRV and ARV, and the aquareovirus GCRV, were well above the 30% limit. However, the amino acid sequence identities between the RdRps from the orthoreoviruses MRV, ARV and the aquareovirus GCRV is above 40% (not shown), which disqualifies the use of this quantitative taxonomic criterion to distinguish between these genera. However, there are differences between orthoreoviruses and aquareoviruses that justify to keep them as separate genera, as listed by Attoui and co-workers [5], like distinct ecologies, 10 versus 11 segments, the GC-content of orthoreoviruses is 44–48% while that of aquareoviruses is 52–60%, many orthoreoviruses do not induce syncytia in contrast to the majority of known aquareoviruses, and there is no antigenic relationship between them. Of these criterions PRV has the following in common with the orthoreoviruses: 10 dsRNA segments, 47% GC-content, syncytia is not reported as a common histopathological finding, nor in cell culture where virus isolation has been attempted (personal observation). The only common criterion with aquareoviruses is the ecocide, if fish consisting of a large number of heterogenous species from very different environments should be regarded as a single ecocide. The antigenic relationships are unknown.

PRV is not the first reovirus with 10 genomic segments that has been described from fish. A virus isolated in Thailand from the striped snakehead fish (Ophicephalus striatus, also known as Channa striata) with epizootic ulcerative syndrome, was also found to contain 10 genomic segments [146]. However, no nucleotide sequences are available for this virus. The lack of serological cross-neutralization activity to other AqRV, and the difference in number of gene segments made the authors conclude that the virus was not a member of the Aquareovirus genus.

To conclude, although it probably is many million years since the most recent common ancestor for PRV and orthoreo- and aquareoviruses existed, we found conserved structural motifs and somewhat less conserved sequence motifs for all 10 PRV genomic segments. All together, the PRV has more properties in common with the Orthoreovirus genus than with the Aquareovirus and should hence be renamed Piscine orthoreovirus.

Materials and Methods

Computer Analyses

GenBank accession numbers for all PRV, MRV T3D, ARV-138 and GCRV-873 sequences used in the present study is shown.
in Table S1. Multiple sequence alignments of protein sequences were performed in AlignX (Vector NTI Advance™ 11 Package, Invitrogen Dynal AS). RNA secondary structure predictions were performed using mFold version 2.3 with the eighty ultimate 5′- and 3′- nucleotides of the mRNAs from each segment as input sequence
(http://mfold.rna.albany.edu/). Default parameters were used in the predictions except for temperature, which was set to 15°C. PSIPRED v3.0 was used for predictions of protein secondary structures (http://bioinf.cs.ucl.ac.uk/psipred/). BindN was used to predict putative RNA binding properties of the proteins (http://bioinfo.ogp.org/bindn/) [59]. The NetNES 1.1 server was used to predict putative leucine-rich nuclear export signals (NESs) in the proteins (http://www.cbs.dtu.dk/services/NetNES/) [105]. Multicoil and COILS were used to predict the presence of putative coiled coil regions in proteins (http://groups.csail.mit.edu/ch/multicoil/cgi-bin/multicoil.cgi, http://embnet.vital-it.ch/software/COILS_form.html) [147,148]. Default settings were used in the predictions except for Multicoil where window size of 21 was used. Prediction of theoretical molecular weights and isolectric points (pIs) for putative PRV proteins was performed using the Compute pI/Mw tool available at http://web.expasy.org/compute_pi/. The presence of putative nuclear localization signals (NLS) in PRV proteins was investigated using PSORTII available at http://psort.hgc.jp/form2.html. Prediction of N-terminal myristoylation was performed using the tools available at http://web.expasy.org/myristoylator/ and http://mendel.imp.ac.at/myrystate/SUPPl predictor.htm. ProtScale, available at http://web.expasy.org/protscale with the algorithm by Kyte and Doolittle [149], was used to generate hydropophicity plots for hypothetical PRV proteins encoded by accessory ORFs and the FAST proteins from ARV-138 and GCRV-873. Prediction of S-palmitoylation sites was performed using CSS-Palm 3.0 using the universally conserved GDD domain (in Motif III) boxed. (TIF)

Supporting Information

Figure S1 Multiple sequence alignment of PRV L1 ORF (1.3) with corresponding ORFs from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. RNA-dependent RNA polymerase (RdRp) domains are indicated with the universally conserved GDD domain (in Motif III) boxed. (TIF)

Figure S2 Multiple sequence alignment of PRV L2 ORF (1.2) with guanylyltransferases from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. * = lysine residues in MRV essential (K171) or significant contributor (K171) for autoguanylation in the MRV, ARV and GCRV proteins, ▲ = conserved histidines essential for guanylyltransferase activity in the MRV protein, ■ = ATP/GTP binding site motif A in ARV, boxed = Sadenosyl-L-methionine (SAM) binding pocket, ● = ATP/GTP binding site motif A in MRV, and ↓ = hypersensitive cleavage site in recombinant MRV λ2 and ARV 1733 λC. (TIF)

Figure S3 Multiple sequence alignment of PRV L3 ORF (1.1) with the helicase-NTPase/core capsid shell proteins from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. * = conserved CCHH zinc-finger motif. (TIF)

Figure S4 Multiple sequence alignment of PRV M1 ORF encoding the µ2 protein with the homologues proteins from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. ▲ = conserved proline residue suggested to play a key role in the formation and structural organisation of reovirus inclusion bodies, a determinant of type I IFN antagonism and a modulator of myocarditis in neonatal mice. ● = leucine vs. phenylalanine, a determinant of tissue tropism of MRV µ2 in MDCK cells. ■ = possible NLS in MRV. Red lines = nucleotide binding/triphosphate phosphohydrolase regions, and * = conserved lysine residues essential for ATPase activity in ARV µA. A nuclear export signal (NES) has been predicted for MRV µ2 (residues 328–335) [151]. NetNES 1.1 predicts a NES in ARV µA in the same region, while in GCRV, L238, L238 and L1697 are predicted to participate in a NES, and in PRV L50 (numbering
according to the GCRV and PRV sequences, respectively) (not shown).

(TIF)

Figure S5 Multiple sequence alignment of PRV M2 ORF encoding the m1/m2 major outer capsid protein with the homologues proteins from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. * = myristoylation site in the MRV protein. ↓ = post-translational cleavage site producing N- and C-terminal fragment m1N and m1C (MRV) or mBN and mBC (ARV). The C-terminal end of the MRV protein is extended by 33 amino acids compared to the homologous proteins in ARV and GCRV. The PRV protein is also extended, by 28 amino acids.

(TIF)

Figure S6 Multiple sequence alignment of PRV M3 ORF encoding the putative mNS protein aligned with mNS/mNS80 proteins from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. ↓ = N-terminal end of second translation product of MRV (mNSC). Met-57, conserved in MRV and PRV is boxed red. * = conserved putative zinc-hook motif crucial in the formation of inclusion-like structures in the MRV protein [86,90,152]. Black lines indicate sequence regions with higher level of conservation with the motif XGXDPX being boxed. In ARV, the larger region forms part of a region that has been shown to be involved in inclusion maturation [86]. Grey solid and dotted lines = coil-coil(s) regions as predicted by MultiCoil (window size: 21, probability cutoff: 0,5). The MRV L711IDFS715 motif shown to be required for the recruitment of clathrin to viral factories is boxed red.

(TIF)

Figure S7 Multiple sequence alignment of PRV S1 ORF encoding the major outer capsid σ3 protein with the homologues proteins from the reovirus prototype strains MRV T3D, ARV-138 and GCRV-873. * = conserved Zn-finger motif.

(TIF)

Figure S8 Multiple sequence alignment of PRV S2 ORF encoding the inner capsid σ2 protein with the homologues proteins from the reovirus prototype strains MRV T3D σ2, ARV σ6 and GCRV-873 VP6. * = R275, one of two arginines in ARV σA linked to dsRNA binding and nucleolar localization, conserved in fusogenic orthoreoviruses.

(TIF)

Acknowledgments
We thank Ingvild Berg Nyman for assistance in subcloning work and Even Thoen for assistance with confocal microscopy.

Author Contributions
Conceived and designed the experiments: TM MKD TT ML ØWF SL ER. Performed the experiments: TM MKD TT ML ØWF CRW Thoen for assistance with confocal microscopy.

Sequence Analysis of the Piscine Reovirus Genome

References


