Segment 8 encodes a structural protein of infectious salmon anaemia virus (ISAV); the co-linear transcript from Segment 7 probably encodes a non-structural or minor structural protein

Eirik Biering[1,*], Knut Falk[3], Eirik Hoel[1,2], Jegatheswaran Thevarajan[1], Maaike Joerink[1], Are Nylund[2], Curt Endresen[2], Bjørn Krossøy[1,2]

[1]Intervet Norbio, Thormøhlensgt. 55, 5008 Bergen, Norway
[2]Department of Fisheries and Marine Biology, University of Bergen, 5020 Bergen, Norway
[3]National Veterinary Institute, 0033 Oslo, Norway

ABSTRACT: In this study we present the cloning, expression and partial identification of Genomic Segment 7 of infectious salmon anaemia virus (ISAV). The nucleotide sequence corresponding to Segment 7 was isolated from a bacteriophage lambda cDNA library and contained 2 overlapping open reading frames (ORFs) of 903 and 522 bases respectively. It also contained an ISAV-specific conserved nucleotide motif in the mRNA 5′ region. The co-linear transcript representing the large ORF undergoes a splicing event that removes a 526 nucleotide intron to form a mRNA corresponding to the smaller reading frame. Thus, ISAV Genomic Segment 7 has a similar coding strategy as influenza A virus Segments 7 and 8. The largest ORF of Segment 7 and the first ORF of Segment 8 was expressed in *E. coli* as fusion proteins and rabbit antiserum was raised against the recombinant protein from Segment 8. Immunoblot studies using this antiserum and a serum against purified virus, show that Segment 8 encodes one of the major structural proteins of the virus whereas the co-linear ORF of Segment 7 probably encodes a non- or minor structural protein

KEY WORDS: ISAV · Splicing · Matrix · Segment 7 · Segment 8 · GenBank Accession No. AY044132

INTRODUCTION

Infectious salmon anaemia virus (ISAV) is the causative agent of an economically important disease of farmed salmonid fishes that has been registered in Norway (Thorud & Djupvic 1988), Canada (Mullins et al. 1998), Scotland (Rodger et al. 1998), the USA (Bouchard et al. 2001) and Chile (Kibenge et al. 2001). Available information suggests that ISAV is a member of the *Orthomyxoviridae*, and based on sequence comparison of the putative RNA polymerase basic protein 1 (PB1) from ISAV with RNA polymerases

from other negative-stranded RNA viruses, it has been proposed that a new genus be created within this family (Krossøy et al. 1999). The single-stranded RNA genome consists of 8 segments ranging from approximately 1.0–2.3 kb (Mjaaland et al. 1997) and the virion contains 4 major proteins estimated at 71–74, 53, 43–46, and 24–26.5 kDa (Falk et al. 1997, Kibenge et al. 2000). The 43–46 kDa protein has been cloned and identified as the haemagglutinin (Krossøy et al. 2001, Rimstad et al. 2001), while cloning and sequence analysis have suggested the 71–74 kDa protein to represent the nucleoprotein (Ritchie et al. 2001b, Snow & Cunningham 2001). The functions of the 53 and 24–26.5 kDa proteins have not been determined.

*E-mail: eirik.biering@intervet.com

Segment 8 was the first part of the ISAV genome to be cloned and sequenced (Mjaaland et al. 1997), and it contains 2 overlapping reading frames theoretically encoding proteins with molecular weights of 22 and 27.4 kDa respectively. Although the encoded proteins were not identified by Mjaaland et al., other authors have proposed, with reference to the similarities between ISAV and the influenza viruses, that the sequence encodes NS1 and NS2 (Blake et al. 1999, Cunningham & Snow 2000, Inglis et al. 2000, Ritchie et al. 2001a,b, Snow & Cunningham 2001). In the influenza A virus, the 2 smallest segments encode the 2 matrix proteins M1 and M2 (Segment 7), and the 2 proteins NS1 and NS2 (Segment 8). Influenza viral mRNAs are synthesised in the nucleus and the mRNAs transcribed from Segments 7 and 8 exist both as co-linear and spliced transcripts. The co-linear transcript of Segment 7 encodes the M1, which is a major structural protein that underlies the lipid envelope of the virion. Segment 8 encodes the NS1, which is found mainly in the nucleus of infected cells, where it inhibits pre-mRNA splicing (Lu et al. 1994) and nuclear export of poly(A)-containing mRNAs (Qiu & Krug 1994). It has also been shown that that NS1 functions as an interferon antagonist (Garcia-Sastre et al. 1998). The M1 and NS1 mRNAs are both spliced to form mRNAs coding for M2 and NS2 respectively. The M2 is an integral membrane protein with ion channel activity (Pinto et al. 1992), whilst the NS2 functions as an export protein mediating the nuclear export of the viral ribonucleo-

proteins at a late stage of infection (O'Neill et al. 1998). NS2 is a minor structural protein, and is now commonly referred to as NEP (nuclear export protein). Except for a short region at their 5′ ends, the spliced M2 and NS2 mRNAs are transcribed in the +1 reading frame relative to their unspliced precursors.

Here we report the cloning and partial characterisation of an ISAV specific cDNA corresponding to Genomic Segment 7. The nucleotide sequence contains 2 overlapping ORFs, and we show that the co-linear transcript of the larger ORF is spliced to form a mRNA corresponding to the smaller reading frame. The larger ORF spans 903 nucleotides, theoretically encoding a protein with a predicted molecular mass of 34.2 kDa. This protein, together with the protein encoded by the first ORF of Segment 8, was expressed in *E. coli* as fusion proteins and used in immunoblot studies for characterisation purposes.

RESULTS AND DISCUSSION

A unidirectional bacteriophage Lambda cDNA library based upon mRNA from ISAV infected ASK-cells was constructed as previously described (Krossøy et al. 2001). Probe screening was used to eliminate known viral genes, and sequencing of remaining clones identified a possible ISAV mRNA with 2 open reading frames (ORFs) of 903 and 522 bases respectively. The clone was tentatively designated 9AI, and BLAST searches revealed no significant homology to other sequences. However, the GenBank Accession No. AX083264, representing an unidentified ISAV sequence from a patent belonging to Griffiths & Ritchie is 89% similar to our 9AI sequence, but is probably not full-length, as it contains 132 nucleotides less in the 5′ end of the larger ORF.

The viral origin of the 9AI cDNA sequence was demonstrated using PCR and hybridisation reactions. The specific PCR primers 9AI-F1 (5′-TGGTGTGCTG-GTTGACCAACTAAA-3′) and 9AI-R1 (5′-CCATCT-CATTGTGCTCAGGGCCAG-3′) amplified a product from ISAV-infected ASK cells and not from uninfected cells (data not shown). A 9AI-specific DNA probe, synthesised and digoxigenin (DIG)-labelled by PCR using the primers S7-615F (5′-CTTCTTTCCTGTCGGGCT-CAAAGGT-3′) and S7-883R (5′-CTCTCTCATTGT-GATCCCGTCTCCA-3′), hybridised to genomic RNA from purified ISAV (Fig. 1A: Lane 2). The 9AI probe visualised a band larger than a Segment 8-specific probe (Fig. 1A: Lane 3), but smaller than a probe specific for the ISAV hemagglutinin gene that contains an ORF of 1167 bases (Krossøy et al. 2001) (Fig. 1A: Lane 4), thus confirming that the 9AI sequence corresponds to ISAV Genomic Segment 7. The Segment 8 probe
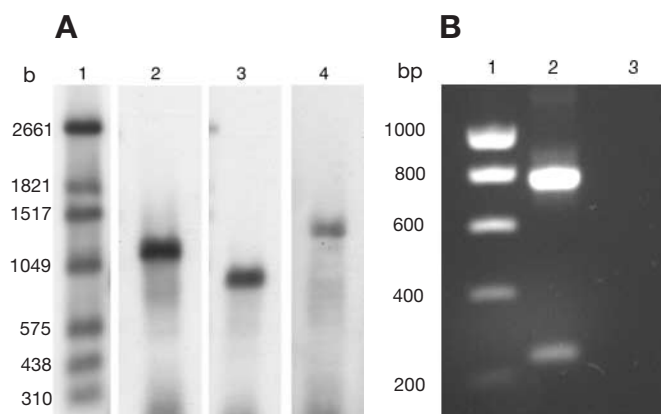

Fig. 1. Infectious salmon anaemia virus (ISAV) genomic RNA. (A) Northern blot demonstrating 9AI sequence encoded by Segment 7; Lane 1: DIG-labelled ssRNA molecular weight marker 1 (Roche); Lanes 2–4: ISAV genomic RNA detected with segment-specific probes; Lane 2: 9AI probe; Lane 3: Segment 8 probe; Lane 4: Segment 6 (haemagglutinin) probe. (B) PCR, demonstrating presence of 2 mRNAs transcribed from Segment 7 in infected cells; Lane 1: DNA ladder; Lane 2: PCR using Primers 9AI-F1 and 9AI-R2 on RNA isolated from ISAV-infected cells; Lane 3: PCR using Primers 9AI-F1 and 9AI-R2 on RNA isolated from uninfected cells

was prepared using primers 5′-GAAGAGTCA-GGATGCCAAGACG-3′ and 5′-GAAGTCGAT-GATCTGCAGCGA-3′, and the hemagglutinin probe was prepared using primers 5′-GATCAA-CGGATGCGGATATTTCA-3′ and 5′-GTCAA-AATCTTTAACCATCTTAGGGCA-3′. When the primers 9AI-F1 and 9AI-R2 (5′-TGCTAGGCT-TCTCGTAGATG-3′) was used for PCR on the total RNA from ISAV-infected cells, 2 bands of 776 and 250 bp respectively were produced, demonstrating the existence of 2 mRNAs transcribed from Segment 7 (Fig.1B: Lane 2). The specificity of the reaction was confirmed, as PCR on RNA from uninfected cells was negative (Fig. 1B: Lane 3).

To obtain a full-length cDNA sequence corresponding to the 9AI mRNA, 5′RACE was performed with the 5′RACE System, Version 2.0 (Life Technologies) using 9AI-R1 as internal primer. RACE products were cloned into the pCR 2.1-TOPO vector using the TOPO TA Cloning Kit (Invitrogen) and sequenced. Analysis of the nucleotide sequence in the 5′ region of 4 clones revealed the motif 5′-GCUAAGA-3′ preceded by heterogenous sequences. This motif has been found in ISAV Segment 4 (Ritchie et al. 2001b) and is, except for substitution of the U with an A in Position 3, identical to the motif found to be conserved for mRNAs derived from ISAV Segments 2, 3, 6 and 8 (Krossøy et al. 1999, Sandvik et al. 2000, Ritchie et al. 2001b, Snow & Cunningham 2001). The presence of this motif further supports our conclusion that the 9AI sequence is of viral origin. The 9AI sequence represents the mRNA of the larger Segment 7 ORF. In order to identify the mRNA representing the smaller ORF and thus be able to predict the theoretical MW of the encoded protein, one DIG-labelled DNA probe complementary to both reading frames was prepared using primers 5′-TGTCTGGAAGCCTCTACTGA-3′ and 5′-GAATGAT-ACGCCGTCTCTGT-3′. This probe was used to screen the bacteriophage lambda cDNA library. PCR products were then generated from positive clones using T7 and T3 vector primers, and sequencing revealed the presence of 2 partly overlapping mRNA species in the library and thus also in infected cells. A sequence analysis of the 2 mRNAs showed that they had identical 5′ and 3′ ends (except for the 5′ heterogeneous region of cellular origin), while a 526 nucleotide region was missing in the smaller (9AI-2) compared to the larger (9AI-1) mRNA. To demonstrate that this deletion could be due to a splicing event, we compared the 5′ and 3′ sequences in the splice junctions with the consensus sequences deduced from a large number of cellular and viral mRNAs. The consensus sequences are found at the 5′ and 3′ splice sites at the junctions of exons and
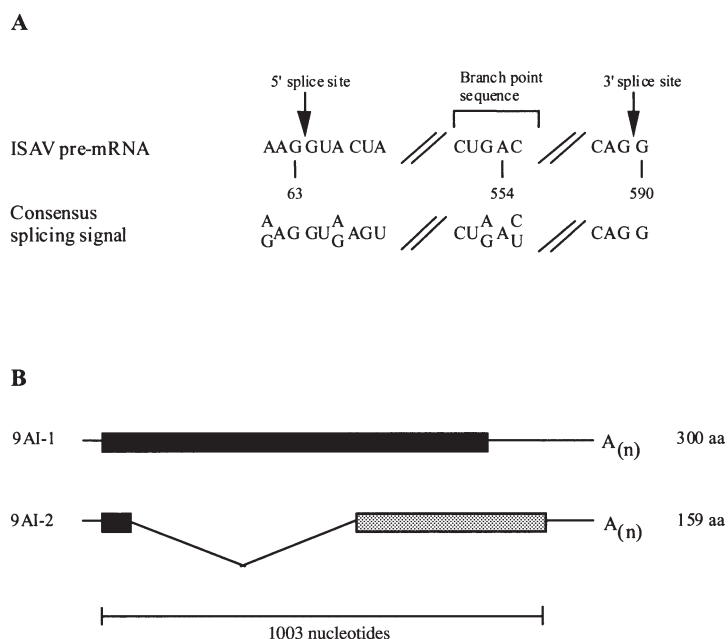

Fig. 2. Infectious salmon anaemia virus (ISAV) Segment 7 encoding of 2 mRNAs due to splicing. (A) Comparison of the ISAV sequences in the splice junctions with the conserved sequences found in cellular and viral pre-mRNAs; the 5′ and 3′ splice sites are indicated with arrows; nucleotide numbers in the ISAV sequence are relative to the initiation codon. (B) The largest ORF (9AI-1) is co-linear and predicted to encode 300 amino acids (aa, black area); spliced transcript (9AI-2), which shares the first 22 aa with 9AI-1 (black area), has an 526 nucleotide intron (V-shaped line) removed by splicing; mRNA then continues in the +1 reading frame (grey area) encoding a predicted 159 aa protein; thin lanes before and after the shaded areas represent untranslated regions. Scale bar represents total coding region in the co-linear transcript

introns, and in Introns 18 to 40 nucleotides upstream of the 3′ splice site (Flint et al. 2000). A comparison of the ISAV sequence and the consensus sequence is shown in Fig. 2A. The ISAV sequence shows distinct similarities with the consensus sequence, and it is deduced that the donor site in the splicing reaction is the G at nucleotide 63 and that the acceptor site is the G at nucleotide 590. The predicted 9AI-1 and 9AI-2 proteins share the same AUG codon for initiation of protein synthesis and 21 subsequent amino acids before the intron. The splicing event then removes a 526 nucleotide intron and the 9AI-2 mRNA continues in the +1 reading frame, encoding a predicted protein of 159 amino acids with a calculated molecular weight of 17.5 kDa. Fig. 2B shows the arrangement of the 9AI-1 and 9AI-2 mRNAs and their open reading frames. Thus, ISAV Genomic Segment 7 has an arrangement of unspliced and spliced mRNAs analogous to that of the influenza A virus Segments 7 and 8 (Lamb & Krug 1996).

Segment 7 theoretically encodes 2 proteins with predicted molecular weights of 34.2 and 17.5 kDa respec-

tively. So far, 4 structural proteins have been identified in purified Norwegian and Canadian ISAV isolates (Falk et al. 1997, Kibenge et al. 2000), and the molecular weights of these proteins have been reported to be 71–74, 53, 43–46, and 24–26.5 kDa respectively. Given that the genomic organisation in ISAV reflects the influenza viruses and, accordingly, Segment 7 encodes the matrix proteins, one should expect to find a protein of approx. 34 or 17.5 kDa in purified virus preparations. So far, such proteins have not been found, indicating that the 34 kDa protein should be subjected to post-translational modifications in order to appear as a 24 kDa structural protein. In order to investigate this, the 9AI-1 ORF was amplified from the Lambda vector using primers 5′-GCGATATCATTCGGCACGAGTC-TACAA-3′ and 5′-ATAAGAATGCGGCCGCCAGCC-AATCACATTCTGAAG-3′. The product was cloned into the *Eco*RV and *Not*I sites of the pET 30a vector (Novagen), generating a construct fused to a polyhistidine tag. The construct was transformed into TOP 10 cells (Invitrogen) and the isolated plasmid was used to transform *Escherichia coli* BL21(DE3)pLys S (Invitrogen). Expression and isolation of protein inclusion bodies were performed according to the pET System Manual (Novagen) (Fig. 3A: Lane 2). A polyclonal antiserum prepared against purified ISAV virions that reacts specifically with the 4 major ISAV proteins

found in purified virus preparation in immunoblotting failed to react with the 9AI recombinant protein (Fig. 3B: Lane 2), implying that this protein is non-structural and not present in virions. Conversely, it might be a minor structural component of the virion present at sub-detectable levels. In a recent publication, Kibenge et al. (2001) used a polyclonal anti-ISAV antiserum to immunoprecipitate 12 proteins assumed to be virus-specific, including 3 proteins of 33.5–36, 29–30, and 19–20.5 kDa. The predicted products of Segment 7 (34.2 and 17.5 kDa) might correspond to the largest and smallest of these proteins, respectively.

Genomic Segment 8 has the potential to encode 2 proteins with theoretical molecular weights of 22 and 27.4 kDa, so both may represent the 24 kDa structural protein. In order to test this hypothesis we attempted to express both proteins in *Escherichia coli*. ORF-1 (the first ORF) was amplified using primers 5′-CATGCCATGGCAA-TGAACGAATCACAATGGATACAA-3′ and 5′-CCG-CTCGAGCTGCAGGGCTTATCTCAGGTACC-3′, and ORF-2 (the second ORF) was amplified with 5′-CAT-GCCATGGATACAAAAACATCTACCATGCATGAGA-3′ and 5′-CCGCTCGAGGATTAATTTATTGTACAG-AGTCTTCCAATTGG-3′. The products were then cloned into the *Nco*1 and *Xho*1 sites of the pET 30a vector, generating constructs fused to a polyhistidine tag. Transformation and expression was performed as described for the 9AI-1protein. ORF-2 appeared to be toxic to *E. coli*, and we were unable to express the protein, even in BL21(DE3)pLysE cells (Invitrogen) that are designed for expression of toxic gene products. The ORF-1 fusion protein was expressed at low levels, purified using polyhistidine binding Ni-NTA resin (Qiagen) under denaturing conditions according to the manufacturers recommendations, and used to immunise rabbits for preparation of antiserum. The polyclonal antiserum prepared against purified ISAV virions reacted specifically with a protein in lysate of *E. coli*, expressing the Segment 8 ORF-1 fusion protein after induction with IPTG (Fig. 3B: Lane 4), indicating that this protein is a component of the virion. Without induction, no specific band was detected when the same amount of bacteria was applied to the gel (Fig. 3B: Lane 3). When the antiserum prepared against the ORF-1 fusion protein was used on purified virus (Fig. 3C: Lane 2), it reacted specifically with the same 24 kDa protein detected by the polyclonal anti ISAV serum (Fig. 3C: Lane 1), confirming that the first
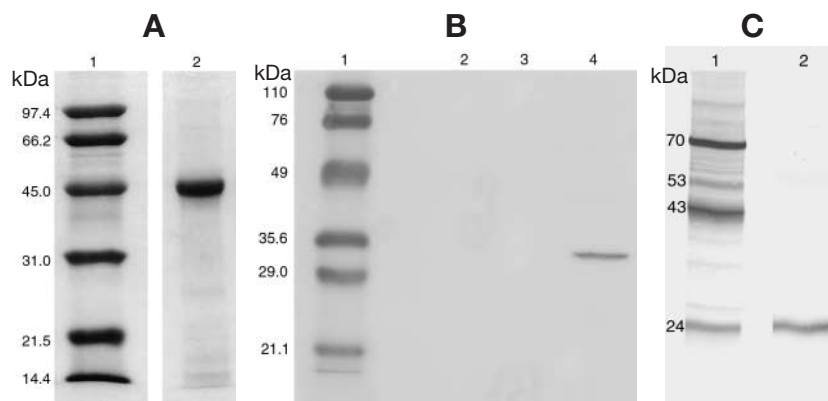

Fig. 3. (A) Expression of 9AI-1 fusion protein in *Escherichia coli*. Lane 1: molecular weight markers; Lane 2: Coomassie Brilliant Blue-stained and purified 9AI-1 fusion protein. (B) Immunoblots using a polyclonal antiserum prepared against purified Infectious salmon anaemia virus (ISAV); Lane 1: prestained molecular weight markers; Lane 2: purified 9AI-1 fusion protein (no reaction detected). Lane 3: lysate of uninduced *E. coli* containing the Segment 8 ORF-1 expression construct (no reaction detected); Lane 4: lysate of induced *E. coli* expressing the Segment 8 ORF-1 fusion protein that reacts specifically with the antiserum. (C) Immunoblots of ISAV purified by gradient centrifugation; antisera used for detection are polyclonal antiserum prepared against purified ISAV (Lane 1) and polyclonal antiserum prepared against purified Segment 8 ORF-1 recombinant fusion protein. The polyclonal antiserum prepared against purified ISAV reacts specifically with the 4 major structural ISAV proteins, and the antiserum against the Segment 8 ORF-1 recombinant fusion protein reacts specifically with the 24 kDa structural protein

ORF of Segment 8 encodes this major structural component.

To conclude, we have cloned and sequenced a new ISAV-specific cDNA sequence representing ISAV Genomic Segment 7. The sequence potentially encodes 2 proteins of 34.2 and 17.5 kDA respectively. The viral origin of this sequence was confirmed using PCR and RNA hybridisation and by the presence of an ISAV-specific conserved motif in the 5′ region. We have shown that the co-linear transcript from this ISAV segment is spliced to form a mRNA corresponding to the smaller reading frame. Thus, ISAV Genomic Segment 7 has a similar coding strategy as Segments 7 and 8 of the influenza A viruses. Expression studies to identify the protein encoded by the larger ORF of Segment 7, indicate that it is a non- or minor structural component of the virion, due to lack of reaction with a polyclonal anti-ISAV antiserum. If 1 of the ORFs of Segment 7 encodes a protein that counteracts interferon activity similar to NS1 of influenza (Garcia-Sastre et al. 1998), this might be a promising target for mutations in order to generate attenuated viruses useful as vaccines (Egorov et al. 1998). Expression of the first ORF on Segment 8 indicates that this is a major structural protein, as the polyclonal anti-ISAV antiserum specifically reacts with the recombinant protein. This conclusion is further strengthened as immune-serum raised against the recombinant Segment 8 protein specifically reacts with the 24 kDa protein found in purified virus preparations. It is tempting to speculate that this protein represents the matrix protein, as it is the smallest major structural component of the virion and encoded by a genomic segment that contains 2 ORFs.

LITERATURE CITED

Blake S, Bouchard D, Keleher W, Opitz M, Nicholson BL (1999) Genomic relationships of the North American isolate of infectious salmon anemia virus (ISAV) to the Norwegian strain of ISAV. Dis Aquat Org 35:139–144

Bouchard D, Brockway K, Giray C, Keleher W, Merril PL (2001) First report of infectious salmon anemia (ISA) in the United States. Bull Eur Assoc Fish Pathol 21:86–88

Cunningham CO, Snow M (2000) Genetic analysis of infectious salmon anaemia virus (ISAV) from Scotland. Dis Aquat Org 41:1–8

Egorov A, Brandt S, Sereinig S, Romanova J and 6 others (1998) Transfectant influenza A viruses with long deletions in the NS1 protein grow efficiently in Vero cells. J Virol 72:6437–6441

Falk K, Namork E, Rimstad E, Mjaaland S, Dannevig BH (1997) Characterization of infectious salmon anemia virus, an orthomyxo-like virus isolated from Atlantic salmon (*Salmo salar* L.). J Virol 71:9016–9023

Flint SJ, Enquist LW, Krug RM, Racaniello VR, Skalka AM (2000) Principles of virology; molecular biology, pathogenesis and control. ASM Press, Washington, DC

Garcia-Sastre A, Egorov A, Matassov D, Brandt S, Levy DE, Durbin JE, Palese P, Muster T (1998) Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. Virology 252:324–330

Inglis JA, Bruce J, Cunningham CO (2000) Nucleotide sequence variation in isolates of infectious salmon anaemia virus (ISAV) from Atlantic salmon Salmo salar in Scotland and Norway. Dis Aquat Org 43:71–76

Kibenge FS, Lyaku JR, Rainnie D, Hammell KL (2000) Growth of infectious salmon anaemia virus in CHSE-214 cells and evidence for phenotypic differences between virus strains. J Gen Virol 1:143–150

Kibenge FS, Gárate ON, Johnson G, Arrigada R, Kibenge MJT, Wadowska D (2001) Isolation and identification of infectious salmon anaemia virus (ISAV) from Coho salmon in Chile. Dis Aquat Org 45:9–18

Krossøy B, Hordvik I, Nilsen F, Nylund A, Endresen C (1999) The putative polymerase sequence of infectious salmon anemia virus suggests a new genus within the Orthomyxoviridae. J Virol 73:2136–2142

Krossøy B, Devold M, Sanders L, Knappskog PM and 6 others (2001) Cloning and identification of the infectious salmon anaemia virus haemagglutinin. J Gen Virol 82:1757–1765

Lamb RA, Krug RM (1996) *Orthomyxoviridae*: the viruses and their replication. In: Fields BN, Knipe DM, Howley PM, Chanock RM, Melville KJ, Monath TP, Roizman B, Straus SE (eds) Fields virology. Lippincott-Raven, Philadelphia, p 1353–1395

Lu Y, Qian XY, Krug RM (1994) The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing. Genes Dev 8: 1817–1828

Mjaaland S, Rimstad E, Falk K, Dannevig BH (1997) Genomic characterization of the virus causing infectious salmon anemia in Atlantic salmon (*Salmo salar* L.): an orthomyxo-like virus in a teleost. J Virol 71:7681–7686

Mullins JE, Groman D, Wadowska D (1998) Infectious salmon anaemia in salt water Atlantic salmon (*Salmo salar* L.) in New Brunswick, Canada. Bull Eur Assoc Fish Pathol 18: 110–114

O'Neill RE, Talon J, Palese P (1998) The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. EMBO (Eur Mol Biol Organ) J 17:288–296

Pinto LH, Holsinger LJ, Lamb RA (1992) Influenza virus M2 protein has ion channel activity. Cell 69:517–528

Qiu Y, Krug RM (1994) The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). J Virol 68:2425–2432

Rimstad E, Mjaaland S, Snow M, Mikalsen AB, Cunningham CO (2001) Characterization of the infectious salmon anemia virus genomic segment that encodes the putative hemagglutinin. J Virol 75:5352–5356

Ritchie RJ, Cook M, Melville K, Simard N, Cusack R, Griffith S (2001a) Identification of infectious salmon anaemia virus in Atlantic salmon from Nova Scotia (Canada): evidence for functional strain differences. Dis Aquat Org 44: 171–178

Ritchie RJ, Heppell J, Cook MB, Jones S, Griffiths SG (2001b) Identification and characterization of segments 3 and 4 of

the ISAV genome. Virus Genes 22:289–297

Rodger HD, Turnbull T, Muir F, Millar S, Richards RH (1998) Infectious salmon anaemia (ISA) in the United Kingdom. Bull Eur Assoc Fish Pathol 18:115–116

Sandvik T, Rimstad E, Mjaaland S (2000) The viral RNA 3′- and 5′- end structure and mRNA transcription of infectious salmon anaemia virus resemble those of influenza viruses.

Arch Virol 145:1659–1669

Snow M, Cunningham CO (2001) Characterisation of the putative nucleoprotein gene of infectious salmon anaemia virus (ISAV). Virus Res 74:111–118

Thorud KE, Djupvik HO (1988) Infectious salmon anaemia in Atlantic salmon (*Salmo salar* L.). Bull Eur Assoc Fish Pathol 8:109–111