Identification of a new Infectious Pancreatic Necrosis Virus (IPNV) isolate in Atlantic salmon (*Salmo salar* L.) that causes mortality in resistant fish

Borghild Hillestad[1], Stein Johannessen[1], Geir Olav Melingen[2] and Hooman K. Moghadam[1*]

[1]*Benchmark Genetics Norway AS, Sandviksboder 3A, N-5035 Bergen, Norway;*

[2]*Benchmark Genetics, Sandviksboder 3A, N-5035 Bergen, Norway;*

Article type:

Original Research

Running title:

Identification of a new pathogenic IPNV isolate

[*]**Corresponding author:**

Hooman K. Moghadam, hooman.moghadam@bmkgenetics.com

Abstract

Infectious pancreatic necrosis (IPN) is an important viral disease of salmonids that can affect fish during various life cycles. In Atlantic salmon, selecting for genetically resistant animals against IPN has been one of the most highly praised success stories in the history of fish breeding. The findings that resistance against this disease has a significant genetic component, which is mainly controlled by variations in a single gene, has helped to reduce the IPN outbreaks over the past decade to a great extent. In this paper, we present the identification of a new isolate of the IPN virus, from a field outbreak, that had caused mortality, even in the genetically resistant animals. We recovered and assembled the full-length genome of this virus, following deep-sequencing of an infected tissue. The comparative sequence analysis revealed that for the critical amino acid motifs, previously found to be associated with the degree of virulence, the newly identified isolate is similar to the virus's avirulent form. However, we detected a set of deduced amino acid residues, particularly in the hypervariable region of the polyprotein, that collectively are unique to this strain compared to all other reference sequences assessed in this study. We suggest that these mutations have likely equipped the virus with the capacity to escape the host defence mechanism more efficiently, even in the genetically deemed IPN resistant animals.

Keywords:

Infectious pancreatic necrosis (IPN); RNA sequencing; Genome assembly; Quantitative Trait Locus (QTL); Atlantic salmon;

Introduction

Infectious pancreatic necrosis (IPN) is one of the leading viral diseases of the Norwegian farmed Atlantic salmon (*Salmo salar* L.). The disease was first reported in 1941, following an outbreak in brook trout (*Salvelinus fontinalis* L.) in Canada (M'Gonigle, 1941). Further, the characterisation of the causative agent, the IPN virus (IPNV), was performed in 1960 (Wolf et al., 1960). In Norway, the virus was first isolated in 1975 from freshwater rainbow trout (*Oncorhynchus mykiss* L.) (Hastein and Krogsrud, 1976) and was designated as a notifiable disease from 1991 till 2008 (Sommerset et al., 2020). The virus can infect Atlantic salmon during all of its developmental stages, but the fish are especially susceptible as fry, during start-feeding, and post-smolts, soon after transfer to seawater (Roberts and Pearson, 2005; Sommerset et al., 2020). As suggested by the name, the pancreas is the primary target tissue of IPNV. However, the liver has also been shown to be one of the key organs affected by this virus (Ellis et al., 2010; Munang'Andu et al., 2013). In addition to significant economic losses, this disease is a concern for animal welfare, as survivors following infection can continue to infect naïve fish (Roberts and Pearson, 2005).

IPNV is a double-stranded RNA virus that belongs to the *Birnaviridae* family. A characteristic feature of the birnaviruses is their possession of a bi-segmented genome (i.e., A and B segments), contained within a non-enveloped, icosahedral capsid (Dobos and Roberts, 1983; Duncan et al., 1987; Duncan and Dobos, 1986). The smaller genomic segment, the B segment, is about 2.5 kb in size and consists of a single open reading frame (ORF), encoding an RNA-dependent, RNA polymerase VP1 (approximately 90 kDa) (Duncan et al., 1987; Duncan and Dobos, 1986). The A-segment is approximately 3 kb, and it contains two, partially overlapping ORFs (Duncan and Dobos, 1986). The first ORF, encodes VP5, a small, cysteine-rich, non-structural protein of about 17 kDa. A precursor polyprotein (NH2-preVP2-NS VP4 protease-VP3-COOH) of approximately

106 kDa constitutes the second and the larger ORF. The encoded VP4 protease cleavages the polyprotein into its two main components, the preVP2 and VP3. The preVP2 further matures to become VP2 and forms the outer capsid protein, containing the neutralizing epitopes and sites that facilitate cell attachment (Dopazo, 2020; Heppell et al., 1995b). The VP3 is the inner capsid protein (Duncan et al., 1987; Duncan and Dobos, 1986) but also in association with the VP1 it seems to be involved in viral packaging and replication (Tacken et al., 2002).

Based on cross-neutralizing tests, the aquatic birnaviruses are broadly divided into two main serogroups, A and B (Hill and Way, 1995). The serogroup A comprises of nine different serotypes, A_1-A_9. Most of the isolates from the USA fall within the A_1 serotype (West Buxton; WB), the A_2-A_5 include mainly the European isolates (A_2 (Spjarup; Sp), A_3 (Abild; Ab), A_4 (Hecht, He) and A_5 (Te)), and A_6-A_9 are predominantly variants from Canada (A_6 (Canada 1; C1), A_7 (Canada 2; C2), A_8 (Canada 3; C3) and A_9 (Jasper)) (Blake et al., 2001; Hill and Way, 1995). The serogroup B consists of only one serotype (B_1 (Tellinavirus; TV1)) that is non-pathogenic in fish. Based on the analysis of variations in the nucleotide and the deduced amino acids of the VP2, Blake et al. (2001) later proposed that aquatic birnaviruses constitute six major genogroups. These groups correspond relatively well with the geographical origins and serological classifications previously suggested. According to this classification, WB and the Jasper strains constitute genogroup 1, the A_3 serotype forms genogroup 2, C1 and Te make genogroup 3, the two Canadian strains, C2 and C3 form genogroup 4, all European isolates of serotype A_2 cluster into genogroup 5 and the He strain constitutes genogroup 6 (Blake et al., 2001). A seventh genogroup was also suggested based on isolates recovered from Japanese aquatic sources (Nishizawa et al., 2005).

Despite extensive vaccination efforts, disinfection and targeted breeding programs, IPN outbreaks remains a concern in farmed Atlantic salmon. Related mortalities can be substantial and in part,

can be a function of the genetic makeup of the host (Guy et al., 2006; Houston et al., 2008; Moen et al., 2009), management and environmental factors (Jarp et al., 1995; Sundh et al., 2009), and the strain and specific genetic variations carried by the virus (Skjesol et al., 2011; Song et al., 2005). There are ample examples of isolates, belonging to the same serotype, to cause different degrees of virulence in the host (e.g., Bruslind and Reno, 2000; Santi et al., 2004; Shivappa et al., 2004), and indeed, there have been many studies, investigating the molecular basis of IPNV virulence in Atlantic salmon (e.g., Mutoloki et al., 2016; Santi et al., 2004, 2005; Skjesol et al., 2011; Song et al., 2005). Although the genetic details of such processes are not yet fully understood, most studies speculated, and chiefly investigated, the genetic variations in segment A concerning the degree of virulence in IPNV. In particular, specific amino acid motifs in the VP2 of the polyprotein have been suggested to be most prevalent and in association with the virulent forms of the pathogen (Bruslind and Reno, 2000; Munang'andu et al., 2016; Mutoloki et al., 2016; Santi et al., 2004; Shivappa et al., 2004; Song et al., 2005). Variations in the amino acid residues in positions 217, 221 and 247 of the VP2 have been of particular focus (Dopazo, 2020; Julin et al., 2013; Munang'andu et al., 2016; Mutoloki et al., 2016; Santi et al., 2004; Song et al., 2005). Further, assessment of the recombinant viral strains, have indicated that the residue 217 of the VP2 might be the most crucial amino acid that can result in the loss or gain of virulence (e.g., Song et al., 2005).

Despite concerns and economic losses due to infection by IPNV, increased resistance against this virus through selective breeding is among the most remarkable success stories in the history of aquaculture and salmon breeding in particular. The initial attempt to estimate the genetic parameters controlling this trait, which was based on mortality in the field and controlled challenge testing, had shown that resistance to IPNV has a significant and heritable genetic component (h^2

of 0.17-0.62) (Kjøglum et al., 2008; Wetten et al., 2007). However, the breakthrough came with the finding that resistance to IPNV is mainly controlled by variations in a single quantitative trait locus (QTL) on chromosome 26 (Houston et al., 2008, 2010; Moen et al., 2009). Subsequent work identified mutations in the epithelial cadherin gene (*cdh1*) as the causative genetic variation for IPNV resistance in Atlantic salmon (Moen et al., 2015). Following the inclusion of this QTL into the breeding programs in Norway, the industry witnessed a sharp decline in the number of IPN outbreaks throughout the country, dropping from 223 to only 19 reported cases from 2009 to 2019 (Sommerset et al., 2020).

This paper reports the identification and whole-genome sequence assembly and analysis of a new IPNV variant. The isolate was recovered from a field outbreak in West of Norway following reported mortalities, due to an IPN infection, among QTL homozygous or heterozygous IPN resistant fish. We describe the amino acid motifs that distinguish this isolate from other phylogenetically related variants and suggest that these mutations are likely to play an essential role in this isolate's pathogenicity.

Materials and Methods

Field outbreak and sample collection

In May of 2019, there was an IPN field outbreak in a Western region of Norway. The affected fish were Atlantic salmon of the SalmoBreed strain, delivered as IPN QTL, all favourable homozygous or heterozygous for the QTL locus (Houston et al., 2008; Moen et al., 2009). Fish were at the post-smolt stage with an approximate average weight of about 700 – 800 gr at the time of the outbreak. The reported mortality reached about 10%, a high percentage for field outbreak among QTL carrying fish. The IPNV infection was identified following an inspection by the veterinary authorities. The diagnosis was later validated by detection of the VP2 segment of the viral genome

using polymerase chain reaction (qRT-PCR) of the infected head-kidney and through histopathological and immunohistochemical examinations of the liver and pancreatic tissues (Figure 1A and B), all performed at the Fish Vet Group Norway (http://fishvetgroup.no) (Supplementary Table 1). Tissues were fixed in a 4% formalin solution (4% formalin, 0.08M sodium phosphate, pH 7.0), processed using Thermo Scientific Excelsior® and embedded in Histowax with the aid of Tissue – Tek®, TEC 5 (Sakura) embedding system. The embedded tissues were then sectioned at 1.5-2 µm using a Leica RM 2255 Microtome and stained with Hematoxylin-Eosin (HE). The stained sections were scanned in an Aperio ScanScope AT Turbo slide scanner and read using Aperio Image Scope (Leica Biosystems). The scans were examined and scored in a blind fashion, i.e., without information about the animal's history. A clip from the adipose fin was also sampled and stored in 95% ethanol for DNA extraction and genotyping of ten fish, five moribund and five dead. We genotyped these fish for the genetic markers used to assign individuals as QTL carrier (Houston et al., 2008; Moen et al., 2009, 2015), and confirmed all animals carry IPN QTL.

Transcriptome sequencing and viral genome assembly

The head-kidney was collected from six infected fish (http://fishvetgroup.no/pcr-2/#1452163253406-d81f6dce-bdf9), all at a moribund state, and stored in RNALater® (Ambion). Total RNA extraction was performed using the RNeasy® Plus mini kit (Qiagen). Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies) was used to assess the concentration and purity of the extracted genetic materials. The Norwegian Sequencing Centre performed library construction and sequencing the total RNA on an Illumina HiSeq 4000 platform as paired-end (PE) 150 bp reads.

We first cleaned the sequence data by removing low-quality nucleotides and sequencing adapters using Trimmomatic (v.0.36) (Bolger et al., 2014). The ribosomal RNA (rRNA) was then identified through a similarity search against the SILVA rRNA database (Quast et al., 2013) and excluded from subsequent analysis. The host-specific transcriptome was detected by aligning the remaining reads against the Atlantic salmon reference genome assembly (ICSASG_v2) (Lien et al., 2016) using TopHat (v.2.0.13) (Trapnell et al., 2009, 2012). The un-aligned reads were then blasted against the IPNV reference assemblies (ASM397166v1, ASM397170v1, ASM397174v1, ASM397168v1, ASM397172V1 and ViralMultiSegProj15024) by setting "task" as "megablast", "word size" to 7 and "expectation value threshold" to $1.00e^{-07}$. All the reads that aligned to the IPNV reference sequences were extracted, pooled into a single sequence datafile per animal, and then fed into Trinity (v.2.11.0) (Grabherr et al., 2011) with the default parameter settings to construct assemblies.

Comparative genomics and phylogenetic analysis

We compared the assembled viral genomes to one another (Supplementary Tables 1 and 2), and against various reference isolates, with different degrees of virulence (Dopazo, 2020; Julin et al., 2013; Munang'andu et al., 2016; Mutoloki et al., 2016; Santi et al., 2004; Song et al., 2005) (Table 1) or strains representing different genogroups, as previously suggested (Blake et al., 2001) (Table 2). Further, we assessed the deduced amino acids of these isolates to any IPNV sequence in the NCBI database with a complete polyprotein or VP1 information (Supplementary Files 1 and 2). The sequences included 90 isolates with the full coding sequence for the polyprotein and 120 for the VP1. Both, the nucleotide and amino acid alignments were performed using ClustalX (Larkin et al., 2007) implemented in Unipro UGENE (v36.0) (Okonechnikov et al., 2012) or Muscle (Edgar, 2004) implemented in Seqotron (v1.0.1) (Fourment and Holmes, 2016). Unipro UGENE

and Mega X (v10.1.7) (Kumar et al., 2018) were also used to assess sequence similarity, calculate pair-wise distances for nucleotide and amino acid data, and to construct phylogenetic associations using the neighbour-joining (NJ) method. To assess the confidence of node assignments, we performed 1000 bootstrap replications.

Results and Discussion

Transcriptome sequencing and genome assembly

This paper describes the genomic features of IPNV isolates, responsible for a field outbreak in an Atlantic salmon population. The post-mortem examination identified IPN infection as the chief cause of mortality among all fish investigated. The histopathological and immunohistochemical examinations (Figure 1A and B and Supplementary Table 1) and the viral genetic material detected through qRT-PCR amplification (Supplementary Table 1) further confirmed earlier diagnosis. We found that the genotypes of all animals examined in this study were consistent with an IPN resistant animal (Houston et al., 2008; Moen et al., 2009, 2015). Further, assessment of the transcriptome sequence data confirmed that all fish carried at least one favourable copy of the causal mutation in the *cdh1* gene, conferring resistance against IPNV (Moen et al., 2015). The causal mutation, which is in the form of a single nucleotide polymorphic (SNP) variation, is located in the protein-coding sequence of *cdh1* (ssa26:15192533; C/T) and can change Proline (P) to Serine (S), with the former amino acid found to be associated with the IPNV resistance (Moen et al., 2015). It is expected that in a heterozygote fish, the favourable allele exhibit close to complete dominant effect in response to the virus (Moen et al., 2015). Among the animals investigated in this study, three fish were homozygous for the favourable allele (C/C), and the other three were heterozygous (C/T). The extracted RNA was sequenced to high depth, with an average of 60 million PE reads per animal. Following the rRNA's exclusion, more than 45 million PE reads per fish remained for the

assemblage of the host's viral genome and profiling gene expression. Together, the two segments of the IPNV genome consist of about 5800 nucleotides. On average, we obtained 56-fold coverage of the viral genome from each of the animals in our sequence data. Here, we refer to the assembled viral genomes as BG_x_y, where x denotes the animal's id (1-6) and y refers to the viral genome's specific segment (e.g., VP1 or polyprotein).

Comparative sequence analysis of the polyprotein

Comparative genomic analysis between the six assembled genomes revealed at least two different, closely linked variants, caused the field outbreak in the population investigated (Figure 2A and B; Supplementary Table 2A and B). The nucleotide sequences of the polyprotein are identical in four of the assemblies (BG_1_polyprot to BG_4_polyprot) and contain nine polymorphic sites compared to the other two assembled sequences (i.e., BG_5_polyprot and BG_6_polyprot). Two of the polymorphisms are nonsynonymous, changing the amino acid residues in positions 47 (aspartate (D) to glutamate (E)) and 717 (lysine (K) to glutamine (Q)).

In a study, aiming to trace the IPNV infection signatures from hatcheries to the sea (Kristoffersen et al., 2018), the authors have recovered a partial fragment of the polyprotein from an isolate (accession number: MH562009), with very similar nucleotide sequence to the ones reported here. Recently, in controlled infection challenge testing, the mortality induced by this isolate among full-sib families was compared with a standard, virulent form of the virus (AY379740.1). While the infection with the latter variant resulted in approximately 30% mortality, mainly among the non-QTL fish, the former isolate caused more than 75% mortality, treating the QTL and non-QTL fish alike (Ørpetveit et al. in preparation; https://www.vetinst.no/nyheter/endringer-i-ipn-viruset-gjor-fisken-mer-utsatt-for-sykdom, https://www.fishfarmingexpert.com/article/changes-in-ipn-virus-make-salmon-more-susceptible, https://salmobreed.no/articles/benchmark-genetics-

intensiverer-forskningen-pa-nytt-ipn-virus). Comparison between the nucleotide sequences of this isolate with our assembled genomes shows divergence ranging from 0.003-0.004 (Supplementary Figure 1A). However, this isolate has deduced amino acid sequences identical to the translated sequences from our first four assemblies (Supplementary Figure 1B), indicating that this variant is likely to be the ancestral form and the other variants (i.e., BG_5_VP2 and BG_6_VP2) are more recently derived. One of the main characteristics of RNA viruses is their high mutation rates, mainly attributed to their lack of an effective proofreading activity. Therefore, finding different sequence variants from the same outbreak might not be surprising, but it indeed is important, as it will allow one to follow the trajectory of pathogen evolution through time and space.

Comparative assessment of the polyprotein nucleotide (data not shown) as well as the amino acid sequence data with multiple reference strains (Figure 2A) shows our assemblies are most closely related to the Sp serotype (genogroup 5). However, they also constitute a distinct group within this serotype compared to Sp isolates assessed in this study (Figure 2A; Supplementary Figure 2A). Several investigations have previously reported variations in some amino acid residues, particularly in the VP2 of the polyprotein, to be associated with the degree of the IPNV virulence (Bruslind and Reno, 2000; Munang'andu et al., 2016; Mutoloki et al., 2016; Santi et al., 2004; Shivappa et al., 2004; Song et al., 2005). Interestingly, the amino acid motifs found in the six assembled genomes in this study are mainly consistent with the virus's avirulent form (Table 1) (Santi et al., 2004; Shivappa et al., 2004; Song et al., 2005). For instance, for the most critical positions associated with the virulence, i.e., 217, 221 and 247 of the VP2, the amino acid residues of our assembled genomes are proline (P), threonine (T) and alanine (A) (Table 1; Supplementary File 1), while it is expected that in the virulent form, the corresponding residues to be T, A and T, respectively (Mutoloki et al., 2016; Santi et al., 2004; Song et al., 2005). On the other hand, we

identified ten amino acid residues in the assembled genomes' polyprotein, that either exclusively or with a disproportionate frequency are found in the reported isolates (Table 2; Supplementary File 1). Nine of these sites are within the VP2, and one is located in the VP4. In the VP2, these residues are all in proximity to one another, encompassing the hypervariable region (Blake et al., 2001; Heppell et al., 1995b). Specifically, these residues are located in positions 245, 248, 252, 255, 257, 278, 282, 285, 321 (Table 2; Supplementary File 1). This data provides additional support for the previously suggested notion that the amino acid motifs in the VP2 capsid protein hypervariable region, might play an essential role in the degree of the virulence displayed between different IPNV isolates (Bruslind and Reno, 2000; Skjesol et al., 2011). It is also likely that this specific constellation of the amino acid residues reported in these isolates has helped the virus to escape the host defence barrier, despite fish being regarded as resistant to IPNV. In the VP4, the deduced amino acid residue, unique to the assembled isolates (Table 2; Supplementary File 1), is located in position 585. The amino acid substitution is a glycine (G) residue, compared to the other sequences investigated in this study, where they carry an aspartate (D) (Table 2).

Comparative sequence analysis of VP5

All assemblies contain the VP5, the partial overlapping, non-structural protein, of segment A. Similar to the polyprotein, the phylogenetic analysis of the VP5, grouped the assemblies into two distinct groups and further clustered them with the isolates from the Sp serotype (Figure 2B). The two assembled groups differ in only one amino acid residue at position 90, with the first four assemblies carrying an arginine (R) while the two other assemblies are having a glycine (G) (Supplementary File 3). Sequence analysis of the deduced amino acids of VP5 did not reveal any unique residue in the assembled isolates compared to all other sequences. Comparing to the other Sp serotypes, however, the amino acid in position 98 differs between the isolates (Supplementary

File 3). At this position, all the investigated Sp isolates carry an R while in the assemblies, the residue is tryptophan (W). It has been shown that the VP5's protein is generally produced during the initial stages of replication (Heppell et al., 1995a), and at least in the Ab serotype, it seems that this protein to be involved in anti-apoptotic functions (Hong and Wu, 2002). However, this product's importance in viral growth or virulence has so far remained ambiguous (e.g., Julin et al. 2013; Dopazo 2020).

Comparative sequence analysis of VP1

The phylogenetic analysis of VP1 clustered the assemblies into three distinct clades (Supplementary Figure 2B) with the amino acid variations in positions 119 (threonine (T) to methionine (M)) and 701 (T to alanine (A) and serine (S)) (Supplementary File 2; Supplementary Table 3A). The nucleotide sequence comparisons, shows a greater rate of divergence in the VP1 of the assemblies compared to those of the polyprotein, suggesting that the B segment of the genome has a higher rate of nucleotide mutation (Supplementary Tables 1 and 2). Comparing the deduced amino acid sequences of the assembled genomes to the VP1 section of 120 IPNV sequences, further revealed only a single, almost unique residue in our isolates (Supplementary File 2). The amino acid residue in position 656 of the assembled isolates is glutamate (E), while in the majority of other strains, the amino acid is either aspartate (D) or alanine (A). The only other isolates, carrying an E in this position, are two sequences recovered from Atlantic salmon (KY548520) and Brown trout (*Salmo trutta* L.) (KY548519) in Finland following outbreaks (Holopainen et al., 2017). However, it should be noted that so far, the role of VP1 in determining the degree of IPNV virulence has remained ambiguous. For example, in a study conducted by Song et al. (2005), through construction and comparative assessment of chimeric IPNV strains, the authors suggested that VP1 is not involved in this pathogen's virulence. On the other hand, the data

296 reported by Shivappa et al. (2004), is suggestive that the amino acids on the B segment of the
297 genome in combination with specific residues in VP2, might help to modulate the magnitude of
298 virulence. Work in the infectious bursal disease virus (IBDV) has also shown that variations in the
299 amino acid residues of the VP1 can change the kinetics of viral replication and influence the virus's
300 virulence (Liu and Vakharia, 2004). VP1 interacts with VP3 to form VP1-VP3 complexes, and as
301 such, can affect various aspects of the virus biology, including its replication efficiency (Pedersen
302 et al., 2007; Tacken et al., 2002).

303 **Concluding remarks**

304 In light of the findings reported in this study, a key point worth noting is the importance of
305 understanding the genetic interactions between the host and the pathogen in determining an
306 infection's outcome. As a general trend, research studies tend to mainly focus on either the host or
307 the pathogen's genetic polymorphisms while investigating the genetic basis of disease-related
308 phenotypic variations. While the merits of such a targeted approach cannot be disputed, it is
309 becoming increasingly important also to be aware of the detailed genetic and genomic variations
310 and interactions between the host and the microbe. In the case of monogenic infectious diseases,
311 it is relatively trivial to speculate whether a new form of a pathogen interacts in fundamentally
312 different ways with the host's defence mechanism. For instance, considering that resistance to
313 IPNV in Atlantic salmon is mainly determined by variations in a single gene, one might expect
314 that the mutations acquired by the virus strain reported in this paper might have offered an adaptive
315 advantage, assisting the pathogen to establish successful infection, irrespective of the variations in
316 the *cdh1* gene. Of course, pinpointing the causative mutations and understanding and validating
317 the molecular basis of such adaptive variations is a substantial undertaking by itself. The task of
318 identifying and untangling the dynamics of how genotypic variations in the host and the pathogen

influence one another and affect the outcome of an infection becomes more challenging while investigating diseases with polygenetic nature. This point becomes even more critical, considering our rapidly changing environments that can facilitate the transmission dynamics and the geographical spread of the pathogens (Reid et al., 2019). For instance, changes in water temperature, pH and oxygen have been linked to saprolegniasis in the Indian major carps (*Catla catla*, *Labeo rohita* and *Chirrinus mrigal*) (Das et al., 2012) and to white spot syndrome virus in prawns (*Penaeus monodon* and *Fenneropenaeus indicus*) (Selvam et al., 2012). Fortunately, advancements in technologies such as sequencing, genotyping, computational analysis and molecular biology can now provide us with many necessary tools, if we are set to understand the dynamics of the interactions between the host and the microbe at their detailed molecular levels.

Data availability

All datasets generated for this study are included in the article (Supplementary Data) and have also been deposited in GenBank under the accession numbers MW496366, MW496367, MW496368, MW496369, MW496370, MW496371, MW496372, MW496373, MW496374, MW496375, MW496376 and MW496377.

Funding

There is no external funding to report.

Acknowledgements

We would like to thank the Fish Vet Group Norway and especially Dr Marianne Kraugerud and Mr Simon Rey for all the laboratory work and the Norwegian Sequencing Centre for the excellent work and guidance.

Contributions

G.M., S.J. and B.H. arranged for tissue sampling, shipment and all laboratory and sequencing procedures. H.M. performed analyses of the sequence data. B.H. and H.M. drafted the manuscript. All authors contributed to the development of the paper, interpretation of the results and improvement of the manuscript.

Disclosure of potential conflicts of interest

No potential conflicts of interest.

References

Blake, S., Ma, J. Y., Caporale, D. A., Jairath, S., and Nicholson, B. L. (2001). Phylogenetic relationships of aquatic birnaviruses based on deduced amino acid sequences of genome segment A cDNA. *Dis. Aquat. Organ.* 45, 89–102.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–20.

Bruslind, L. D., and Reno, P. W. (2000). Virulence comparison of three Buhl-subtype isolates of infectious pancreatic necrosis virus in brook trout fry. *J. Aquat. Anim. Health* 12, 301–15.

Das, S. K., Murmu, K., Das, A., Shakuntala, I., Das, R. K., Ngachan, S. V., et al. (2012). Studies on the identification and control of pathogen Saprolegnia in selected Indian major carp fingerlings at mid hill altitude. *J. Environ. Biol.* 33, 545–549.

Dobos, P., and Roberts, T. E. (1983). The molecular biology of infectious pancreatic necrosis virus: A review. *Can. J. Microbiol.* 29, 377–84.

Dopazo, C. P. (2020). The infectious pancreatic necrosis virus (IPNV) and its virulence determinants: What is known and what should be known. *Pathogens* 9, 94.

Duncan, R., and Dobos, P. (1986). The nudeotide sequence of infectious pancreatic necrosis virus (IPNV) dsRNA segment a reveals one large orf encoding a precursor polyprotein.

364 *Nucleic Acids Res.* 14, 5934.

365 Duncan, R., Nagy, E., Krell, P. J., and Dobos, P. (1987). Synthesis of the infectious pancreatic
366 necrosis virus polyprotein, detection of a virus-encoded protease, and fine structure
367 mapping of genome segment A coding regions. *J. Virol.* 61, 3655–64.

368 Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high
369 throughput. *Nucleic Acids Res.* 32, 1792–7.

370 Ellis, A. E., Cavaco, A., Petrie, A., Lockhart, K., Snow, M., and Collet, B. (2010). Histology,
371 immunocytochemistry and qRT-PCR analysis of Atlantic salmon, Salmo salar L., post-
372 smolts following infection with infectious pancreatic necrosis virus (IPNV). *J. Fish Dis.* 33,
373 803–18.

374 Fourment, M., and Holmes, E. C. (2016). Seqotron: A user-friendly sequence editor for Mac OS
375 X. *BMC Res. Notes* 9, 106.

376 Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011).
377 Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat.*
378 *Biotechnol.* 29, 644–52.

379 Guy, D. R., Bishop, S. C., Brotherstone, S., Hamilton, A., Roberts, R. J., McAndrew, B. J., et al.
380 (2006). Analysis of the incidence of infectious pancreatic necrosis mortality in pedigreed
381 Atlantic salmon, Salmo salar L., populations. *J. Fish Dis.* 29, 637–47.

382 Hastein, T., and Krogsrud, J. (1976). Infectious pancreatic necrosis: first isolation of virus from
383 fish in Norway. *Acta Vet. Scand.* 17, 109–11.

384 Heppell, J., Tarrab, E., Berthiaume, L., Lecomte, J., and Arella, M. (1995a). Characterization of
385 the small open reading frame on genome segment A of infectious pancreatic necrosis virus.
386 *J. Gen. Virol.* 76, 2091–96.

387 Heppell, J., Tarrab, E., Lecomte, J., Berthiaume, L., and Arella, M. (1995b). Strain Variability
388 and Localization of Important Epitopes on the Major Structural Protein (VP2) of Infectious
389 Pancreatic Necrosis Virus1. *Virology* 214, 40–49.
390 Hill, B. J., and Way, K. (1995). Serological classification of infectious pancreatic necrosis (IPN)
391 virus and other aquatic birnaviruses. *Annu. Rev. Fish Dis.* 5, 55–77.
392 Holopainen, R., Eriksson-Kallio, A. M., and Gadd, T. (2017). Molecular characterisation of
393 infectious pancreatic necrosis viruses isolated from farmed fish in Finland. *Arch. Virol.* 162,
394 3459–71.
395 Hong, J. R., and Wu, J. L. (2002). Induction of apoptotic death in cells via Bad gene expression
396 by infectious pancreatic necrosis virus infection. *Cell Death Differ.* 9, 113–24.
397 Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Mota-Velasco, J. C., Gheyas, A. A., et
398 al. (2010). The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic
399 necrosis is largely explained by a major QTL. *Heredity (Edinb).* 105, 318–27.
400 Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Tinch, A. E., Taggart, J. B., et al. (2008).
401 Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic
402 salmon (Salmo salar). *Genetics* 178, 1109–15.
403 Jarp, J., Gjevre, A. G., Olsen, A. B., and Bruheim, T. (1995). Risk factors for furunculosis,
404 infectious pancreatic necrosis and mortality in post-smolt of Atlantic salmon, Salmo solar L.
405 *J. Fish Dis.* 18, 67–78.
406 Julin, K., Mennen, S., and Sommer, A. I. (2013). Study of virulence in field isolates of infectious
407 pancreatic necrosis virus obtained from the northern part of Norway. *J. Fish Dis.* 36, 89–
408 102.
409 Kjøglum, S., Henryon, M., Aasmundstad, T., and Korsgaard, I. (2008). Selective breeding can

410 increase resistance of Atlantic salmon to furunculosis, infectious salmon anaemia and
411 infectious pancreatic necrosis. *Aquac. Res.* 39, 498–505.
412 Kristoffersen, A. B., Devold, M., Aspehaug, V., Gjelstenli, O., Breck, O., and Jensen, B. B.
413 (2018). Molecular tracing confirms that infection with infectious pancreatic necrosis virus
414 follows the smolt from hatchery to grow-out farm. *J. Fish Dis.* 41, 1601–7.
415 Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular
416 evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–9.
417 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., et
418 al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–8.
419 Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., et al. (2016). The
420 Atlantic salmon genome provides insights into rediploidization. *Nature* 533, 200–5.
421 Liu, M., and Vakharia, V. N. (2004). VP1 protein of infectious bursal disease virus modulates
422 the virulence in vivo. *Virology* 330, 62–73.
423 LMD (Landbruks- og matdepartementet) (1991). Forskrift om fortegnelser over sjukdommer
424 som omfattes av midlertidig lov mot sjukdom hos akvatiske organismer. Norway:
425 Norwegian Ministry of Agriculture and Food.
426 M'Gonigle, R. H. (1941). Acute Catarrhal Enteritis of Salmonid Fingerlings. *Trans. Am. Fish.*
427 *Soc.* 70, 297–303.
428 Moen, T., Baranski, M., Sonesson, A. K., and Kjøglum, S. (2009). Confirmation and fine-
429 mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon
430 (Salmo salar): Population-level associations between markers and trait. *BMC Genomics* 10,
431 368.
432 Moen, T., Torgersen, J., Santi, N., Davidson, W. S., Baranski, M., Ødegård, J., et al. (2015).

433 Epithelial cadherin determines resistance to infectious pancreatic necrosis virus in Atlantic
434 salmon. *Genetics* 200, 1313–26.
435 Munang'Andu, H. M., Fredriksen, B. N., Mutoloki, S., Dalmo, R. A., and Evensen, Ø. (2013).
436 Antigen dose and humoral immune response correspond with protection for inactivated
437 infectious pancreatic necrosis virus vaccines in Atlantic salmon (Salmo salar L). *Vet. Res.*
438 44, 7.
439 Munang'andu, H. M., Santi, N., Fredriksen, B. N., Løkling, K. E., and Evensen, Ø. (2016). A
440 systematic approach towards optimizing a cohabitation challenge model for infectious
441 pancreatic necrosis virus in Atlantic salmon (Salmo salar L). *PLoS One* 11, e0148467.
442 Mutoloki, S., Jøssund, T. B., Ritchie, G., Munang'andu, H. M., and Evensen, Ø. (2016).
443 Infectious pancreatic necrosis virus causing clinical and subclinical infections in atlantic
444 salmon have different genetic fingerprints. *Front. Microbiol.* 7, 1393.
445 NFD (Nærings- og fiskeridepartementet) (2008). Forskrift om drift av akvakulturanlegg.
446 Norway.
447 Nishizawa, T., Kinoshita, S., and Yoshimizu, M. (2005). An approach for genogrouping of
448 Japanese isolates of aquabirnaviruses in a new genogroup, VII, based on the VP2/NS
449 junction region. *J. Gen. Virol.* 86, 1973–8.
450 Okonechnikov, K., Golosova, O., Fursov, M., Varlamov, A., Vaskin, Y., Efremov, I., et al.
451 (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* 28, 1166–7.
452 Pedersen, T., Skjesol, A., and Jørgensen, J. B. (2007). VP3, a Structural Protein of Infectious
453 Pancreatic Necrosis Virus, Interacts with RNA-Dependent RNA Polymerase VP1 and with
454 Double-Stranded RNA. *J. Virol.* 81, 6652–63.
455 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA

ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–6.

Reid, G. K., Gurney-Smith, H. J., Marcogliese, D. J., Knowler, D., Benfey, T., Garber, A. F., et al. (2019). Climate change and aquaculture: Considering biological response and resources. *Aquac. Environ. Interact.* 11, 569–602.

Roberts, R. J., and Pearson, M. D. (2005). Infectious pancreatic necrosis in Atlantic salmon, Salmo salar L. *J. Fish Dis.* 28, 383–390.

Santi, N., Song, H., Vakharia, V. N., and Evensen, Ø. (2005). Infectious Pancreatic Necrosis Virus VP5 Is dispensable for virulence and Persistence. *J. Virol.* 79, 9206–16.

Santi, N., Vakharia, V. N., and Evensen, Ø. (2004). Identification of putative motifs involved in the virulence of infectious pancreatic necrosis virus. *Virology* 322, 31–40.

Selvam, D. G., Mujeeb Rahiman, K. M., and Mohamed Hatha, A. A. (2012). An Investigation into occasional white spot syndrome virus outbreak in traditional paddy cum prawn fields in India. *Sci. World J.*, 340830.

Shivappa, R. B., Song, H., Yao, K., Aas-Eng, A., Evensen, and Vakharia, V. N. (2004). Molecular characterization of Sp serotype strains of infectious pancreatic necrosis virus exhibiting differences in virulence. *Dis. Aquat. Organ.* 61, 23–32.

Skjesol, A., Skjæveland, I., Elnæs, M., Timmerhaus, G., Fredriksen, B. N., Jørgensen, S., et al. (2011). IPNV with high and low virulence: Host immune responses and viral mutations during infection. *Virol. J.* 8, 396.

Smail, D. A., Bain, N., Bruno, D. W., King, J. A., Thompson, F., Pendrey, D. J., et al. (2006). Infectious pancreatic necrosis virus in Atlantic salmon, Salmo salar L., post-smolts in the Shetland Isles, Scotland: Virus identification, histopathology, immunohistochemistry and

479 genetic comparison with Scottish mainland isolates. *J. Fish Dis.* 29, 31–41.

480 Sommerset, I., Walde, C. S., Bang Jensen, B., Bornø, B., Haukaas, A., and Brun, E. (2020).
481 Fiskehelserapporten 2019. Norway.

482 Song, H., Santi, N., Evensen, Ø., and Vakharia, V. N. (2005). Molecular determinants of
483 Infectious Pancreatic Necrosis Virus virulence and cell culture adaptation. *J. Virol.* 79,
484 10289–99.

485 Sundh, H., Olsen, R. E., Fridell, F., Gadan, K., Evensen, O., Glette, J., et al. (2009). The effect of
486 hyperoxygenation and reduced flow in fresh water and subsequent infectious pancreatic
487 necrosis virus challenge in sea water, on the intestinal barrier integrity in Atlantic salmon,
488 Salmo salar L. *J. Fish Dis.* 32, 687–98.

489 Tacken, M. G. J., Peeters, B. P. H., Thomas, A. A. M., Rottier, P. J. M., and Boot, H. J. (2002).
490 Infectious Bursal Disease Virus Capsid Protein VP3 Interacts both with VP1, the RNA-
491 Dependent RNA Polymerase, and with Viral Double-Stranded RNA. *J. Virol.* 89, 11165–
492 68.

493 Tamura K, Nei M, and Kumar S (2004). Prospects for inferring very large phylogenies by using
494 the neighbor-joining method. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11030–5.

495 Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: Discovering splice junctions with
496 RNA-Seq. *Bioinformatics* 25, 1105–11.

497 Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential
498 gene and transcript expression analysis of RNA-seq experiments with TopHat and
499 Cufflinks. *Nat. Protoc.* 7, 562–78.

500 Wetten, M., Aasmundstad, T., Kjøglum, S., and Storset, A. (2007). Genetic analysis of resistance
501 to infectious pancreatic necrosis in Atlantic salmon (Salmo salar L.). *Aquaculture* 272, 111–

502 7.

503 Wolf, K., Snieszko, S. F., Dunbar, C. E., and Pyle, E. (1960). Virus Nature of Infectious

504 Pancreatic Necrosis in Trout. *Proc. Soc. Exp. Biol. Med.* 104, 25743.

505

506 **Figure Legends**

507 **Figure 1.** A. Histology section of the pancreatic tissue infected with the infectious pancreatic
508 necrosis virus (IPNV). Arrows indicate acute necrosis of pancreatic acinar cells. B.
509 Immunohistochemical staining of the pancreas tissue using antibodies specific for IPNV. Red
510 colours show cells infected with this virus (arrow).

511 **Figure 2.** Neighbour-Joining tree, showing the phylogenetic relationship between the A.
512 Polyprotein and B. VP5's deduced amino acid data from different IPNV reference strains, and the
513 assemblies reported in this study. The accession ids of the sequences are presented in Table 2. The
514 associated genogroups are in parentheses. BG 1 to 6 represent isolates identified in the current
515 work. The confidence of association between sequences was estimated using bootstrap testing
516 (1000 replicates) and is shown next to the branches. The Maximum Composite Likelihood method
517 (Tamura K et al., 2004) was used to estimate the isolates' evolutionary distances.

518

519 **Table 1.** Amino acid sequences in segment A of the newly identified IPNV isolate compared with
520 previously reported key residues suggested being correlated with viral virulence (Santi et al.,
521 2004).

Accession	Isolate	217	221[a]	247	252	500	Virulence
AY374435.1	NVI-001	T	A/T	T	V	Y	avirulent
AY379742.1	NVI-016	P	T	A	N	Y	avirulent
AY379744.1	NVI-010	P	T	A	N	H	moderate
AY379735.1	NVI-011	T	T/A	A	V	H	virulent
AY379738.1	NVI-013	T	A/T	T	V	Y	virulent
AY379740.1	NVI-015	T	A/T	T	V	Y	virulent
AY379736.1	NVI-020	T	T/A	A	V	H	virulent
AY379737.1	NVI-023	T	A/T	T	V	Y	virulent
-	BG	P	T	A	D	Y	-

522 [a] Double peak on the chromatograms after two cell culture passages.
523 The dominating amino acid is indicated first.
524

525 **Table 2.** Amino acid residues and their positions in the polyprotein region of the IPNV isolates
526 reported in this study that are either exclusively unique or occur with a very low frequency in other
527 strains.

Num.	Accession	ID	245	248	252	255	257	278	282	285	321	585
1	-	BG	**G**	**R**	**D**	**T**	**H**	**A**	**T**	**H**	**D**	**G**
2	AJ829474	Sp_975/99	**S**	**E**	**I**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
3	AY379735	Sp_NVI-011	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
4	AY379738	Sp_NVI-013	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
5	AY379740	Sp_NVI-015	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
6	AY379736	Sp_NVI-020	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
7	AY379737	Sp_NVI-023	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
8	AY379744	Sp_NVI-010	**S**	**E**	**N**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
9	AY374435	Sp_NVI-001	**S**	**E**	**V**	**K**	**D**	**V**	**N**	**Y**	**G**	**D**
10	AY379742	Sp_NVI-016	**S**	**E**	**N**	**K**	**D**	**V**	**D**	**Y**	**G**	**D**
11	AF343573	Buhl	**Q**	**A**	**N**	**V**	**D**	**T**	**V**	**H**	**A**	**D**
12	AF342729	Ab	**S**	**N**	**A**	**K**	**D**	**V**	**A**	**Y**	**A**	**D**
13	AF342730	He	**N**	**Q**	**N**	**K**	**D**	**V**	**E**	**Y**	**G**	**D**
14	AF342731	Te	**R**	**A**	**A**	**K**	**D**	**V**	**A**	**F**	**T**	**D**
15	AF342732	C1	**S**	**R**	**T**	**R**	**D**	**V**	**A**	**F**	**T**	**D**
16	AF342733	C2	**S**	**T**	**A**	**T**	**E**	**A**	**K**	**H**	**G**	**D**
17	AF342734	C3	**S**	**N**	**A**	**T**	**K**	**A**	**T**	**H**	**A**	**D**
18	AF342735	Jasper	**Q**	**A**	**N**	**R**	**D**	**A**	**A**	**H**	**G**	**D**
19	AF342727	WB	**Q**	**A**	**N**	**R**	**D**	**T**	**A**	**Y**	**G**	**D**

528

529

Supplementary Figure Legends

Supplementary Figure 1. Neighbour-Joining trees, showing the phylogenetic relationships between partial polyprotein nucleotide (A) and the amino acid (B) sequences of the reported assemblies with an isolate (MH562009) previously recovered from a field outbreak (Kristoffersen et al., 2018). BG 1 to 6 represent assembled sequences in the current study. The distances were computed using the Maximum Composite Likelihood method, for nucleotides and Poisson for the amino acid data. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). There were a total of 1513 nucleotides and 504 amino acids in the final dataset. Evolutionary analyses were conducted in MEGA X.

Supplementary Figure 2. Neighbour-Joining trees, showing the phylogenetic relationships between the assembled polyprotein (A) and VP1 (B) segments of the IPNV genomes for the six samples sequenced in this study. The trees are constructed based on the amino acid sequence information. The confidence of association, estimated using bootstrap testing (1000 replicates), is shown next to the branches. There were 2919 and 845 positions in the final datasets for the polyprotein and VP1 respectively. Analyses were conducted in MEGA X.

547 **Supplementary Table 1.** Overview of the laboratory results, confirming the initial diagnosis of the IPNV infection and QTL status of the fish.

Num.	tissue	QTL	qRT-PCR Ct value	histology	Immunohistochemistry
1	adipose fin	Yes	Not tested	Not tested	Not tested
2	adipose fin	Yes	Not tested	Not tested	Not tested
3	adipose fin	Yes	Not tested	Not tested	Not tested
4	adipose fin	Yes	Not tested	Not tested	Not tested
5	adipose fin	Yes	Not tested	Not tested	Not tested
6	adipose fin	Yes	Not tested	Not tested	Not tested
7	adipose fin	Yes	Not tested	Not tested	Not tested
8	adipose fin	Yes	Not tested	Not tested	Not tested
9	adipose fin	Yes	Not tested	Not tested	Not tested
10	adipose fin	Yes	Not tested	Not tested	Not tested
11	kidney	Not tested	20.32	Not tested	Not tested
12	kidney	Not tested	19.79	Not tested	Not tested
13	kidney	Not tested	18.93	Not tested	Not tested
14	kidney	Not tested	18.92	Not tested	Not tested
15	kidney	Not tested	20.77	Not tested	Not tested
16	kidney	Not tested	19.32	Not tested	Not tested
17	multiple	Not tested	Not tested	Multifocal necrosis in the exocrine pancreas	positive
18	multiple	Not tested	Not tested	Multifocal necrosis in exocrine pancreas and multifocal bleeding in abdominal fat	positive
19	multiple	Not tested	Not tested	Multifocal necrosis in exocrine pancreas and focal, extensive necrosis in the liver	positive

548

Supplementary Table 2. Estimates of evolutionary divergence rate between the amino acid (A) and nucleotide (B) sequences of the polyprotein of the assembled genomes. The actual number of differences are provided in the parentheses. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1).

A	BG_1_poly	BG_2_poly	BG_3_poly	BG_4_poly	BG_5_poly	BG_6_poly
BG_1_poly	0 (0)	0 (0)	0 (0)	0 (0)	0.00206 (2)	0.00206 (2)
BG_2_poly		0 (0)	0 (0)	0 (0)	0.00206 (2)	0.00206 (2)
BG_3_poly			0 (0)	0 (0)	0.00206 (2)	0.00206 (2)
BG_4_poly				0 (0)	0.00206 (2)	0.00206 (2)
BG_5_poly					0 (0)	0 (0)
BG_6_poly						0 (0)

B	BG_1_poly	BG_2_poly	BG_3_poly	BG_4_poly	BG_5_poly	BG_6_poly
BG_1_poly	0 (0)	0 (0)	0 (0)	0 (0)	0.00308 (9)	0.00308 (9)
BG_2_poly		0 (0)	0 (0)	0 (0)	0.00308 (9)	0.00308 (9)
BG_3_poly			0 (0)	0 (0)	0.00308 (9)	0.00308 (9)
BG_4_poly				0 (0)	0.00308 (9)	0.00308 (9)
BG_5_poly					0 (0)	0 (0)
BG_6_poly						0 (0)

Supplementary Table 3. Estimates of evolutionary divergence rate between the amino acid (A) and nucleotide (B) sequences of the VP1 of the assembled genomes. The actual number of differences are provided in the parentheses. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1).

A	**BG_1_VP1**	**BG_2_VP1**	**BG_3_VP1**	**BG_4_VP1**	**BG_5_VP1**	**BG_6_VP1**
BG_1_VP1	0 (0)	0.00118 (1)	0 (0)	0.00118 (1)	0.00237 (2)	0.00237 (2)
BG_2_VP1		0 (0)	0.00118 (1)	0 (0)	0.00237 (2)	0.00237 (2)
BG_3_VP1			0 (0)	0.00118 (1)	0.00237 (2)	0.00237 (2)
BG_4_VP1				0 (0)	0.00237 (2)	0.00237 (2)
BG_5_VP1					0 (0)	0 (0)
BG_6_VP1						0 (0)

B	BG_1_VP1	BG_2_VP1	BG_3_VP1	BG_4_VP1	BG_5_VP1	BG_6_VP1
BG_1_VP1	0 (0)	0.00158 (4)	0.000395 (1)	0.00158 (4)	0.00592 (15)	0.00553 (14)
BG_2_VP1		0 (0)	0.00118 (3)	0 (0)	0.00474 (12)	0.00434 (11)
BG_3_VP1			0 (0)	0.00118 (3)	0.00553 (14)	0.00513 (13)
BG_4_VP1				0 (0)	0.00474 (12)	0.00434 (11)
BG_5_VP1					0 (0)	0.000395 (1)
BG_6_VP1						0 (0)

Supplementary Files

Supplementary File 1. Alignment of the deduced amino acid sequence data from the polyprotein section of the IPNV. BG 1 to 6 represent the six assemblies from the current study.

Supplementary File 2. Alignment of the deduced amino acid sequence data from the VP1 section of the IPNV. BG 1 to 6 represent the six assemblies from the current study.

Supplementary File 3. Alignment of the deduced amino acid sequence data from the VP5 section of the IPNV. BG 1 to 6 represent the six assemblies from the current study.

Figure 1.TIFF

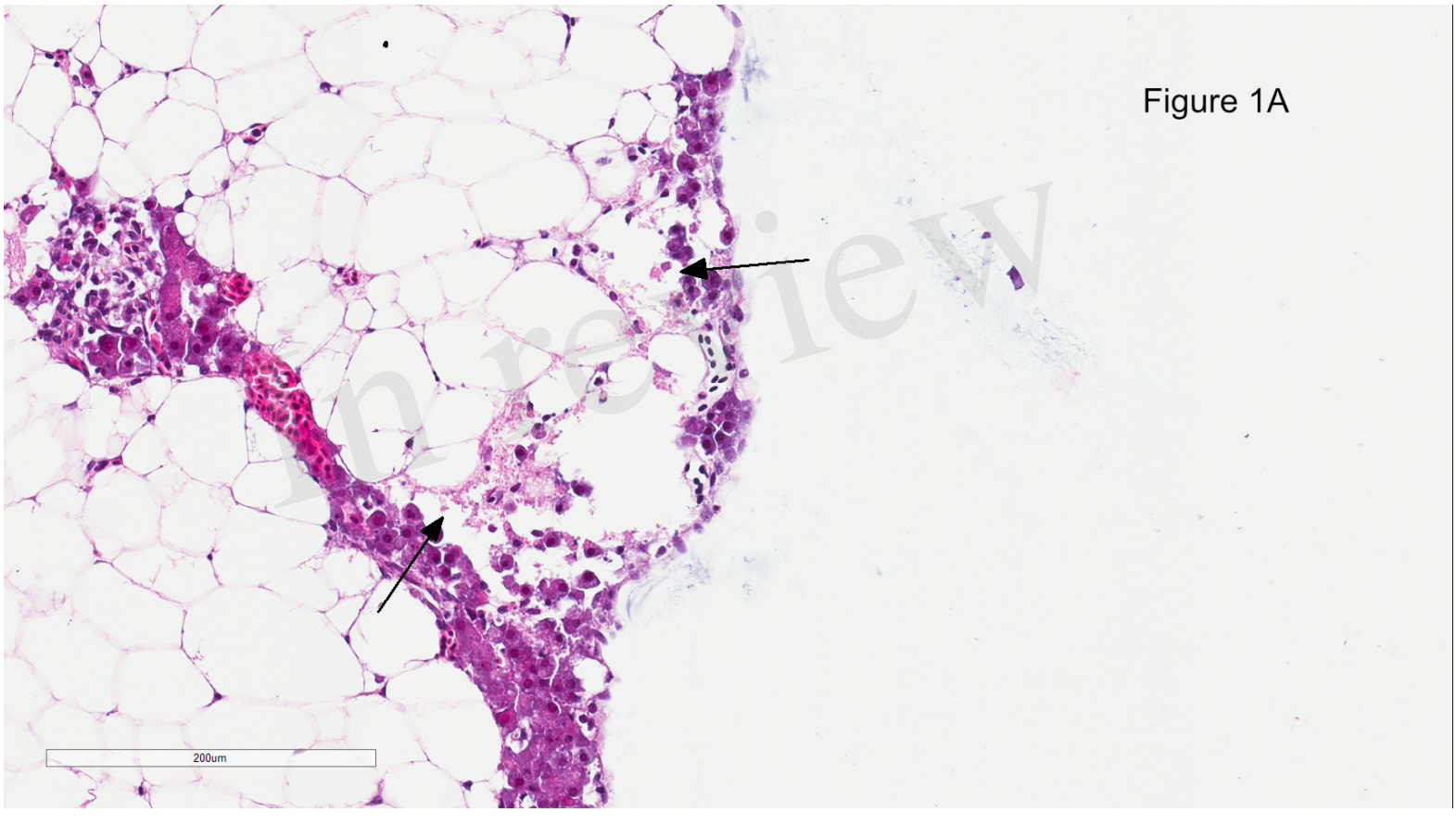

Figure 2.TIFF

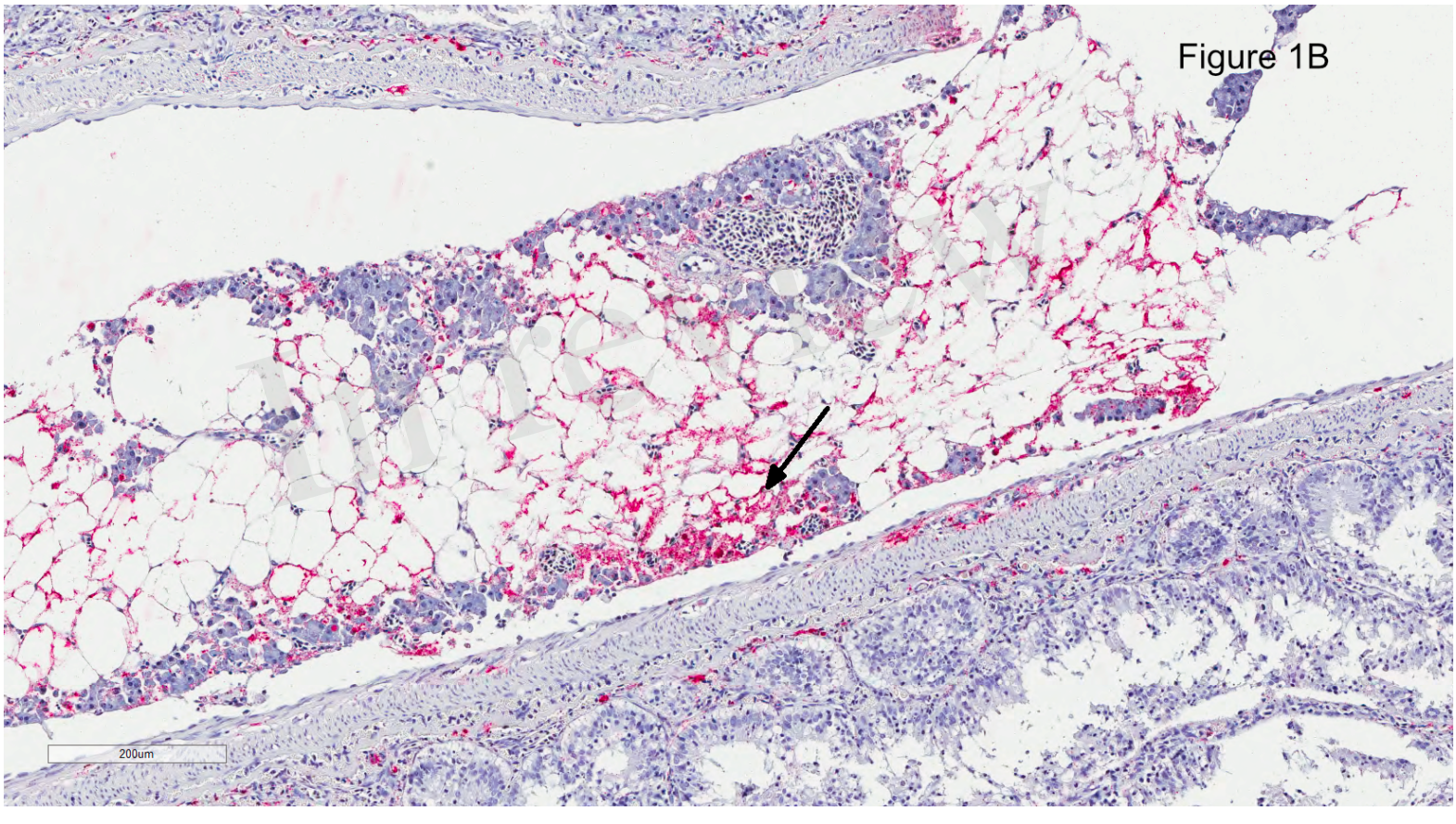

Figure 3.TIFF

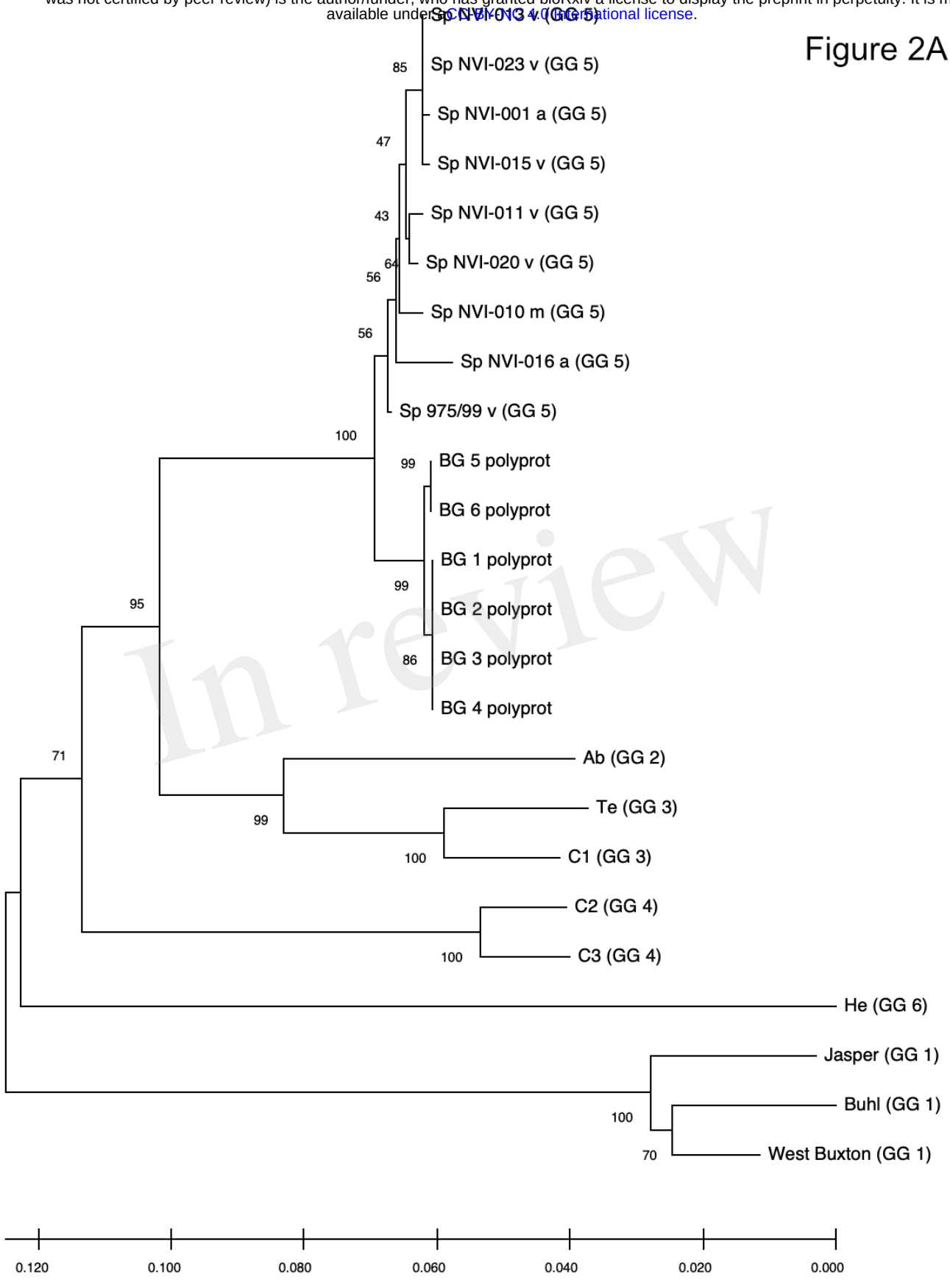

Figure 4.TIFF

Figure 2B