

Estimation of the repeatability and reproducibility of three diagnostic tests for infectious salmon anaemia virus

P N  rette¹, I Dohoo¹, L Hammell¹, N Gagn  ², P Barbash³, S MacLean⁴ and C Yason⁵

1 Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PEI, Canada

2 Fisheries and Oceans Canada, Molecular Biology, Aquaculture Division, Moncton, NB, Canada

3 Fish Health Center, Northeast Fishery Center, U.S. Fish and Wildlife Service, Lamar, PA, USA

4 NOAA/National Marine Fisheries Service, Narragansett, RI, USA

5 Department of Pathology and Microbiology, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PEI, Canada

Abstract

Reverse transcriptase polymerase chain reaction (RT-PCR), virus isolation (VI) and indirect fluorescent antibody tests (IFAT) are three assays currently used by the salmon industry to identify fish infected with infectious salmon anaemia virus (ISAV). However, no data are available on the repeatability (within-laboratory consistency) and reproducibility (between-laboratory consistency) of these assays and very limited information is available on the effect of freezing samples on test results. In order to evaluate these assays, five laboratories participated in a blinded study of 400 kidney samples representing four populations of farmed Atlantic salmon with different prevalence of ISAV. Each laboratory used its own testing protocols. Repeatability and reproducibility were evaluated using kappa as the measure of agreement. The effect of freezing was evaluated using the McNemar test. Freezing did not affect VI but improved the sensitivity of RT-PCR. The repeatability and reproducibility of VI was almost perfect. There was a substantial difference in repeatability of RT-PCR among the three laboratories with kappa ranging from 0.5 to 0.96. The repeatability for RT-PCR was generally low. The repeatability of IFAT was moderate when the results were analysed using 1 +

and above as a positive result. The results of the study show the need to standardize the protocol and interpretation of RT-PCR.

Keywords: diagnostic tests, indirect fluorescent antibody testing, infectious salmon anaemia virus, repeatability, reproducibility, reverse transcriptase polymerase chain reaction.

Introduction

Infectious salmon anaemia (ISA) is an infectious disease of Atlantic salmon, *Salmo salar* L., caused by a newly identified Orthomyxovirus genus *Isavirus* (Anonymous 2002). Initially reported in Norway in 1984 (Thorud & Djupvik 1988) ISA has since been found in Canada (Byrne, Macphee, Ostland, Johnson & Ferguson 1998), Scotland (Rodger, Turnbull, Muir, Millar & Richards 1998), the Faroe Islands (Anonymous 2000) and more recently in the US (Bouchard, Brockway, Giray, Keleher & Merrill 2001). Fish affected by the virus show anaemia, congestion of the liver, spleen and foregut, and haemorrhagic liver necrosis (Evensen & Thorud 1991).

Transmission of the virus from fish to fish can be achieved with skin mucus, faeces, urine and blood (Totland, Hjeltnes & Flood 1996). The gills are the most likely port of entry of the virus (Totland *et al.* 1996). It is hypothesized that ISA may have emerged in farmed Atlantic salmon when mutated isolates were transmitted from wild salmonids or following mutation of benign isolates in farmed

Correspondence P N  rette, Atlantic Veterinary College, Department of Health Management, 550 University Avenue, Charlottetown, Prince Edward Island, Canada C1A 4P3 (e-mail: pnerette@upei.ca)

salmon (Nylund, Devold, Plarre, Isdal & Aarseth 2003).

Three diagnostic assays are commonly used for the detection of infectious salmon anaemia virus (ISAV). These include virus isolation (VI) on the salmon head kidney (SHK-1) cell line (Dannevig, Falk & Namork 1995), reverse transcriptase polymerase chain reaction (RT-PCR) (Mjaaland, Rimstad, Falk & Dannevig 1997) and indirect fluorescent antibody testing (IFAT) (Falk & Dannevig 1995; Falk, Namork & Dannevig 1998).

In New Brunswick, the current ISAV management programme is based on early detection and removal of ISA-infected stocks on a cage-by-cage basis. Currently, IFAT tests on moribund fish samples are used for preliminary screening. Both RT-PCR and viral tissue culture are then used to confirm positive IFAT results. The screening with IFAT is carried out on fresh samples but confirmation with RT-PCR and VI is often carried out on frozen samples. Samples are stored at $-20\text{ }^{\circ}\text{C}$ for RT-PCR (up to 7 weeks) and at $-80\text{ }^{\circ}\text{C}$ for VI (up to 4 weeks) until they can be processed.

The surveillance programme results in early slaughter of a cage if there have been two positive tests on at least two fish. Until recently clinical signs and elevated mortality rates were considered for decisions on most slaughterings. However, recent modifications to the programme have reduced the emphasis on mortality and clinical signs in attempts to remove sources of virus much earlier in an outbreak. These changes will put more reliance on diagnostic tests and their ability to correctly detect negative and positive cages.

It is important to know the precision of these diagnostic assays currently used for disease control and research programmes by the aquaculture industry in Canada. However, to date, no study has evaluated the repeatability (within-laboratory consistency) and reproducibility (between-laboratory consistency) of these diagnostic assays. It is also important to know the effect of freezing on RT-PCR and VI as samples are often stored at $-20\text{ }^{\circ}\text{C}$ (RT-PCR) or $-80\text{ }^{\circ}\text{C}$ (VI) before they are processed for diagnostic (i.e. until the results of preliminary screening assays are available) and research purposes ($-80\text{ }^{\circ}\text{C}$) (i.e. studies requiring randomization/blinding of samples are more able to utilize frozen samples).

The overall objective of the research programme, of which this study was one component, was to determine the operating characteristics of current

and recently developed diagnostic assays for detecting ISAV. The specific objectives of this study were to evaluate the repeatability and reproducibility of currently used diagnostic assays and to determine how freezing can influence the repeatability of the test results.

Materials and methods

Study material

Details of animal selection, tissue sampling, storage and distribution have been published previously (N  rette, Dohoo & Hammell 2005).

Multiple samples were taken from each fish, which were submitted to several laboratories for different tests. Assay methods evaluated included the IFAT (one laboratory; IFATa), RT-PCR (three laboratories; PCRa, PCRb, PCRC) and VI (two laboratories; VIa, VIb). For PCRb the first 75 samples were initially analysed using a different PCR assay. After several months of storage at $-20\text{ }^{\circ}\text{C}$, the PCR reaction was performed again using the assay described below. The results of this second analysis were inconsistent due to the partial degradation of RNA and cDNA over time at $-20\text{ }^{\circ}\text{C}$. Consequently, these first 75 samples were not included in the analysis.

Freezing effect study

Fifty-two duplicate samples of moribund fish (10 from each outbreak except one outbreak where 12 duplicates were taken) were used to evaluate the effect of freezing on RT-PCR and VI. One set of samples was sent to one laboratory to be processed within 24 h of collection while the other set was stored at $-80\text{ }^{\circ}\text{C}$ for later evaluation by the same laboratory.

Repeatability study

Duplicate samples were submitted to the laboratories to evaluate repeatability; these consisted of: (i) moribund salmon ($n \geq 15$), (ii) healthy salmon from an outbreak cage ($n \geq 15$), and (iii) healthy salmon from a healthy cage ($n \geq 20$). The samples were selected randomly from each of these three populations and the complete set (ranging from 20 to 96 samples) was tested by each of the assay methods evaluated. The lower number of samples analysed at PCRb was because 37 of 75 samples

that could not be included in the analysis were part of the repeatability study.

Reproducibility/agreement study

A full set of 400 samples (100 from each of the four populations) was tested by each assay being evaluated. Test results were used to evaluate reproducibility (agreement between the same assays performed at different laboratories) and agreement between different assays. Eighty salmon had insufficient individual tissue samples collected, thus some laboratories received fewer than 400 samples.

Methods for ISAV detection used by the five participating laboratories

The participating laboratories were asked to determine the presence or absence of ISAV in each of the 400 samples according to their own procedures. Details of the methods used for each test evaluated are presented elsewhere (N  rette *et al.* 2005).

Statistical analysis

The data were stored, edited and manipulated using Version 7.0 of the statistical package Stata (Stata Corp., College Station, TX, USA, 2001). Kappa was used to measure the level of agreement beyond chance. It ranges from -1 to $+1$. Common interpretations of kappa, when applied to tests that are subjective in nature (i.e. radiographic interpretation), are as follows: a value of 0 represents no agreement, values of 0–0.20 indicate slight agreement, values of 0.21–0.40 indicate fair agreement, values of 0.41–0.60 indicate moderate agreement, values of 0.61–0.80 indicate substantial agreement, values of 0.81–1.00 indicate almost perfect agreement, and a value of $+1$ represents perfect agreement (Dohoo, Martin & Stryhn 2003). Before assessing kappa we assessed if there was a serious disagreement between the tests by determining whether the proportion positive to each test was different. Because the data were paired this was assessed by McNemar's test or an exact binomial test for correlated proportions. A non-significant χ^2 indicates that the proportions positive did not differ. A significant result suggests a serious disagreement between the tests and thus the detailed assessment of agreement was of little value (Dohoo *et al.* 2003).

Results

Freezing effect

Duplicate samples were taken from 52 moribund fish to evaluate the effect of freezing on RT-PCR and VI.

Reverse transcriptase polymerase chain reaction

The proportion positive on the frozen samples (94.2%, 49/52) was higher than the proportion positive on the fresh samples (80.8%, 42/52) and the difference was statistically significant (McNemar χ^2 : $P = 0.02$) (Table 1). This discordance between fresh and frozen samples was observed in seven pairs of samples. Five of seven tested positive on VI, which suggests that they were misclassified as negative by RT-PCR on the fresh samples. Two of seven tested negative on all other assays, which suggests that they were misclassified by RT-PCR as positive on the frozen samples (Table 2).

Virus isolation

The proportion of positives found on the fresh samples (88.5%, 46/52) was slightly higher than the proportion of positives found on the frozen samples (80.8%, 42/52) but this difference was not statistically significant (McNemar χ^2 : $P = 0.13$) (Table 1). This discordance between fresh and frozen samples was observed on four kidney samples (Table 3).

Repeatability

Duplicate samples were taken from three different populations for evaluating the repeatability.

Table 1 Contingency table of fresh and frozen samples for evaluation of the effect of freezing on virus isolation (VI) and reverse transcriptase polymerase chain reaction (RT-PCR) based on 52 samples from a moribund (high prevalence) population

Fresh	Frozen		McNemar χ^2 <i>P</i> -value
	Positive	Negative	
RT-PCR			
Positive	42	0	0.016 ^a
Negative	7	3	
VI			
Positive	42	4	0.125
Negative	0	6	

^a McNemar's χ^2 significant ($P < 0.05$) indicating different proportions positive on the two tests.

Table 2 Summary of diagnostic test results from discordant samples on fresh and frozen samples for evaluation of freezing on reverse transcriptase polymerase chain reaction (RT-PCR)

Kidney	Fresh PCRa	Frozen PCRa	Other test results ^a				IFATa
			PCRB	PCRC	VIa	VIb	
1	-	+	-	-	-	-	-
2	-	+	-	-	+	-	-
3	-	+	-	+	+	-	-
4	-	+	+	+	+	+	2+
5	-	+	+	+	+	+	-
6	-	+	-	-	-	-	-
7	-	+	-	-	+	+	-

PCRa, PCRB, PCRC = RT-PCR performed at laboratories A, B and C respectively; VIa, VIb = virus isolation performed at laboratories A and B respectively; IFATa = indirect fluorescent antibody testing performed at laboratory A.

^a Performed on frozen samples.

Table 3 Summary of diagnostic test results from discordant samples on fresh and frozen samples for evaluation of freezing on virus isolation

Kidney	Fresh VIa	Frozen VIa	Other test results ^a				IFATa
			VIb	PCRa	PCRB	PCRC	
1	+	-	-	-	-	-	-
2	+	-	-	+	-	-	-
3	+	-	+	+	-	-	-
4	+	-	+	+	-	-	-

VIa, VIb = virus isolation performed at laboratories A and B respectively; PCRa, PCRB, PCRC = RT-PCR performed at laboratories A, B and C respectively; IFATa = indirect fluorescent antibody testing performed at laboratory A.

^a Performed on frozen samples.

Reverse transcriptase polymerase chain reaction

There was a substantial difference in the repeatability of RT-PCR tests among the three laboratories with kappa ranging from 0.5 (moderate agreement) to 0.96 (almost perfect agreement) (Table 4). Discordant test results are shown in Table 5. For example, fish number 8 tested positive the first time but negative the second time. The results from the other laboratories' assays were all negative suggesting that the discordance was probably due to a false-positive result on the first test. Based on all other test results it appears that most of the discordant pairs arose from one sample being a false positive.

Virus isolation

Repeatabilities were very high in both laboratories (kappa = 0.85 and 0.87) (Table 4). Discordant test results are given in Table 5. Most of the discordant results appear to have risen from one sample being a false negative.

Indirect fluorescent antibody test

The repeatability was almost perfect when the agreement was evaluated using a weighted kappa (Dohoo *et al.* 2003) on the original ordinal data (results not presented). When the agreement was evaluated on dichotomized data the agreement depended on the cut-off point selected. When the IFAT results were analysed using 1+ and higher as

Test	Laboratory		Duplicate		n	McNemar χ^2 P-value	Kappa
			Positive	Negative			
RT-PCR	A	Positive	35	7	63	1	0.5 (0.25, 0.75)
		Negative	7	14			
RT-PCR	B	Positive	5	1	20	1	0.76 (0.32, 1.00)
		Negative	1	13			
RT-PCR	C	Positive	19	0	50	1	0.96 (0.68, 1.00)
		Negative	1	30			
VI	A	Positive	13	2	50	1	0.85 (0.58, 1.00)
		Negative	1	34			
VI	B	Positive	17	1	49	1	0.87 (0.59, 1.00)
		Negative	2	29			
IFAT1	A	Positive	60	10	96	0.09	0.68 (0.49, 0.88)
		Negative	3	23			
IFAT2	A	Positive	18	3	96	1	0.85 (0.65, 1.00)
		Negative	2	73			

RT-PCR, reverse transcriptase polymerase chain reaction; VI, virus isolation; IFAT1, indirect fluorescent antibody testing if 1+ was used as the cut-point; IFAT2, indirect fluorescent antibody testing if 2+ was used as the cut off-point.

Table 4 Contingency table of duplicates for evaluation of repeatability on virus isolation, RT-PCR and IFAT (kappa: 95% confidence intervals in brackets)

Table 5 Summary of diagnostic test results from discordant samples on virus isolation from different laboratories for evaluation of the repeatability of virus isolation, RT-PCR and IFAT

ID	PCRa		PCRb		PCRC		VIa		VIb		IFATa	
	O	D	O	D	O	D	O	D	O	D	O	D
1	+	-	-	-	-	-	-	-	-	-	-	-
2	-	+	-	-	-	-	-	-	-	-	-	-
3	-	+	+	-	-	-	-	-	-	-	-	-
4	-	+	-	-	-	-	-	-	-	-	-	-
5	-	+	-	-	-	-	-	-	-	-	-	-
6	+	-	-	-	-	-	-	-	-	-	-	-
7	+	-	-	-	-	-	-	-	-	-	-	-
8	+	-	-	-	-	-	-	-	-	-	-	-
9	+	-	-	-	-	-	-	-	-	-	-	-
10	-	+	+	-	-	-	-	-	-	-	-	-
11	-	+	-	-	-	-	-	-	-	-	-	-
12	+	-	-	-	-	-	-	-	-	-	-	-
13	-	+	-	-	-	-	-	-	-	-	-	-
14	+	-	-	-	-	-	-	-	-	-	-	-
15	+	-	-	+	+	-	-	-	-	-	-	1+
16	-	-	+	-	-	-	-	-	-	-	-	-
17	+	-	-	-	+	-	-	-	-	-	-	1+
18	+	+	+	+	-	-	+	-	-	-	-	-
19	+	+	+	+	+	-	-	-	-	-	-	-
20	+	-	-	+	+	-	-	-	-	-	-	-
21	+	+	+	+	-	-	-	+	-	-	-	-
22	+	+	+	+	+	-	+	-	+	-	4+	-
23	+	-	-	+	-	-	-	-	+	-	-	-
24	+	+	+	+	+	+	+	+	+	+	1+	-
25	+	+	+	+	+	+	+	+	+	+	-	1+
26	+	-	-	-	-	-	-	-	-	-	-	1+
27	+	+	+	+	-	-	-	-	-	-	-	1+
28	+	+	-	-	+	+	+	+	+	+	1+	-
29	+	-	+	+	-	-	-	-	-	-	-	1+
30	-	+	-	-	-	-	-	-	-	-	-	1+
31	-	-	-	-	-	-	-	-	-	-	1+	-
32	-	-	-	-	-	-	-	-	-	-	-	1+
33	+	+	+	+	+	+	+	+	+	+	-	1+
34	+	-	-	-	-	-	-	-	-	-	-	1+
35	-	-	-	-	-	-	-	-	-	-	-	1+
36	-	-	-	-	-	-	-	-	-	-	-	1+

PCRa, PCRb, PCRC = RT-PCR performed at laboratories A, B and C respectively; VIa, VIb = virus isolation performed at laboratories A and B respectively; IFATa = indirect fluorescent antibody testing performed at laboratory A; O, original; D, duplicate; RT-PCR, reverse transcriptase polymerase chain reaction; IFAT, indirect fluorescent antibody testing.

^a Positive result noted as coming from a weak band.

a positive result (IFAT1) the agreement was substantial ($\kappa = 0.68$), while when the results were analysed using negative and 1+ as a negative result and 2+ and higher as a positive result (IFAT2) the agreement was almost perfect ($\kappa = 0.85$) (Table 4). Nine of thirteen discordant duplicates on IFAT1 tested positive on at least one other assay (Table 5) suggesting that most of the discordant results were due to one sample giving a false-negative result.

Reproducibility/agreement

Reproducibility of RT-PCR

One laboratory (PCRa) produced significantly (McNemar's χ^2 : $P < 0.001$) more positive results on RT-PCR than the other two laboratories (Table 6). Consequently, computation of kappas for a more detailed assessment of agreement among these laboratories was of little value. Of 71 fish which tested positive on PCRa, but negative on either or both of PCRb and PCRC, 24 had at least one positive VI test result suggesting that the discordance was a mixture of both false-positive results from PCRa and false-negative results from the other two RT-PCR tests, although the former appeared to be more common (complete listing of discordant results not presented). Reproducibility of test results between PCRb and PCRC was moderate ($\kappa = 0.79$) (Table 6).

Most of the disagreement was in the population of healthy salmon from an outbreak cage (HS) and in moribund salmon (M) compared with the population of healthy salmon from non-outbreak cages (HH) (Table 7).

Reproducibility of VI

Virus isolation test results were highly reproducible with 380 of 399 samples tested by both laboratories giving the same test result ($\kappa = 0.88$) (Table 6). Of the 19 fish with discordant results, all except one had at least one other positive test result suggesting that the disagreement was primarily due to occasional false-negative test results (Table 8). All the disagreement was in the HS and M populations (Table 7).

Agreement-all tests

The PCRa had relatively poor agreement with most of the other assays (ranging from 0.47 to 0.66) while PCRb, PCRC, VIa and VIb had good agreement with most of the other assays (ranging from 0.67 to 0.92) (Table 9). The lack of agreement between PCRa and other assays was related to the fact that PCRa tended to classify many more samples as positive.

Discussion

As a measure of agreement, kappa measures the level of agreement beyond chance. Kappa is difficult to interpret as a measure if the prevalence of

Test	Laboratory	Positive	Negative	<i>n</i>	McNemar χ^2 <i>P</i> -value	Kappa	
RT-PCR B	Laboratory A						
	Positive	80	6	332	<0.001 ^a	Serious disagreement	
	Negative	60	186				
Laboratory A							
RT-PCR C	Laboratory A						
	Positive	105	0	317	<0.001 ^a	Serious disagreement	
	Negative	49	163				
Laboratory B							
RT-PCR C	Laboratory A						
	Positive	71	15	277	1	0.79 (0.68, 0.91)	
	Negative	9	182				
Laboratory A							
VI	B	Positive	98	6	399	0.17	0.88 (0.78, 0.98)
		Negative	13	282			

RT-PCR, reverse transcriptase polymerase chain reaction; VI, virus isolation.

^a McNemar's χ^2 highly significant ($P < 0.001$) indicating different proportions positive on the two tests.

Table 7 Summary of the number of discordant samples between assays within each population (row proportions in bracket)

Comparison	Populations, <i>n</i> (%)			
	M	HS	HH	F
PCRa/PCRB	25 (37.9)	19 (28.8)	12 (18.2)	10 (15.1)
PCRa/PCRC	13 (26.5)	18 (36.7)	8 (16.4)	10 (20.4)
PCRB/PCRC	13 (54.1)	10 (41.7)	1 (4.2)	0
VIa/VIb	12 (63.2)	7 (36.8)	0 (0)	0

PCRa, PCRB, PCRC = RT-PCR performed at laboratories A, B and C respectively; VIa, VIb = virus isolation performed at laboratories A and B respectively; M, moribund fish in an outbreak cage; HS, apparently healthy fish in a 'sick' (outbreak) cage; HH, apparently healthy fish in a 'healthy' cage (non-outbreak cage on same site or nearest neighbouring site); F, apparently healthy fish from a population assumed to be free of infectious salmon anaemia.

positive results is either very high or very low (Dohoo *et al.* 2003). However, mixing samples from populations with very high prevalence (moribund fish in outbreak cages) through to zero prevalence (ISA free area) will have ensured that this problem was minimized in this study. Samples were taken in 2000–02 and testing was conducted in 2001–03. This implies that recent modifications to assay protocols were not assessed by this study.

Freezing effect

Because kidney samples are often frozen for later analysis it was important to evaluate the effect of freezing on VI and RT-PCR. The results of the current study showed that freezing does not significantly affect VI but appears to increase the frequency of positive RT-PCR. Other laboratories

Table 6 Contingency table of test results obtained from different laboratories for evaluation of the reproducibility of RT-PCR and virus isolation (kappa: 95% confidence intervals in brackets)

Table 8 Summary of diagnostic test results from discordant samples on virus isolation from different laboratories for evaluation of the reproducibility of virus isolation

Kidney	VIa	VIb	PCRC	PCRa	PCRB	IFATa
1	+	-	-	+	-	0
2	+	-	+	+	+	4+
3	+	-	-	+	-	0
4	+	-	+	+	-	0
5	+	-	+	+	-	0
6	+	-	+	+	+	0
7	+	-	+	+	+	0
8	+	-	-	+	-	0
9	+	-	+	+	+	0
10	+	-	+	+	+	0
11	+	-	+	+	-	0
12	+	-	+	+	-	2+
13	+	-	+	+	+	4+
14	-	+	-	+	-	0
15	-	+	-	+	-	0
16	-	+	+	+	+	0
17	-	+	-	+	-	0
18	-	+	-	+	-	0
19	-	+	-	-	-	0

VIa, VIb = virus isolation performed at laboratories A and B respectively; PCRa, PCRB, PCRC = RT-PCR performed at laboratories A, B and C respectively; IFATa = indirect fluorescent antibody testing performed at laboratory A.

have reported that tissue homogenates may be frozen at -20 °C or lower for up to 3 months without substantial loss of viral recovery (Merrill 2003). The apparent increased sensitivity of RT-PCR when performed on frozen samples might be attributed to the liberation of more viral particles during the freeze/thaw cycle, allowing greater access to target in subsequent tests. It is important to note that the common practice of freezing surveillance

Table 9 Summary of the kappa statistics evaluating the agreement among all pairs of test results

	PCRa	PCRb	PCRc	VIa	VIb	IFAT1a
PCRa	1					
PCRb	n.a.	1				
PCRc	n.a.	0.79	1			
VIa	0.66	0.76	0.92	1		
VIb	0.62	0.74	0.85	0.88	1	
IFAT1a	0.47	0.67	0.74	0.76	0.75	1

PCRa, PCRb, PCRc = RT-PCR performed at laboratories A, B and C respectively; VIa, VIb = virus isolation performed at laboratories A and B respectively; IFATa = indirect fluorescent antibody testing performed at laboratory A; n.a. = not applicable – McNemar's χ^2 highly significant ($P < 0.001$) indicating different proportions positive on the two tests.

samples at $-80\text{ }^\circ\text{C}$ may enhance recovery of positive test results, rather than decrease them. It is not known how long that effect would continue if $-80\text{ }^\circ\text{C}$ frozen samples were stored for longer periods.

Virus isolation

Virus isolation for ISA virus appears to be a highly repeatable and reproducible diagnostic method. There was only one sample where virus was isolated by virus culture but was not detected by any of the RT-PCRs, suggesting that any discordance between laboratories was due to occasional false-negative test results.

The very limited discordance between duplicate samples within the same laboratory could be due to the uneven distribution of the virus in samples obtained from the same organ of a fish. Five of six discordant samples were negative by IFAT suggesting the amount of virus in the discordant sample was low (Table 5). Thus, some kidney samples probably had a sufficient amount of virus to be detectable on VI while others did not.

Factors that might have contributed to the variability between laboratories in VI are the difference in laboratory methods and the uneven distribution of the virus in samples. Sixteen of 19 discordant samples between laboratories were negative by IFAT suggesting the amount of virus in the discordant samples was low (Table 8).

Confirmatory tests are used because non-specific CPE changes can occur (i.e. cytotoxicity, infectious pancreatic necrosis virus). Tests used to confirm the CPE do not seem to influence the sensitivity of virus culture. Only three of 13 discordant samples that were negative at VIb had a CPE and were IFAT-negative. Even if IFAT is less sensitive than

RT-PCR it performs much better on cell lysates because they are very rich in virus particles. Because the virus can often be difficult to grow and may not give a CPE even if present, VIb performed a confirmatory test even on samples that had no CPE. All CPE negative samples were also negative by IFAT suggesting that performing a confirmatory IFAT on samples lacking CPE did not alter the interpretation of the results of the virus culture assay.

Reverse transcriptase polymerase chain reaction

There was a substantial difference in repeatability among the three laboratories. If one assay has poor repeatability the agreement between that assay and any other assay is also likely to be poor. If both assays have poor repeatability, the disagreement is even greater (Bland & Altman 1986).

Factors that may provide an explanation of the observed lack of agreement within or between laboratories include: (i) different laboratory protocols; (ii) the subjectivity of the interpretation; and (iii) the heterogeneity of the virus distribution in tissue samples.

The most important laboratory procedures crucial for the reliability of RT-PCR are: (i) the specificity of the primers; (ii) the inclusion of controls (negative, positive and/or internal controls) that run in parallel with the sample during each analysis; (iii) the organization of the laboratory to avoid contamination between samples from different locations and from amplified products from earlier analysis; and (iv) the laboratory skills of the personnel performing the analysis (Royal Society of Edinburgh 2002). Laboratories used different primers, different methods to avoid contamination, and different controls which were reflected in disagreements among the various laboratories.

Measures of agreement do not give information about which assay is the best or what kind of errors either assay is more likely to produce, but when we compared each assay results with the other assays it was possible to observe some trends. PCRa did not have good agreement with most of the other assays. It was often positive when the other five tests were negative suggesting the assay was either very sensitive and/or had a high rate of false positives. PCRa was also the only laboratory that had positive results (10%) in the population assumed to be free of disease (not illustrated) suggesting these were false positives.

The most commonly reported concern with RT-PCR is the possibility of false-positive results resulting from cross-contamination (Wilson 1997). There were many basic procedural differences between laboratories that are not detailed here, so it is difficult to speculate on whether or not cross-contamination may have been more of a problem in PCRa than PCRb and PCRc. For example, gloves that are not being changed enough times during the process, tubes that are not quickly spun to draw contaminants away from the lids before opening, and air currents in the laboratory are some of the many sources of contamination.

Each laboratory had different steps to interpret the final result when the DNA band on the gel was weak. PCRa recorded it as positive while PCRb and PCRc retested the sample. If the second test was positive, the result was recorded as positive. When the retest was negative the result was recorded as negative at PCRc while a third PCR test was performed at PCRb and the result of this last test was used. If this weak band was the result of contamination, then PCRb and PCRc reduced the risk of false-positive results. However, if this weak band was due to a low virus load in the sample, then PCRa increased the sensitivity of the assay. Most samples that tested positive at PCRb but negative at PCRa or PCRc were weakly positive and tested negative on the other assays suggesting these weak bands were falsely positive.

The virus was probably not equally distributed in samples obtained from the same organ of a fish, thus some kidney samples may have had a sufficient amount of viral RNA to be detected by RT-PCR while others did not. This uneven distribution of the virus probably played a minor role in the repeatability and reproducibility of RT-PCR. If it had played an important role it would have been expected that all laboratories had a low repeatability, which was not the case.

It is very important to know current test performance in field situations of very low prevalence because the surveillance programme is using positive diagnostic tests for decisions in the absence of mortality. Interestingly, most of the disagreement among laboratories was in the medium and high prevalence populations (ranging from 63.2 to 95.8%) compared with the low prevalence population (ranging from 4.2 to 18.2%). A biological explanation of this observed phenomenon is not known.

Issues affecting the use of RT-PCR

Polymerase chain reaction is an assay with potentially high sensitivity and specificity (N  rette *et al.* 2005), has the ability to detect the presence of infecting micro-organisms that may not be identified by conventional methods and is rapid. However, the potential for false positive and negative test results and the technically complex procedure are limitations on its use. When the test results are used for regulatory decisions the implications of inaccurate results are serious and thus it is important that the assay be performed only by highly trained technicians working in an accredited laboratory that will use a standardized and validated assay.

The development and validation of an assay is an incremental process and consists of the determination of the feasibility of the method, the development and standardization of the assay, the determination of the characteristics of the assay and a constant monitoring, maintenance and enhancement of the assay (Office International des Epizooties 2003a). Where there are published standardized and validated methods, these should be followed. A protocol for RT-PCR has been published recently in the *Manual of Diagnostic Tests for Aquatic Animals* of the Office International des Epizooties (2003b). Improved methods and methods using other sets of primers have recently been reported (Mjaaland, Rimstad & Cunningham 2002).

It may be the case that only certain types of ISAV are pathogenic to Atlantic salmon (Ritchie, Cook, Melville, Simard, Cusack & Griffiths 2001; Nylund *et al.* 2003). This hypothesis might explain some of the observed disagreement between tests (e.g. PCRa might be able to detect non-pathogenic strains and this might explain why there were positive test results on PCRa and negative test results on other tests). Only gene sequencing can determine what strain is being detected. However this type of analysis is lengthy, costly and not suitable on a routine basis (Mjaaland *et al.* 2002). The use of primers specific for pathogenic strains could solve this issue.

Indirect fluorescent antibody test

The operator is the key factor influencing the repeatability of IFAT. Due to the subjectivity of the assay, there is no clear distinction between scores. The most difficult judgment is to differentiate

between a 0 (negative) and 1+ rating (positive if 1+ used as the cut-off point) in borderline cases. Four of 13 discordant duplicate samples (based on IFAT1) did not test positive on any other assay suggesting those four samples were false positives. The other nine fish had one or more other assays positive suggesting the discrepant result was a false negative. There were two operators used for IFAT reading at IFATa and differences between the readers might explain some of the differences.

Conclusions

Freezing (–80  C) does not negatively affect VI and may actually improve the sensitivity of RT-PCR assays. The repeatability and reproducibility of VI was almost perfect. There was substantial difference in repeatability of RT-PCR among the three laboratories and consequently only a moderate reproducibility between these laboratories, suggesting the diagnostic protocols and the interpretation of RT-PCR should be standardized across laboratories. The repeatability of IFAT was almost perfect when the agreement was evaluated on ordinal data. However, due to the subjectivity of the assay it might be prudent that at least two viewers are involved in questionable cases (1+) before a final rating is given. The assay should also be performed by highly trained personnel to read the sample consistently.

Acknowledgements

This study was supported financially by Aquanet, NBDFAA, and a number of producers and veterinarians. We are grateful to Dr Carol McClure, Department of Health Management, Atlantic Veterinary College, for providing us with some of the samples and assisting during the sample collection. Additionally we thank Laurie Wright, Katherine Cleghorn and Holly Beaman for valuable help during this study. Dr Gilles Olivier, Marcia Cook, Dr Carol McClure and Dr Stephen Griffiths are acknowledged for stimulating discussions.

References

- Anonymous (2000) ISA hits the Faroes. *Fish Farming International* **27**, 47.
- Anonymous (2002) *Infectious Salmon Anemia Virus*. International Committee on Taxonomy of Viruses, 00.046.0.05.001 (<http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/index.htm>).
- Bland M. & Altman D. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1**, 307–310.
- Bouchard D.A., Brockway K., Giray C., Keleher W. & Merrill P.L. (2001) First report of infectious salmon anaemia (ISA) in the United States. *Bulletin of the European Association of Fish Pathologists* **21**, 86–88.
- Byrne P.J., Macphee D.D., Ostland V.E., Johnson G. & Ferguson H.W. (1998) Haemorrhagic kidney syndrome of Atlantic salmon, *Salmo salar* L. *Journal of Fish Diseases* **21**, 81–91.
- Dannevig B.H., Falk K. & Namork E. (1995) Isolation of the causal virus of infectious salmon anaemia (ISA) in a long-term cell line from Atlantic salmon head kidney. *Journal of General Virology* **76**, 1353–1359.
- Dohoo I., Martin W. & Stryhn H. (2003) *Veterinary Epidemiologic Research*. AVC Inc., Charlottetown, Canada.
- Evensen O. & Thorud K.E. (1991) A morphological study of the gross and light microscopic lesions of infectious anaemia in Atlantic salmon (*Salmo salar*). *Research in Veterinary Science* **51**, 215–222.
- Falk K. & Dannevig B.H. (1995) Demonstration of infectious salmon anemia (ISA) viral-antigens in cell-cultures and tissue-sections. *Veterinary Research* **26**, 499–504.
- Falk K., Namork E. & Dannevig B.H. (1998) Characterization and applications of a monoclonal antibody against infectious salmon anaemia virus. *Diseases of Aquatic Organisms* **34**, 77–85.
- Merrill P. (2003) A comparative review of diagnostic assays used to detect infectious salmon anemia virus in the United States. In: *International Response to Infectious Salmon Anemia: Prevention, Control, and Eradication*, pp. 25–37. New Orleans, LA, 3–4 September 2002.
- Mjaaland S., Rimstad E., Falk K. & Dannevig B.H. (1997) Genomic characterization of the virus causing infectious salmon anaemia in Atlantic salmon (*Salmo salar* L.): an orthomyxo-like virus in a teleost. *Journal of Virology* **71**, 7681–7686.
- Mjaaland S., Rimstad E. & Cunningham C.O. (2002) Molecular diagnosis of infectious salmon anaemia. In: *Molecular Diagnosis of Salmonid Diseases*, pp. 1–22. Kluwer Academic Publishers, Dordrecht.
- N  rette P., Dohoo I. & Hammell L. (2005) Estimation of specificity and sensitivity of three diagnostic tests for infectious salmon anaemia virus in the absence of a gold standard. *Journal of Fish Diseases* **28**, 89–99.
- Nylund A., Devold M., Plarre H., Isdal E. & Aarseth M. (2003) Emergence and maintenance of infectious salmon anaemia virus (ISAV) in Europe: a new hypothesis. *Diseases of Aquatic Organisms* **56**, 11–24.
- Office International des Epizooties (2003a) Validation and quality control of polymerase chain reaction methods used for the diagnosis of infectious diseases. In: *Manual of Diagnostic Tests for Aquatic Animals 2003* (http://www.oie.int/eng/normes/fmanual/A_00012.htm).
- Office International des Epizooties (2003b) Infectious salmon anaemia. In: *Manual of Diagnostic Tests for Aquatic Animals 2003* (http://www.oie.int/eng/normes/fmanual/A_00026.htm).

- Ritchie R.J., Cook M., Melville K., Simard N., Cusack R. & Griffiths S. (2001) Identification of infectious salmon anaemia virus in Atlantic salmon from Nova Scotia (Canada): evidence for functional strain differences. *Diseases of Aquatic Organisms* **44**, 171–178.
- Rodger H., Turnbull T., Muir F., Millar S. & Richards R.H. (1998) Infectious salmon anaemia (ISA) in the United Kingdom. *Bulletin of the European Association of Fish Pathologists* **18**, 115–116.
- Royal Society of Edinburgh (2002) *The Scientific Issues Surrounding the Control of Infectious Salmon Anaemia (ISA) in Scotland*. Report of the Royal Society of Edinburgh Working Party on Infectious Salmon Anaemia. RSE, Edinburgh.
- Thorud K. & Djupvik H.O. (1988) Infectious salmon anaemia in Atlantic salmon (*Salmo salar* L.). *Bulletin of the European Association of Fish Pathologists* **8**, 109–111.
- Totland G.K., Hjeltnes B.K. & Flood P.R. (1996) Transmission of infectious salmon anaemia (ISA) through natural secretions and excretions from infected smolts of Atlantic salmon *Salmo salar* during their presymptomatic phase. *Diseases of Aquatic Organisms* **26**, 25–31.
- Wilson I.G. (1997) Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology* **63**, 3741–3751.

Received: 19 May 2004

Revision received: 24 November 2004

Accepted: 24 November 2004