Vol. 123: 101–122, 2017
https://doi.org/10.3354/dao03093

DISEASES OF AQUATIC ORGANISMS
Dis Aquat Org

Published March 6

Diagnostic validation of three test methods for detection of cyprinid herpesvirus 3 (CyHV-3)

Sharon C. Clouthier[1,*], Carol McClure[2], Tamara Schroeder[1], Megan Desai[1], Laura Hawley[3], Sunita Khatkar[1], Melissa Lindsay[1], Geoff Lowe[3], Jon Richard[3], Eric D. Anderson[4]

[1]Fisheries and Oceans Canada, Freshwater Institute, 501 University Crescent, Winnipeg, Manitoba R3T 2N6, Canada
[2]AquaEpi Research, RR3, 300 Heartz Road, Charlottetown, Prince Edward Island C1C 0H8, Canada
[3]Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada
[4]Box 28, Group 30, RR2, Ste Anne, Manitoba R5H 1R2, Canada

ABSTRACT: Cyprinid herpesvirus 3 (CyHV-3) is the aetiological agent of koi herpesvirus disease in koi and common carp. The disease is notifiable to the World Organisation for Animal Health. Three tests — quantitative polymerase chain reaction (qPCR), conventional PCR (cPCR) and virus isolation by cell culture (VI) — were validated to assess their fitness as diagnostic tools for detection of CyHV-3. Test performance metrics of diagnostic accuracy were sensitivity (DSe) and specificity (DSp). Repeatability and reproducibility were measured to assess diagnostic precision. Estimates of test accuracy, in the absence of a gold standard reference test, were generated using latent class models. Test samples originated from wild common carp naturally exposed to CyHV-3 or domesticated koi either virus free or experimentally infected with the virus. Three laboratories in Canada participated in the precision study. Moderate to high repeatability (81 to 99%) and reproducibility (72 to 97%) were observed for the qPCR and cPCR tests. The lack of agreement observed between some of the PCR test pair results was attributed to cross-contamination of samples with CyHV-3 nucleic acid. Accuracy estimates for the PCR tests were 99% for DSe and 93% for DSp. Poor precision was observed for the VI test (4 to 95%). Accuracy estimates for VI/qPCR were 90% for DSe and 88% for DSp. Collectively, the results show that the CyHV-3 qPCR test is a suitable tool for surveillance, presumptive diagnosis and certification of individuals or populations as CyHV-3 free.

KEY WORDS: CyHV-3 · Diagnostic validation · Precision · Accuracy · Quantitative PCR · Virus isolation

INTRODUCTION

Cyprinid herpesvirus 3 (CyHV-3) is the aetiological agent of koi herpes virus disease (KHVD) in common carp *Cyprinus carpio* L. and koi (Bretziner et al. 1999, Walster 1999, Hedrick et al. 2000, Perelberg et al. 2003, Ilouze et al. 2006, Boutier et al. 2015). The disease is notifiable to the World Organisation for Animal Health (OIE 2015a) and reportable to the Canadian Food Inspection Agency (CFIA 2016). Currently, there are no CyHV-3 tests whose analytical and diagnostic performance characteristics are validated to internationally recognized standards set forth by the OIE (OIE 2015b). This study was designed to address this need. The goal was to provide scientific information on CyHV-3 diagnostic test performance that could be used by stakeholders in support of the World Trade Organization Agree-

*Corresponding author: sharon.clouthier@dfo-mpo.gc.ca

ment on the Application of Sanitary and Phytosanitary Measures.

An effective CyHV-3 diagnostic assay should detect the virus in diseased fish as well as in asymptomatic carriers of the virus with latent or persistent infections. Three diagnostic assays were evaluated for their performance characteristics in this study: a modified conventional polymerase chain reaction (cPCR) assay targeting the CyHV-3 thymidine kinase gene (cTK; Bercovier et al. 2005), a modified quantitative PCR (qPCR) assay targeting CyHV-3 open reading frame 89 (ORF89) (qORF89; Gilad et al. 2004) and virus isolation by culture on common carp brain (CCB) cells (VI; Neukirch et al. 1999). The unmodified cTK and qORF89 tests have been evaluated in inter-laboratory ring studies (Way 2008) and intra-laboratory comparisons (Bergmann et al. 2010, Monaghan et al. 2015a). The qPCR test displayed high analytical specificity (ASp), detecting all known CyHV-3 isolates but not related virus species such as CyHV-1 or CyHV-2 (Gilad et al. 2004, Bergmann et al. 2010). It is currently the most commonly used diagnostic test for CyHV-3 and may prove to be suitable for use in surveillance programs to declare healthy populations of susceptible fish free of the virus (OIE 2015c). The cTK test is one of the PCR assays recommended by the OIE for detection of CyHV-3 (OIE 2015c). In contrast, virus isolation by cell culture is not considered by the OIE and others to be a reliable option for detection of CyHV-3 (Haenen et al. 2004, OIE 2015c).

CyHV-3 is a member of the genus *Cyprinivirus* within the *Alloherpesviridae* family and *Herpesvirales* order (Waltzek et al. 2005, 2009, Davison et al. 2009, Davison 2010). Phylogenetic analysis suggests the existence of 2 CyHV-3 lineages designated as Asian and European (Bigarré et al. 2009, Kurita et al. 2009, Avarre et al. 2011). Virus isolates from both lineages display a complex lifecycle which may contribute to the difficulty in detecting the virus. CyHV-3 is epitheliotropic (Hanson et al. 2011) and gains entry into carp via the skin covering the fins and body (Costes et al. 2009, Raj et al. 2011, Fournier et al. 2012, Ronsmans et al. 2014), the digestive tract through infection of the periodontal pharyngeal mucosa (Fournier et al. 2012) or the gills (Miyazaki et al. 2008, Monaghan et al. 2015b). The virus then moves rapidly to internal tissues including the liver, kidney, gut, spleen and brain (Gilad et al. 2004). A latent infection can eventually be established in host peripheral white blood cells, specifically the IgM+ B cells (Eide et al. 2011, Reed et al. 2014, Prescott et al. 2016). Adult fish that survive infection can serve as asymptomatic carriers of the virus, shedding virus particles through horizontal transmission in their urine and faeces (Dishon et al. 2005) as well as via their skin and gill mucus (Perelberg et al. 2003). The virus may cycle in a temperature-dependent manner between latent and acute infection stages in convalescent fish (Gilad et al. 2003, St-Hilaire et al. 2005, Uchii et al. 2009, 2011, 2014, Eide et al. 2011, Sunarto et al. 2014).

We hypothesized that 1 or more of the tests described in this study could provide the analytical and diagnostic performance suitable for use in national aquatic animal health programs as diagnostic assays for detection of CyHV-3. This validation study was multi-phasic, beginning with an assessment of each test's analytical characteristics including their ASp, analytical sensitivity (ASe) and repeatability. Estimates of each test's precision (i.e. repeatability and reproducibility) and accuracy (i.e. diagnostic sensitivity [DSe] and diagnostic specificity [DSp]) were generated in subsequent phases. The results are reported according to the Standards for Reporting of Animal Diagnostic Accuracy (STRADAS)-aquatic (Gardner et al. 2016).

MATERIALS AND METHODS

Diagnostic test optimization

Viruses and viral nucleic acid

The name, origin and source of the viruses used in this study are provided in Table 1. CyHV-3 was cultured on CCB cells (Neukirch et al. 1999) at 20°C. Spring viremia of carp virus (SVCV) and pike fry rhabdovirus (PFRV) were cultured using epithelioma papulosum cyprini cells at 20°C (Fijan et al. 1983, Winton et al. 2010). Cell culture medium was minimal essential medium with Hanks' salts (MEM-H) supplemented with 2 to 10% fetal bovine serum, 2 mM L-glutamine and, after virus infection, antibiotic/antimycotic (Life Technologies). Cells were inoculated with virus at a multiplicity of infection of 0.001 to 0.0001 and harvested after development of complete cytopathic effect (CPE). Virus-infected whole cell lysates were stored at −80°C. Virus titers were determined using the endpoint dilution method of Spearman (1908) and Karber (1931) and expressed as 50% tissue culture infective dose ($TCID_{50}$).

Infected whole cell lysates were used to prepare CyHV-3 U-NY1999 DNA positive reference material for the analytical and diagnostic validation studies.

Table 1. Virus isolates used in the present study

Virus	Geographic origin	Year	Reference
Alloherpesviridae			
Cyprinid herpesvirus 1 (CyHV-1)	Niigata, Japan	1981	Sano et al. (1985)
Cyprinid herpesvirus 2 (CyHV-2)	Japan	1992–1993	Jung & Miyazaki (1995)
Cyprinid herpesvirus 3 (CyHV-3)			
J-Ibaraki2007	Ibaraki, Japan	2007	R. P. Hedrick (pers. comm.)
J-Tottori2007	Tottori, Japan	2007	R. P. Hedrick (pers. comm.)
I-1999	Israel	1999	R. P. Hedrick (pers. comm.)
U-NY1999	NY, USA	1999	R. P. Hedrick (pers. comm.)
U-CA2001	CA, USA	2001	R. P. Hedrick (pers. comm.)
U-CA2002	CA, USA	2002	R. P. Hedrick (pers. comm.)
U-MI2011	MI, USA	2011	M. Faisal (pers. comm.)
MB2008	MB, Canada	2008	Garver et al. (2010)
MB2010	MB, Canada	2010	This study
Salmon herpesvirus 2 (SalHV-2)	Niigata, Japan	1981	Sano et al. (1983)
Sturgeon herpesvirus 1 (AciHV-1)	CA, USA	1991	Hedrick et al. (1991)
Sturgeon herpesvirus 2 (AciHV-2)			
	ID, USA	2001	Kurobe et al. (2008)
	Canada	1999	Kurobe et al. (2008)
Ictalurid herpesvirus (IctHV-1)	USA	1968	Fijan et al. (1970)
Rhabdoviridae			
Spring viremia of carp virus (SVCV)			
D148	UK	2001	Miller et al. (2007)
RHV	Ukraine	1989	Stone et al. (2003)
P4-7	Russia	1983	Stone et al. (2003)
M2-78	Moldova	1983	Stone et al. (2003)
Pike fry rhabdovirus (PFRV)			
II	Germany	1982	Stone et al. (2003)
IV	UK	1995	Stone et al. (2003)

Infected whole cell lysates were initially clarified by centrifugation (4500 × *g*, 20 min, 4°C). The pellet was discarded, and the supernatant was centrifuged (10 000 × *g*, 18 h, 4°C). The resulting pellet was homogenized in lysis buffer (10 mM Tris-HCl [pH 8], 1 mM pH 8.0 EDTA buffer; with 0.1% [w/v] sodium dodecylsulfate) with a glass Dounce tissue grinder and treated with RNaseA (16 µg ml^{-1}, 37°C, 30 min) followed by Proteinase K (80 µg ml^{-1}, 56°C, 3 h; Life Technologies). An equal volume of phenol, chloroform and isoamyl alcohol (25:24:1 ratio, respectively; Sigma-Aldrich) was added, and the solution was mixed (30 min, 22°C). Following centrifugation (10 000 × *g*, 20 min, 22°C), the aqueous phase of the solution was collected and mixed with an equal part isopropyl alcohol. After another centrifugation step (10 000 × *g*, 30 min, 4°C), the pellet, consisting of semi-purified CyHV-3 U-NY1999 DNA, was washed with 70% ethanol, centrifuged (10 000 × *g*, 10 min, 4°C), air dried and resuspended in TE buffer (Life Technologies). The CyHV-3 U-NY1999 positive reference DNA was quantified using the Nanodrop 8000 (Nanodrop Technologies) and then stored at –80°C.

Infected whole cell lysates were also used to prepare SVCV and PFRV RNA for synthesis of cDNA for use in the analytical validation studies. After clarification (2500 × *g*, 10 min, 4°C), the infected whole cell lysate supernatant was centrifuged (21 100 × *g*, 90 min, 4°C), and viral RNA was extracted using the QiaAmp Viral RNA Extraction Kit (Qiagen). The semi-purified SVCV and PFRV RNA was quantified and then stored at –80°C. Synthesis of cDNA was performed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Each reaction consisted of 20 µl containing 1× reverse transcriptase buffer, 4 mM dNTP mix, 1× random hexamer primers, 50 U MultiScribe Reverse Transcriptase (Applied Biosystems) and 20 ng RNA. The test was run using the Applied Biosystem Veriti 96-well thermocycler. The thermocycling profile was 1 cycle of 10 min at 25°C and 1 cycle of 120 min at 37°C followed by 1 cycle of 5 s at 85°C.

Tissue samples, virus free or infected with CyHV-3 U-NY1999 or MB2010, were used as the source of DNA for the cPCR and qPCR tests evaluated in the diagnostic validation study. The samples were processed using the DNeasy Blood & Tissue Kit (Qia-

gen). Tissue samples (<100 mg) were homogenized in lysis buffer (Buffer ATL; Qiagen) using a 5 mm stainless steel bead and a TissueLyser (2 min, 25 Hz; Qiagen). Proteinase K digestion (3 h at 56°C) was followed by DNA extraction. Tissue sample DNA was eluted in 200 µl elution buffer (Qiagen) and then quantified using a Nanodrop 1000 or 8000 (Nanodrop Technologies) or Multiskan (Thermo Scientific) spectrophotometer. The DNA samples were stored at −80°C.

Infected whole cell lysates were used to prepare CyHV-3 U-NY1999 or MB2010 DNA for qORF89 analysis of accuracy study virus isolation samples displaying suspect CPE in CCB monolayers. Supernatant (1 ml) from inoculated CCB cells in 24-well tissue culture plates was centrifuged (21 100 × *g*, 18 h, 4°C), the pellet was resuspended in 200 µl phosphate-buffered saline and DNA was extracted from the suspension using the DNeasy Blood & Tissue Kit (Qiagen). Infected whole cell lysate DNA was eluted in 200 µl elution buffer (Qiagen) and then quantified using the Nanodrop 8000 (Nanodrop Technologies) spectrophotometer. The DNA samples were stored at −80°C.

Plasmids

Plasmid DNA encoding CyHV-3 MB2008 DNA sequences was engineered for use as positive control material for the PCR tests in the validation study. The pORF89 plasmid contains the full-length sequence of ORF89 (1458 bp), and the pTK plasmid contains DNA encoding the thymidine kinase gene (TK, 651 bp) from CyHV-3 MB2008. PCR primers were designed based on sequences from CyHV-3 U-NY1999 (accession number DQ657948) or CyHV-3 I-1999 (accession number AJ535112). The MB2008 isolate sequences were amplified by cPCR using pairwise primer combinations of KHV orf89full-F (5′-ATG GCC TCC ACT TCA ACC GCT GTG-3′) and KHV orf89full-R (5′-TTA AGC GAG CAG TCC CCT CGG G-3′) or KHV TK-FullF (5′-ATG GCT ATG CTG GAA CTG GTG ATC GG-3′) and KHV TK-FullR (5′-TCA CAG GAT AGA TAT GTT ACA AGA ACG AGG TGG AG-3′). Each 25 µl reaction contained 1× PCR buffer (Applied Biosystems), 1.5 or 3.0 mM $MgCl_2$, 200 µM dNTPs, 1.25 or 2.5 U AmpliTaq Gold DNA polymerase (Applied Biosystems), 400 nM of each primer and between 100 and 500 ng of DNA isolated in 2008 from Lake Manitoba carp infected with CyHV-3 MB2008. Thermocycling parameters were as follows: 95°C for 5 min; 35 cycles of 95°C for 30 s, 67 or 68°C for 30 s and 72°C for 1 or 2 min; followed by a post-cycle extension at 72°C for 8 min. Amplicons generated from the cPCR were analyzed for purity and size by electrophoresis in 1% agarose gels, purified with the QIAquick Gel Extraction Kit (Qiagen), TA cloned into vector pGEM-T Easy (Promega) and then transformed into *Escherichia coli* DH5alpha competent cells (Invitrogen) according to manufacturers' instructions. Plasmid DNA was purified with a QIAprep Spin Miniprep Kit (Qiagen) as described by the manufacturer and screened for inserts by cPCR using the relevant primer pairs. Both strands of DNA from at least 3 positive clones per amplicon were sequenced by the dideoxynucleotide chain termination method using an automated sequencer (Sanger et al. 1977). Two additional primers were used for sequence analysis of ORF89 amplicons: KHV orf89 Mid For (5′-GTG GTA GTG GTA GCG AC-3′) and KHV orf89 Mid Rev (5′-GAG CTC ATG GCG ATC ACC-3′). Analyses of the DNA sequences were performed using BioEdit v7.0.9.0 software (Hall 1999). Plasmid DNA was linearized with the *Nde*I restriction enzyme (New England Biolabs), analyzed for complete digestion by electrophoresis in 1% agarose gels and purified with the QIAquick Gel Extraction Kit (Qiagen). The linearized DNA was used as positive control material for the qORF89 and cTK tests. The pORF89 construct is 4498 bp in length corresponding to an estimated 2.92×10^6 g mol^{-1}. The copy number per µg DNA was calculated as 2.06×10^{11}. The pTK construct is 3668 bp in length corresponding to 2.38×10^6 g $mole^{-1}$. The copy number per µg DNA was calculated as 2.53×10^{11}.

Optimization of qORF89 and cTK

The analytical performance of both PCR tests was optimized. Variables assessed included primer and probe quantity as well as Mg^{2+}, dNTPs, polymerase, target DNA and thermocycling conditions. Target DNA from 3 different sources was also evaluated: CyHV-3 infected carp tissue, CyHV-3 infected whole cell lysate and plasmid encoding CyHV-3 DNA (pORF89 or pTK).

qPCR test parameters were selected for their ability to amplify plasmid DNA from $10^{5.7}$ to $10^{0.7}$ copies per reaction with equimolar concentrations of primers, in increments of 100 nM, from 300 to 700 nM. In all cases, the hydrolysis probe concentration was 80 nM. Primer concentrations for qORF89 giving the lowest quantification cycle (Cq) value were selected for further analysis. Equimolar probe concentrations, in

increments of at least 10 nM from 80 to 150 nM, were tested, and those giving the highest final fluorescence value in the baseline-corrected, ROX-normalized view (i.e. dRn) were selected for further analysis. The effect of adding a 30 s 72°C elongation step into each cycle was also assessed.

The cTK test was optimized with the 3 different CyHV-3 DNA templates, primer concentrations from 100 to 1500 nM, 1 to 3 mM $MgCl_2$, 1 to 2.5 U *Taq*, 35 or 40 cycles, 52 or 55°C annealing temperature and 30 or 60 s annealing time. Primer concentrations and cycling conditions were selected from those tests giving the highest relative fluorescence of target product in the absence of non-specific amplification products.

The effect of the quantity of DNA template on qORF89 and cTK was examined with 500, 1000, 1500, 2000 and 3000 ng of kidney tissue. The effect of digesting the tissue with Proteinase K for 3 or 18 h prior to DNA extraction was also evaluated for each test.

Test methods and outcome measures

qPCR assay

The qORF89 qPCR assay was originally described by Gilad et al. (2004). The reaction conditions were modified to increase its analytical performance. The primers were forward primer orf89-QF (5′-GAC GCC GGA GAC CTT GTG-3′), reverse primer orf89-QR (5′-CGG GTT CTT ATT TTT GTC CTT GTT-3′) (Sigma Aldrich; Gilad et al. 2004) and assay probe orf89-QProbe (5′-6FAM-CTT CCT CTG CTC GGC GAG CAC G-TAMRA-3′) (Life Technologies; Gilad et al. 2004). Each test consisted of 25 µl containing 400 nM orf89-QF, 400 nM orf89-QR, 100 nM orf89-QProbe, 1× TaqMan Universal PCR Master Mix (Applied Biosystems) and 1500 ng genomic DNA. The test was run using the Stratagene Mx3000/5P system in a 96-well format. The thermocycling profile was 1 cycle of 2 min at 50°C and 1 cycle of 10 min at 95°C followed by 40 cycles of 15 s at 95°C and 60 s at 60°C. Fluorescence was measured at the end of each amplification cycle. Tests were run with 3 replicates of each sample. Data were analyzed using MxPro software (Stratagene) with the adaptive baseline and amplification-based threshold algorithm enhancements. The threshold was adjusted manually such that it was above the background and in the linear phase of the amplification plot. The copy number of pORF89 was calibrated using the standard curve method provided in this software. The PCR cycles used for quantification are referred to as the quantification cycle or Cq (Bustin et al. 2009). Samples were considered positive if the Cq value was <40 and negative if the Cq value was ≥40. Results were analyzed as binary data as well as continuous data. For the latter, the Cq values were configured as the average of the 3 replicates. In those cases where 1 or 2 of the 3 replicates had a value of >40 Cq, the Cq value for that replicate was set at 40.

cPCR assay

The cPCR cTK assay was originally described by Bercovier et al. (2005). The reaction parameters are those described by the OIE (2015b). The primers for the cTK test were forward primer TK-EPF (5′-GGG TTA CCT GTA CGA G-3′) and reverse primer TK-EPR (5′-CAC CCA GTA GAT TAT GC-3′) (Sigma-Aldrich; Bercovier et al. 2005). The cTK test consisted of 25 µl containing 100 nM TK-EPF, 100 nM TK-EPR, 2.5 mM $MgCl_2$, 250 µM dNTPs, 1.25 U AmpliTaq Gold (Applied Biosystems) and 500 ng genomic DNA. It was performed using an Applied Biosystems Veriti thermocycler. The thermocycling profile was 1 cycle of 5 min at 94°C; 40 cycles of 60 s at 95°C, 60 s at 55°C and 60 s at 72°C; followed by 1 cycle of 10 min at 72°C. Amplicons were separated by electrophoresis on 1 to 2% agarose gels (Life Technologies) or E-gels (Thermofisher) containing 1× fluorescent nucleic acid stain GelRed (Biotium) or SYBR® Safe (Thermofisher), respectively. Data were recorded using a digital imaging system (e.g. Kodak Gel Logic 200). The expected amplicon was 409 bp in length. The samples were run singly and reported to be positive or negative based on the respective presence or absence of this fragment in the agarose gel following electrophoresis.

The sample set for each test run of qORF89 and cTK included control samples of a known composition. The qORF89 control material consisted of a buffer blank (N1), 50 mg kidney tissue + 25 ng DNA isolated from CyHV-3 U-NY1999 infected cell monolayers (P1), pORF89 diluted from $10^{3.7}$ to 5 copies per reaction (P3) and water (N3). The cTK positive control material consisted of kidney tissue + 50 ng DNA isolated from CyHV-3 U-NY1999 infected cell monolayers (P1) and $10^{4.7}$ copies of pTK per reaction (P3). The N1 and P1 samples were included at the nucleic acid extraction step and carried through the final PCR step, whereas N3 and P3 were added at the PCR step of the qORF89 and cTK tests. The positive con-

trol samples confirmed that the test was performing as expected, whereas the negative control samples were essential for detecting contamination or non-specific amplification in the reaction. If unexpected results were obtained with the control samples (i.e. no amplification detected with the positive control sample or amplification detected with the negative control sample), then the results for that sample batch were considered invalid.

Virus isolation

For the precision study, CyHV-3 isolation was done by culture on CCB cells. This test is referred to as VI to differentiate it from the VI/qPCR test (see next paragraph). A pool of frozen kidney, spleen and gill tissues was homogenized in MEM-H2 containing antibiotics and 0.02 M HEPES. This was either clarified (2500 × *g*, 15 min, 4°C) and diluted to 2% or used to make a 2% solution that was then filtered (0.45 µm). These samples were held at 4°C for less than 48 h prior to testing with duplicate samples of 100 µl. Fresh monolayers of CCB cells in 24-well plates were inoculated with the diluted tissue homogenate, incubated at 20°C for 30 to 60 min, overlayed with 1 ml per well MEM-H2 containing antibiotics and 0.02 M HEPES and then incubated at 20°C for 21 d. Cell monolayers were examined within 48 h of inoculation and then observed twice weekly for virus-induced CPE. One laboratory performed a blind passage with the pooled supernatant collected from the first cell culture passage on CCB cell monolayers. For the precision study, the presence or absence of CPE was reported for each replicate (i.e. VI). If CPE was observed in 1 or both of the replicates, then the sample was considered positive. If CPE was observed in neither replicate, the sample was reported as negative. Each sample set included a positive control sample consisting of cell culture supernatant from CyHV-3 U-1999 infected whole cell lysates as well as a negative control sample such as sterile MEM-H2. These samples were used to verify the validity of the unknown sample test results.

For the accuracy study, CyHV-3 isolation was conducted as described directly above except that the qORF89 qPCR test was used to confirm the CPE result for samples displaying suspect CPE. This form of the test is referred to as VI/qPCR. Results were reported as described above based on the CPE observed or the qORF89 result if the qPCR test was used to confirm suspect CPE.

Analytical validation

Analytical sensitivity, specificity and repeatability

The ASe of qORF89 and cTK was evaluated using standard curves generated with (1) plasmid DNA (pORF89 or pTK), (2) DNA extracted from CyHV-3 infected whole cell lysates or (3) DNA extracted from tissue + exogenous DNA from CyHV-3 U-NY1999 infected whole cell lysates. Comparison of the qPCR reaction efficiencies was used to determine whether plasmid DNA could be used as a proxy for absolute quantification of CyHV-3 in infected tissue. Standard curves were constructed from 2- or 10-fold serial dilutions, and each dilution was assayed by qPCR in triplicate or in replicates of 6. The qPCR assay was performed using the reaction conditions described in 'Test methods; qPCR assay'. The 50 and 100% limit of detection (LOD) for the assay was expressed as copies of plasmid DNA and was derived from the measured concentrations of the last dilution with at least 50% or with 100% detection, respectively.

The ASp of qORF89 and cTK was tested for exclusivity with nucleic acid isolated from tissue or cells infected with viruses listed in Table 1. Inclusivity of these assays was established using DNA from infected tissue or lysates of cells infected with CyHV-3 isolates listed in Table 1. Tests were performed using 3 replicates per sample and the reaction conditions outlined in 'Materials and methods: Test methods and outcome measures'. The level of agreement between the observed and expected results was calculated using 2-way tables and expressed as inclusivity or exclusivity (i.e. degree to which the assay detects all intended viruses and does not detect other viruses, respectively). The 95% confidence intervals for these estimates were also calculated for each assay.

The analytical repeatability of qORF89 was tested with pORF89 serially diluted 10-fold, from $10^{8.7}$ to $10^{1.7}$, and then 2- to 5-fold, from 25 to 0.5 copies per reaction. The intra- and inter-assay repeatability was determined by analyzing each dilution in replicates of 6 in 5 independent runs. The repeatability of the assay was also evaluated with positive control samples P1 and P3. Samples from 2 batches of P1 and P3 were assessed in 54 independent runs performed by 5 analysts over a 4 yr period. The qORF89 test was performed using the reaction conditions outlined in 'Materials and methods: Test methods and outcome measures'. The results were evaluated in a scatter plot of the mean Cq of the replicates plotted against the standard deviation. A linear regression was used to determine whether the inter-run or inter-batch

variability was statistically significant (i.e. $p < 0.05$). The coefficient of variation (CV) was calculated as SD/mean × 100 using the continuous outcome data (i.e. Cq values). All statistical analyses were performed in Stata/IC (v12.1).

Diagnostic validation

All items from the STRADAS-aquatic checklist (Gardner et al. 2016) applied to this study except for STRADAS items 7, 17, 19, 20 and 23

Participating laboratories

Three laboratories participated in the precision study (Fig. 1). All are located in Canada and are members of the country's National Aquatic Animal Health Program (NAAHP). Two of the laboratories are located at the Freshwater Institute (FWI, Winnipeg, MB), and the third is located at the Pacific Biological Station (PBS, Nanaimo, BC). They are referred to as follows: laboratory A: FWI-Aquatic Animal Health Laboratory (AAHL) (research); laboratory B: FWI-AAHL (diagnostic); laboratory C: PBS-AAHL. At the time of the study, laboratories B and C were working towards ISO/IEC 17025 accreditation and have now achieved that status. Laboratory A is a research laboratory. Each laboratory performed all 3 diagnostic tests. Seven analysts were involved, and all had extensive experience in their respective tests. Two analysts in laboratory C participate in annual inter-laboratory comparison studies run by the European Union Reference Laboratory for Fish Diseases. For our validation study, analysts (n = 4) performing the qPCR test were required to pass a qORF89 proficiency panel prior to starting the validation samples. These panels were shipped by courier on dry ice for overnight delivery to laboratory C or transferred into an ultra-cold freezer operated by laboratory A or B. A statement of shipment condition accompanied the courier package, with a request to confirm the presence of dry ice upon receipt. Instructions on how to conduct the qPCR test (as well as the cPCR and VI tests) were sent to all participants. Each panel consisted of 5 blinded, randomly selected samples containing naïve kidney tissue or kidney tissue + 1 or 50 ng DNA isolated from CyHV-3 U-NY1999 infected whole cell lysates and contained at least 1 strong positive, 1 weak positive and a negative sample. The qPCR test was run with 3 replicates of each sample. Panel results were evaluated using qualitative (i.e. positive, negative) and quantitative (i.e. Cq values) data. The latter were assessed relative to a consensus value. Results were rejected if the Cq value reported was greater than 2 standard deviations above the expected value for positive samples or if a Cq value was reported for a negative sample. The assigned value was the average Cq of the participants' results for a sample, and the standard deviation was based on the Cq values reported by participants for that sample.

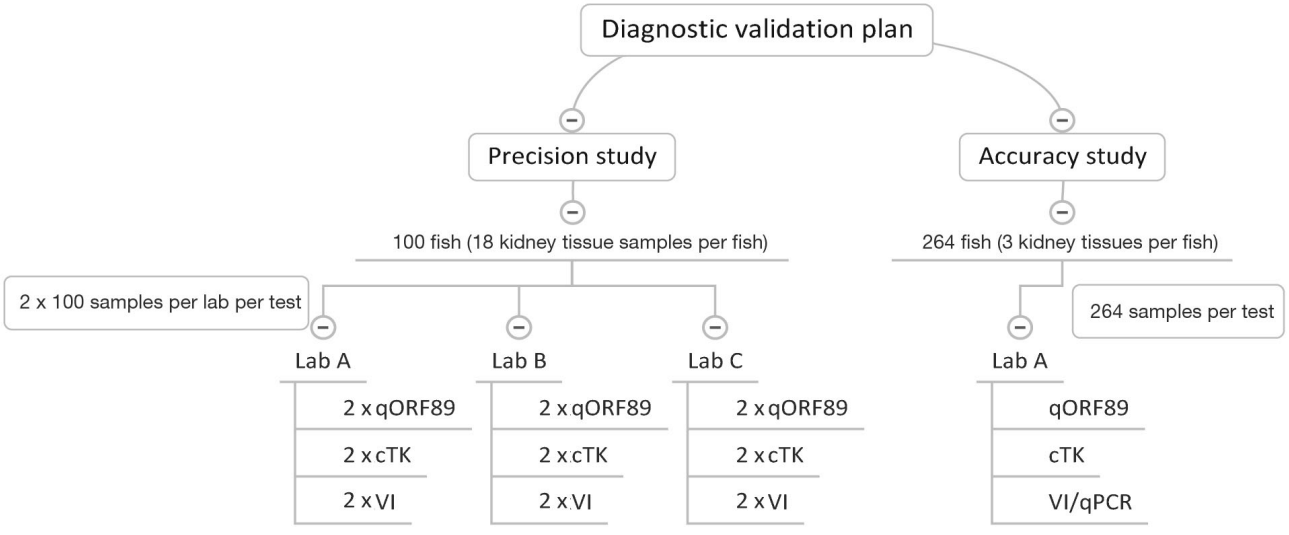


Fig. 1. Diagnostic validation study plan. Number of fish, tissue samples, laboratories and cyprinid herpesvirus 3 (CyHV-3) diagnostic tests used in the precision and accuracy studies. qORF89: quantitative PCR (qPCR) assay targeting CyHV-3 open reading frame 89; cTK: CyHV-3 thymidine kinase conventional PCR assay; VI: virus isolation; VI/qPCR: VI followed by qORF89 testing of cell culture supernatant

Fish populations

The diagnostic validation study was conducted using a negative reference population, a low-prevalence CyHV-3 MB2010 positive reference population and a high-prevalence CyHV-3 U-NY1999 positive experimental population.

The negative reference population was sourced from a koi hatchery (BC, Canada) with no previous history of KHVD. Kidney tissue samples from 22 koi were tested by qORF89 for the presence of CyHV-3 nucleic acid to confirm freedom of infection by the virus. Koi were selected from the hatchery, euthanized, placed on ice (not frozen) and shipped by courier for next-day delivery to the FWI-AAHL in Winnipeg, MB, Canada. Fish were dissected upon their arrival. One shipment of 100 koi (40.6 ± 9.41 g) was received and processed in September 2010. Of these fish, 80 were selected for inclusion in the accuracy study. A second shipment of 20 koi (83.18 ± 19.46 g) was received and processed in October 2010. All 20 fish were included in the precision study.

The positive reference population with a low prevalence of CyHV-3 MB2010 consisted of 134 wild, apparently healthy adult carp (5.17 ± 2.34 kg; 70 male, 64 female) collected from 5 sites on Delta Marsh, Lake Manitoba, in July and August 2010. Fish were caught using gill nets (1.8 m in height, 95 to 140 m in length) with mesh sizes ranging from 1.5 to 5 inches (~3.8 to ~12.7 cm). The average water temperature over this period of time was 22.2 ± 3.2°C. Fish were transported by boat from the collection site to a field laboratory where the necropsy and tissue sample collection took place immediately. The collection of carp was performed opportunistically, coordinated with existing fishing efforts on Delta Marsh by Ducks Unlimited Canada. All of the carp (n = 134) collected during the 7 field trips were included in our prospective precision (n = 30 fish) and accuracy (n = 104 fish + 30 fish from the precision study) studies. Lake Manitoba was the site of a KHVD outbreak in wild carp populations in July and August 2008.

The positive experimental population with a high prevalence of CyHV-3 was comprised of CyHV-3 U-NY1999 infected koi generated through an *in vivo* virus challenge study conducted in November 2010. The study was executed under FWI Animal Use Protocol number 2010-069. Naïve koi (n = 333; 23.3 ± 7.8 g) were purchased from a commercial supplier (BC, Canada) with no previous history of KHVD. Fish were selected from the hatchery, placed in shipping containers and transported live overnight by air to FWI (Winnipeg, MB). Upon arrival at the facility, the koi were distributed among seven 75.8 and two 37.5 l glass aquaria using an average stocking density of 12.3 ± 2.4 g l^{-1}. The koi were acclimated over a 5 d period by increasing the temperature of de-chlorinated city water by approximately 2°C d^{-1} from 14 to 24°C. Fish were held at 24.0 ± 0.27°C for the duration of the study. They were fed a dry pellet diet (Hikari) daily at 1% of their body weight. For the challenge, fish were anesthetized with 100 µg l^{-1} tricaine methane sulfonate (TMS) prior to handling. Koi (n = 308; 44 per tank) housed in the larger tanks were intraperitoneally injected with $1 \times 10^{4.25}$ $TCID_{50}$ of CyHV-3 U-NY1999, whereas fish (n = 24; 10 or 14 per tank) in the smaller tanks received 100 µl sterile MEM-H2. Dead or moribund fish were removed twice daily for 10 d after challenge, euthanized if required, individually bagged, placed on ice and immediately necropsied for tissue collection. With the exception of 1 fish that died on Day 3, mortality in the virus-challenged fish was first observed 5 or 6 d following challenge. The moribund koi displayed a variety of clinical signs that included lethargy, erratic swimming, loss of equilibrium, hanging in the water column with the head down and widespread hemorrhage of the epidermis and fins. The trial was terminated on Day 10, and the average cumulative mortality in the groups of CyHV-3 U-NY1999 challenged fish was 58 ± 0.08%. No mortality was observed in the 2 groups of mock-challenged fish. At the end of the trial, all remaining fish were euthanized by an overdose of TMS, individually bagged, placed on ice and processed immediately for tissue collection. The precision study included samples from 50 representative fish randomly selected from this population, and an additional 50 fish were similarly selected for the accuracy study.

Tissue sample collection, preservation and shipping conditions

Fish were dissected in a dry laboratory setting using sterile technique. A new set of necropsy tools was used between fish and between the outside and inside of each fish to avoid the risk of cross-contamination. Dissection tools (i.e. disposable scalpel, forceps) were not re-used. For the precision study, 18 pieces of kidney (≤25 mg), 6 pieces of gill and 6 pieces of spleen tissue (≤50 mg) were harvested per fish. For the accuracy study, 3 pieces of kidney (25 to 50 mg), 1 gill arch and 1 piece of spleen tissue (≤50 mg) were harvested from each fish. Kidney tissue was collected and placed in individual microtubes for the qORF89 and cTK test samples, whereas kidney (≤50 mg),

spleen (≤50 mg) and gill tissue (at least 1 gill arch) were collected from the same fish and pooled in a Whirl-Pak bag for the virus isolation samples. Each sample was individually labeled with a computer-generated 5-digit random number so that technicians processing them were blinded to the samples' CyHV-3 status. Tissues in the tubes or bags were flash frozen in a dry ice–ethanol bath and transferred to an ultra-cold freezer (e.g. −80°C) for interim storage. Bulk samples of kidney, spleen and gill tissue were collected in the field from each fish and placed in separate 50 ml tubes (i.e. 3 tubes per fish). The samples were immediately flash frozen in a dry ice–ethanol bath and transported on dry ice prior to storage in an ultra-cold freezer. A few days before shipment, frozen pieces of field-collected tissue of approximately 25 mg (≤25 mg) or 50 mg were transferred into microtubes or Whirl-Pak bags, respectively, sorted on dry ice into boxes for each laboratory and then transferred back into the ultra-cold freezer. At the same time, archived samples from the other 2 fish populations were retrieved from −80°C storage and sorted on dry ice into the same boxes. Samples (n = 600, duplicate sets of 100 samples for each test) and positive control material for the precision study were shipped by courier on dry ice in January 2011 for overnight delivery to laboratory C. A statement of shipment condition accompanied the package, with a request to confirm the presence of dry ice upon receipt. At the same time, the other 2 laboratories received duplicate sets of samples (n = 600 per laboratory) which were transferred to ultra-cold freezers. The 3 laboratories completed all the testing by November 2012, and the results were used to generate estimates of test repeatability and reproducibility. Laboratory A tested an additional 792 samples (n = 264 per test) in December 2012 through January 2013, and the results were used to evaluate the DSe and DSp of the 3 assays.

Statistical methods for diagnostic validation study

Precision estimates

Three laboratories participated in a blinded joint testing scheme in which sampled fish were subjected to all 3 imperfect diagnostic tests performed in parallel. The test panels for the repeatability and reproducibility study consisted of duplicate sets of 100 tissue samples (sample sets 1 and 2) collected from the negative reference population (n = 20 fish), the positive reference population (n = 30 fish) and the positive experimental population (n = 50 fish).

Estimates of within-laboratory agreement (repeatability) and between-laboratory agreement (reproducibility) for binary outcomes were generated using the proportion of agreement and Cohen's kappa (κ), which estimates agreement beyond chance (Cohen 1960). Kappa values range from −1 to +1, and the results were interpreted using the classification scheme of Landis & Koch (1977). Confidence intervals for the κ statistic were calculated by the method of Reichenheim (2004). Kappa functions as an index of concordance only if the observed proportion positive for each test run is similar. The significance of any differences in the proportion of positive test results for each pair of test runs was assessed using McNemar's test statistic (χ^2). This computation was performed prior to estimating κ since evidence of disagreement between test results (i.e. p-value ≤ 0.05) suggests that the test pairs cannot possibly agree and that kappa may not be a useful measurement (Dohoo et al. 2009).

Precision of the qORF89 test was also evaluated using its continuous outcome data. Agreement was assessed using 2 approaches. The first method was the concordance correlation coefficient (CCC), which quantified the degree of agreement between 2 Cq values according to the method of Lin (1989, 2000). In our study, concordance values were determined for duplicate sample sets tested by qORF89 within each laboratory and between the 3 laboratories. Only fish that had Cq values <40 for both sample pairs were included in this analysis. Pairwise agreement of the repeatability data was also evaluated graphically with concordance plots. Perfect concordance was represented by a 45° angle straight line through the origin. A best-fit line representing the pairwise agreement between the test results was created by linear regression analysis. Precision of the data was represented by how tight the data points were to the best-fit line, whereas accuracy was assessed by how far the best-fit line deviated from the 45° line through the origin of the graph.

The second statistical method for evaluating agreement in Cq values was the Bland and Altman limits of agreement (Bland & Altman 1986, Barnhart et al. 2007). In our study, the agreement was reported as the mean difference between duplicate sets of Cq data within 1 laboratory. Bland–Altman limits of agreement and plots were used to assess overall mean differences and the range of test pair differences, to identify outliers and to reveal relationships between mean differences and the magnitude of the Cq values reported in each laboratory (Dohoo et al. 2009).

Accuracy estimates

The disease status of each sample used in the accuracy study was defined using 2 approaches. The first method established the disease status of fish using the aforementioned gold standard negative or positive reference populations. For this approach, the DSe or proportion of positive fish testing positive and the DSp or proportion of negative fish testing negative were calculated using 2 × 2 tables as described by McClure et al. (2005). The second method involved maximum-likelihood (ML) and Bayesian estimation procedures, which were used to fit latent class models (LCMs) and determine the disease status of each fish that would most likely yield the observed test results (Enøe et al. 2000).

DSe and DSp estimates were generated in the second approach without defined reference animals. Latent class analysis was conducted using binary test outcomes of laboratory A collected from the qORF89, cTK and VI/qPCR tests for each of the sampled fish in the accuracy study. ML and Bayesian estimation procedures were used to fit the LCMs. ML estimates for population prevalences and diagnostic test DSe and DSp were generated using the software program TAGS (tests in the absence of a gold standard; Pouillot et al. 2002). Having multiple different populations of animals increased the degrees of freedom (df) of our model, which allowed for goodness of fit measurement estimates. Test results were collected from 3 assays (qORF89, cTK, VI/qPCR) for each of the 264 fish from 3 different prevalence populations. Two ML models were run with no prior information added or with the negative reference population prevalence set to 0 and the positive experimental reference population prevalence set to 100%. A third model in which hypothetical changes were made to test results from 9 fish was used to test the stability of the accuracy estimates. Goodness of fit p-values less than 0.05 provided evidence that the fit of the model was not adequate and suggested that the estimates from the model were most likely biased (Pouillot et al. 2002).

Bayesian estimates of test accuracy and prevalence were obtained using Markov chain Monte Carlo (MCMC)-based algorithms and WinBUGS software (Spiegelhalter et al. 2003). Preliminary models were run with uniform or non-informative priors set for DSe, DSp and population prevalence (Table 2). They were initially run with all tests independent of one another. Conditional dependence between each possible pair of tests (i.e. qORF89–cTK, qORF89–VI/qPCR, cTK–VI/qPCR) was explored by adding covariance terms to the model (Table 2; Gardner et al. 2000, Nérette et al. 2008). Additional models were constructed by restricting the prevalence of the negative reference population to 0 and the prevalence of the positive experimental reference population to 1 by adjusting the prior distributions (Table 2). Model stability and robustness of the

Table 2. Bayesian models used in the accuracy study to generate estimates of diagnostic sensitivity (DSe) and diagnostic specificity (DSp) for qORF89 (Q), cTK (P) and VI/qPCR (VI) tests. Prior prevalence low = 0 sets the prevalence for the negative reference population to 0, and prior prevalence high = 1 sets the prevalence for the positive experimental population to 1. cov: covariance; other abbreviations as in Fig. 1

Variables	Model category, name and composition							
	Base	Two-test dependence				Prior	Prior prevalence high = 1	Sensitivity analysis
	Bayes-1	Bayes-2 covQP	Bayes-2 covQVI	Bayes-2 covPVI	Bayes-2 covQP DSe only	prevalence Bayes-3	Bayes-3A	Bayes-4[a]
No covariance	✓					✓	✓	✓
No prior	✓	✓	✓	✓				✓
DSe covariance QP		✓						
DSp covariance QP		✓			✓			
DSe covariance QVI			✓					
DSp covariance QVI			✓					
DSe covariance PVI				✓				
DSp covariance PVI				✓				
Prior prevalence low = 0					✓	✓		
Prior prevalence high = 1					✓	✓	✓	
Altered data								✓

[a]Run using data in which hypothetical changes were made to test results from 9 fish

accuracy estimates were evaluated by making hypothetical changes to test results from 9 fish. Models (n = 7) corresponding to different covariance patterns (with and without conditional dependence) were then compared using the deviance information criterion (DIC) (Spiegelhalter et al. 2002, Nérette et al. 2008) and Bayesian goodness of fit p-values (Nérette et al. 2008). As described by Caraguel et al. (2012), models with DIC values less than 3 units from the DIC value for a different model were considered significantly better in our study. Models with similar DIC values were discriminated based on their simplicity such that those with fewer variables were deemed better models. Bayesian p-values near the extreme values of 0 and 1 provided evidence for lack of fit (Gelman et al. 2003). MCMC analysis was run using a burn-in period of 10 000 iterations for all models to eliminate Markov chains that were unrepresentative of the equilibrium distribution. Posterior probability estimates were then made using the next 50 000 iterations. Convergence of the Markov chains was evaluated using trace plots of MCMC iterations, cumulative quantile plots of the parameters and autocorrelation plots of the MCMC samples. Point estimates were taken from the posterior median, and the probability intervals were from percentiles of the posterior distributions.

RESULTS

Diagnostic test optimization

qPCR reactions containing 100 nM of probe orf89-QProbe detected CyHV-3 with higher fluorescence relative to those containing 80 nM, which is the quantity recommended for the qORF89 test by Gilad et al. (2004). The optimized parameters for the cTK test corresponded to those described by the OIE (2015b). The quantity of template DNA added to each reaction was 500 ng for the cTK test and 1500 ng for the qORF89 test. Interference in the performance of both tests was observed when the recommended quantity of DNA was exceeded (data not shown). The duration of Proteinase K digestion of kidney tissue (3 to 18 h) did not alter the tests' analytical performance.

Analytical validation

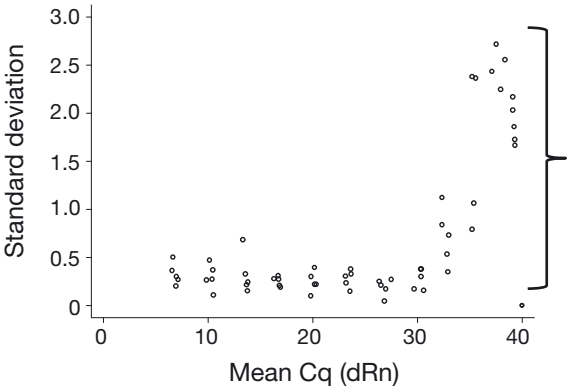
The qORF89 test for detection of CyHV-3 nucleic acid was linear across 10 orders of magnitude, with a strong correlation between cycle number and quantity of target template (R^2 = 0.993; Fig. S1A in the Supplement at www.int-res.com/articles/suppl/d123p101_supp.pdf). qPCR amplification efficiency ranged from −3.3 to −3.4 (Fig. S2 in the Supplement). The similarity in amplification efficiency independent of the template type showed that plasmid DNA could be used to define the LOD for qORF89.

The ASe of qORF89 and cTK was determined using pORF89 and pTK plasmid DNA. The cTK test LOD was 500 copies of pTK. The qPCR test was positive, with $10^{8.7}$ and $10^{-0.3}$ copies of pORF89 or 10 orders of magnitude (Fig. S1 in the Supplement). The observed ≥50% LOD for qORF89 was 5 copies of pORF89 (26 of 30 replicates tested positive; 35.71 ± 1.81 Cq). The 100% LOD was observed to be 25 copies of pORF89 (32.68 ± 0.72 Cq).

The ASp of qORF89 and cTK was determined using the viruses shown in Table 1. The tests performed equally well with respect to exclusivity (100% [75–100%]) and inclusivity (100% [63–100%]). In both cases, the test results were positive for the CyHV-3 isolates and negative for other members of the *Alloherpesviridae* or SVCV which may coexist with CyHV-3.

Repeatability of the qPCR assay was evaluated with pORF89 plasmid template DNA ranging from $10^{8.7}$ to $10^{-0.3}$ copies. The average intra-assay CV for Cq values was 0.71 to 5.04, with the highest CV values associated with samples having the lowest and highest plasmid copy number of 5 to 0.5 copies and $10^{8.7}$ copies, respectively (Fig. 2). The inter-assay CV for Cq values varied from 0.94 to 3.52. As in the previous case, the highest CV values were associated with samples with the lowest and highest plasmid copy numbers (Fig. 2).

Repeatability of qORF89 was also evaluated with positive control samples P1 and P3. The average intra-assay CV for Cq values obtained with P1 was 1.09, whereas the inter-assay CV was 4.38 (Fig. S3 in the Supplement). The average intra-assay CV for Cq values of P3 within a run ranged from 0.82 to 2.74 (Fig. S3). The highest CV value was associated with samples with the lowest plasmid copy number: 5 plasmid copies. The variability in Cq values obtained for P1 was not statistically significant ($p < 0.05$) between batches, runs, analysts or laboratories. For P3, variability was not significant between runs but was significant at 50 copies between batches (−6.67 Cq) and at all dilutions between analysts (−0.36 to −0.28 Cq) and laboratories (−1.11 to −0.87 Cq).



Plasmid copies	Cq ± SD[a]	Coefficient of variation	
		Intra-assay	Inter-assay
5×10^{-1}	39.02±1.60	4.13	2.00
1×10^{0}	38.86±1.83	4.77	2.43
5×10^{0}	35.71±1.81	5.04	2.25
2.5×10^{1}	32.68±0.72	2.19	0.94
5×10^{1}	30.26±0.28	0.91	1.11
5×10^{2}	26.84±0.19	0.71	1.69
5×10^{3}	23.44±0.28	1.18	1.12
5×10^{4}	20.08±0.25	1.23	1.15
5×10^{5}	16.70±0.25	1.51	1.42
5×10^{6}	13.62±0.32	2.39	1.42
5×10^{7}	10.27±0.30	2.91	2.64
5×10^{8}	6.86±0.33	4.81	3.52

Fig. 2. Analytical repeatability of the quantitative PCR assay targeting cyprinid herpesvirus 3 open reading frame 89 (qORF89). Quantification cycle (Cq) values for the FAM-labelled probe were generated with plasmid pORF89 DNA diluted from $10^{8.7}$ to 0.5 copies (dRn: Rox-normalized, baseline corrected fluorescence view). Measurements were performed in replicates of 6 in 5 independent runs. Intra- and inter-assay coefficient of variation for Cq values as well as the standard deviation and mean Cq values for each of 5 runs are presented for each dilution in the table. In the table, Cq is the mean of the mean Cq per run, SD is the mean of the mean Cq standard deviation per run and intra-assay is the average of the within-assay coefficient of variation for all runs

qORF89 proficiency testing

Four analysts were selected for their proficiency with the CyHV-3 qORF89 test. None of the analysts reported false positive or false negative results. All the analysts reported outcomes that were lower than 2 standard deviations above the expected value for the strong (28.72 ± 2[0.95]) and weak (35.05 ± 2[1.11]) positive samples. The same nucleic acid extraction kit and qPCR platform were used by the laboratories, while the primers were from 2 different sources.

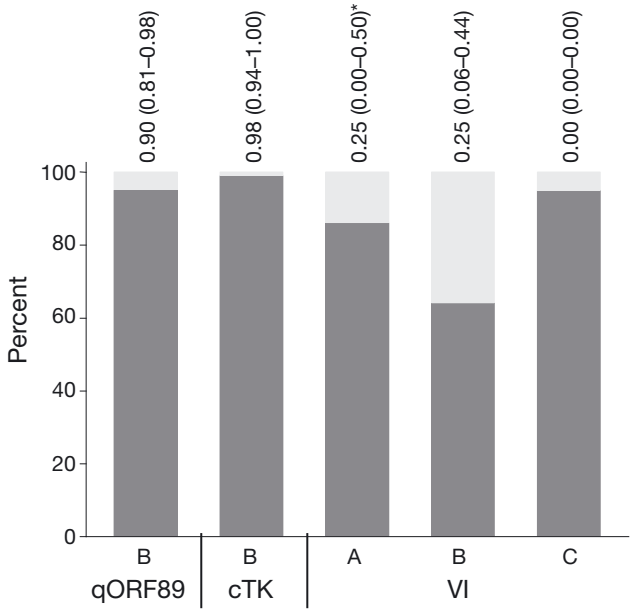
Diagnostic validation

An overview of the precision and accuracy studies is provided in Fig. 1. The positive and negative test results from the precision study are provided in Fig. S4 in the Supplement and those from the accuracy study are provided in Fig. S5 in the Supplement using the alignment format of Caraguel et al. (2009). The 3 fish corresponding to the VI samples reported as contaminated by Laboratory C were removed from the analyses.

Precision study

The precision study showed evidence of false positive results with the qORF89 and cTK tests in 2 of the 3 laboratories. Positive PCR test results in laboratories A and C2 (i.e. laboratory C, sample set 2) were reported for some of the N1 negative control samples introduced at the nucleic acid extraction step. The N3 negative control samples for the corresponding cPCR or qPCR steps were always negative. The N1 false positive results invalidated the results reported for the study samples processed in parallel with those negative control samples. Results for valid and invalid data sets from all 3 laboratories are presented in the 'Results' body text; only the results from valid data sets (i.e. those with no evidence of false positive results) are presented in the corresponding figures and tables.

Repeatability, binary data. Repeatability for the valid binary outcomes from laboratory B for the qORF89 and cTK tests and from all 3 laboratories for the VI test is presented in Fig. 3. The observed pairwise agreement of CyHV-3 test results (n = 2 × 100 per test) within each of the 3 laboratories ranged from 79 to 95% for qORF89, 86 to 99% for cTK and 64 to 94.85% for the VI test. The range in number of sample pairs testing positive in a set was 54 to 60 for qORF89, 50 to 51 for cTK and 0 to 83 for VI. Estimates of kappa for the same pairwise comparisons were in the range of substantial to almost perfect agreement for qORF89 (0.83 to 0.90) and cTK (0.71 to 0.98) with the exception of 1 qPCR result (0.50) in the moderate agreement category. The kappa values for the VI test were in the slight to fair agreement range (0 to 0.25) (Fig. 3). McNemar's test was significant ($p < 0.05$) for virus isolation from laboratory A.



Test	Lab	Test result (Test 1, Test 2)			
		+,+	+,–	–,+	–,–
qORF89	B	54	4	1	41
cTK	B	50	0	1	49
VI	A	83	1	13	3
	B	43	13	23	21
	C	0	0	5	92

Fig. 3. Repeatability of qORF89 and cTK tests in laboratory B and VI test in 3 laboratories (A, B, C). Estimates of within-laboratory agreement for binary outcomes with duplicate samples were generated using the proportion of agreement and Cohen's kappa methods. Each bar represents the proportion (%) of samples in agreement (dark grey) or disagreement (light grey). Cohen's kappa values appear above each bar, with the 95% confidence intervals in brackets. McNemar's test was significant ($p \leq 0.05$) for pairs of test results denoted with an asterisk (*). The number of paired samples with results corresponding to each combination of test results is provided in the table. Abbreviations as in Fig. 1

Repeatability, continuous data. Laboratory repeatability for the CyHV-3 qORF89 Cq test results expressed as estimates of agreement from CCCs was 0.879 (0.821 to 0.937) for laboratory A, 0.814 (0.722 to 0.905) for laboratory B (Table 3) and 0.814 (0.725 to 0.903) for laboratory C. A representative concordance correlation plot generated using paired test results from laboratory B is presented in Fig. 4A. The Bland–Altman average difference in paired Cq values was 0.887 (–5.156 to 6.930) for laboratory A, 0.126 (–5.874 to 6.126) for laboratory B (Table 3) and 0.461 (–6.505 to 7.426) for laboratory C. The number of paired test results with Cq values reported by each laboratory was 54 to 60. A representative Bland and Altman plot for test pair results reported by laboratory B is presented in Fig. 4B.

Reproducibility, binary data. Reproducibility for the valid binary outcomes from laboratories B and C1 for the qORF89 test, from laboratories A2 and B for the cTK test and from all 3 laboratories for the VI test is presented in Table 4. Laboratory reproducibility estimates expressed as agreement proportions for CyHV-3 test results ($n = 2 \times 100$ per test) ranged from 79 to 97% for qORF89, 85 to 98% for cTK and 4 to 66% for VI. The range in number of sample pairs testing positive in a set between laboratories was 53 to 59 for qORF89, 49 to 51 for cTK and 0 to 64 for VI. Estimates for kappa were in the range of substantial to almost perfect agreement for qORF89 (0.613 to 0.939) and cTK (0.695 to 0.960) with the exception of 2 qPCR results (0.561 and 0.601) in the moderate agreement category. The kappa value ranges for the

Table 3. Repeatability and reproducibility of the quantitative PCR assay targeting cyprinid herpesvirus 3 open reading frame 89 (qORF89 continuous outcome data). Test pairs indicate laboratory (A, B, C) and sample set (1, 2). Confidence intervals (95%) accompany each estimate

Diagnostic characteristic	Test pairs	Concordance correlation coefficient (degree of agreement)	Bland–Altman limit of agreement (average difference)	Tests positive in both sample sets
Repeatability	B1, B2	0.814 (0.722, 0.905)	0.126 (–5.874, 6.126)	54
Reproducibility	B1, C1	0.882 (0.823, 0.941)	0.749 (–4.496, 5.994)	56
	B2, C1	0.723 (0.593, 0.852)	0.683 (–5.822, 7.188)	53

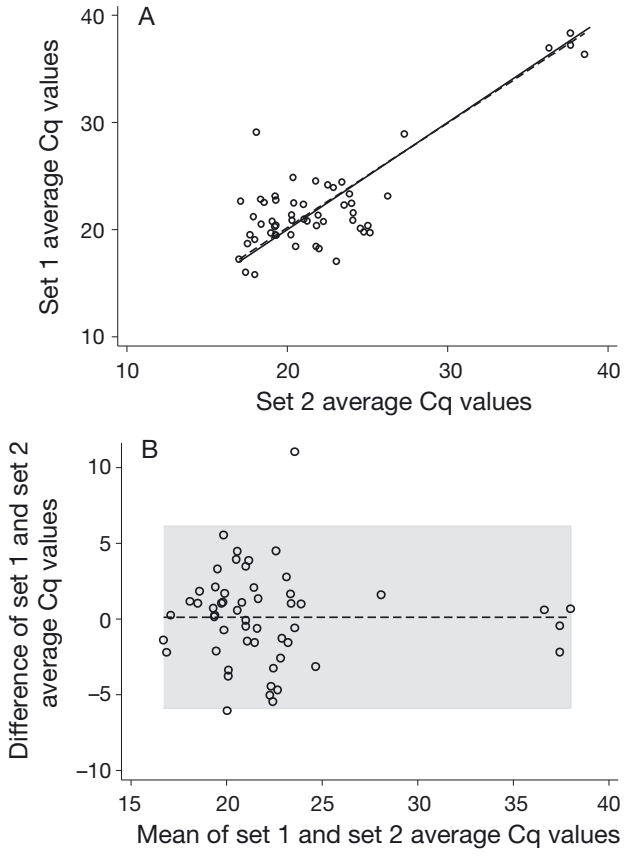


Fig. 4. Repeatability of quantification cycle (Cq) values for the quantitative PCR assay targeting cyprinid herpesvirus 3 open reading frame 89 (qORF89) in laboratory B. Pairwise analysis of average Cq values <40 is presented in (A) the concordance correlation plot and (B) the Bland and Altman limits of agreement plot. In (A), the linear regression line of the test results is shown by the dotted line, whereas the line of perfect concordance is solid black. In (B), $y = 0$ is the line of perfect agreement, the mean difference in Cq values between duplicate samples in set 1 and 2 is represented by the dotted line and the grey zone represents the 95% confidence interval

VI test were in the no to slight agreement categories (−0.074 to 0.129). McNemar's test revealed disagreement for 24 of 36 (67%) pairwise comparisons including 8 of 12 qPCR, 4 of 12 cPCR and all 12 of the VI paired test results.

Reproducibility, continuous data. Reproducibility for the valid continuous data from laboratories B and C1 for the qORF89 test is presented in Table 3. Laboratory reproducibility of CyHV-3 qORF89 Cq test results expressed as estimates of agreement from CCCs was in the range of 0.764 (0.650 to 0.878) to 0.882 (0.823 to 0.941) for laboratory A–B test pairs, 0.740 (0.622 to 0.857) to 0.912 (0.869 to 0.955) for laboratory A–C test pairs and 0.723 (0.593 to 0.852) to 0.882 (0.823 to 0.941) for laboratory B–C test pairs. The Bland and Altman average difference in paired Cq values ranged from −0.200 (−6.306 to 5.907) to 0.803 (−5.347 to 6.953) for laboratory A–B test pairs, 0.485 (−5.074 to 6.044) to 1.830 (−5.153 to 8.812) for laboratory A–C test pairs and 0.683 (−5.822 to 7.188) to 1.248 (−4.646 to 7.141) for laboratory B–C test pairs. The range in number of paired test results with Cq values reported for 2 laboratories was 53 to 57 for laboratories A and B, 56 to 59 for laboratories A and C and 53 to 56 for laboratories B and C.

Accuracy study

Gold standard reference population method. Estimates of DSp and DSe for qORF89, cTK and VI/qPCR were generated using the binary test results of laboratory A from fish belonging to the gold standard reference population (n = 80 fish) or the gold standard positive experimental population (n = 50 fish) of the accuracy study, respectively. Each fish was tested 3 times. The DSe and DSp of the 2 PCR-based tests qORF89 and cTK were perfect (i.e. 100%). The accuracy estimates for the VI/qPCR test were lower at 89% for the DSe and 92% for the DSp.

Latent class modelling. Latent class analysis was the second approach used to generate estimates of DSe and DSp for the 3 diagnostic tests. The data set for this evaluation consisted of qORF89, cTK and VI/qPCR results reported by laboratory A, for an additional 264 fish sampled for the accuracy study. A summary of the results is presented in Fig. S5 in the Supplement. ML and Bayesian estimation procedures were used to fit the LCMs.

ML analysis was performed assuming conditional independence between tests. Each model had 12 df available to evaluate other parameters. However, all of the models displayed unsatisfactory goodness of fit p-values (i.e. 0.002 to 0.02), providing evidence that they did not fit the data and that the corresponding estimates of diagnostic accuracy were biased.

Bayesian estimates were generated using models run with or without prior prevalences and with or without covariance between tests (Table 2). Conditional dependence between tests was evaluated since qORF89, cTK and VI/qPCR are nucleic acid based and measure the same biological trait (i.e. viral DNA), a factor that may result in a possible violation of an underlying assumption of the model (i.e. conditional independence of all tests). MCMC convergence diagnostics revealed stable history and quantile plots as well as short-range autocorrelation

Table 4. Reproducibility of qORF89, cTK and VI (binary outcome data). Test pairs indicate laboratory (A, B, C) and sample set (1, 2). Symbols in parentheses describe the test results presented for test 1 and test 2, respectively. Estimates of between-lab agreement for binary outcomes with duplicate samples were generated using the proportion of agreement and Cohen's kappa methods. Confidence intervals (95%) accompany each kappa estimate. McNemar's test was significant for pairs of test results denoted with an asterisk (*). Abbreviations as in Fig. 1

Test pairs (test 1, test 2)	2 × 2 contingency table (+,+)	(+,−)	(−,+)	(−,−)	McNemar's chi-squared p-value	Agreement (%)	Cohen's kappa
qORF89							
Lab B1, Lab C1	56	2	1	41	1.000	97.00	0.939 (0.870 to 1.000)
Lab B2, Lab C1	53	2	4	41	0.688	94.00	0.878 (0.784 to 0.973)
cTK							
Lab A2, Lab B1	49	8	1	42	0.039*	91.00	0.820 (0.709 to 0.931)
Lab A2, Lab B2	49	8	2	41	0.109	90.00	0.799 (0.682 to 0.917)
VI							
LabA1, Lab B1	50	34	6	10	<0.0001*	60.00	0.129 (−0.030 to 0.288)
Lab A1, Lab B2	54	30	12	4	0.008*	58.00	−0.074 (−0.238 to 0.090)
Lab A1, Lab C1	0	84	0	16	<0.0001*	16.00	0.000
Lab A1, Lab C2	3	78	2	14	<0.0001*	17.53	0.030 (−0.089 to 0.029)
Lab A2, Lab B1	55	41	1	3	<0.0001*	58.00	0.056 (−0.035 to 0.147)
Lab A2, Lab B2	64	32	2	2	<0.0001*	66.00	0.036 (−0.077 to 0.150)
Lab A2, Lab C1	0	96	0	4	<0.0001*	4.00	0.000
Lab A2, Lab C2	5	89	0	3	<0.0001*	8.25	0.004 (−0.001 to 0.008)
Lab B1, Lab C1	0	56	0	44	<0.0001*	44.00	0.000
Lab B1, Lab C2	4	50	1	42	<0.0001*	47.42	0.046 (−0.030 to 0.121)
Lab B2, Lab C1	0	66	0	34	<0.0001*	34.00	0.000
Lab B2, Lab C2	4	60	1	32	<0.0001*	37.11	0.023 (−0.036 to 0.081)

plots for all models, suggesting that each attained a stationary distribution. The model with the best Bayesian p-value and the lowest DIC accounted for DSe dependences between the qORF89 and cTK tests. However, the DIC for this model was not lower by 3 units from the DIC for the conditionally independent model in which the negative reference population prevalence was restrained to 0 and the positive experimental population prevalence was restrained to 1. The aforementioned qORF89/cTK DSe covariance model provided evidence of minor dependence with a covariance of 0.004 (0, 0.04). Although the dependence covariance was significant (i.e. larger than 0), the DIC values for the 2 models were similar and the conditionally independent model contained fewer variables, making it the model of choice for this study.

Bayesian posterior distributions of population prevalence and test DSe and DSp are summarized in Table 5 with their median estimates and the corresponding 95% credibility intervals. Point estimates for each parameter were similar across the conditionally independent and dependent models. The probability intervals overlapped and precluded conclusive evidence that the estimates were different, but in general, qORF89 and cTK had higher DSe values (i.e. 97 to 99%) relative to those for VI/qPCR (i.e. 89 to 90%). The probability of qORF89 and cTK to test negative was high, ranging from 92 to 93%, whereas the DSp for the VI/qPCR test was lower at 87 to 88%. Prevalence estimates for the negative reference population ranged from 0 to 1%, whereas the positive reference and positive experimental populations had values ranging from 1 to 2% and 98 to 100%, respectively. The Bayesian p-values for the 7 models were in the range of 0.1178 to 0.1651, and the DIC values were between 71.343 and 77.061. Altering the data from 9 fish produced accuracy estimates with the Bayes-4 model similar to those obtained with the other 7 models providing evidence of model stability and robustness (i.e. Bayes-4 model estimates were lower by 1 to 2%).

DISCUSSION

The results indicated that the qORF89 and cTK tests but not the VI test are fit for the intended purposes of surveillance, presumptive diagnosis and certifying the CyHV-3 status of individual common carp and koi and populations of these fish. Reproducibility estimates for the VI assay showed unacceptable agreement (Caraguel et al. 2009) for all of its paired

Table 5. Diagnostic sensitivity and specificity of the qORF89, cTK and VI/qPCR tests and CyHV-3 prevalence for the 3 fish populations. Median posterior estimates (95% confidence interval) from a conditionally independent Bayesian model run with prior distributions for negative and positive reference populations set at 0 and 1, respectively. Abbreviations as in Fig. 1

Parameter	qORF89	cTK	VI/qPCR
Diagnostic sensitivity	0.99 (0.93, 1.0)	0.99 (0.93, 1.0)	0.90 (0.80, 0.96)
Diagnostic specificity	0.93 (0.90, 0.96)	0.93 (0.90, 0.96)	0.88 (0.83, 0.92)
Bayesian p-value	0.1477		
Deviance information criterion	71.435		
	Negative reference population	Positive reference population with a low prevalence	Positive experimental reference population with a high prevalence
Prevalence	0	0.02 (0, 0.05)	1

inter-laboratory test results. The poor precision of VI may be attributed to differences in cell line stability and variation in their susceptibility to CyHV-3 infection. In laboratories A and B, the CCB cell line exhibited a pseudo-cytopathic effect which became evident in the negative control cell monolayers towards the end of the test period. Consequently, these laboratories reported more false positive test results than laboratory C. Laboratory C reported a higher proportion of false negative results, which was attributed to the low susceptibility of the laboratory's CCB cell line to CyHV-3. The inconsistent performance of the CCB as well as the koi fin cell lines has also been reported by others (Haenen et al. 2004, OIE 2015c). If it is necessary to use these cell lines within a diagnostic workflow, then a validated test should be used to confirm that CPE is due to CyHV-3.

The analytical performance estimates for the qORF89 test were close to their theoretical limits for sensitivity, specificity and repeatability. The qORF89 test was linear over 10 orders of magnitude with ASe estimates of 25 plasmid copies (100% LOD) and 5 plasmid copies (>50% LOD). The cTK test was 100 times less sensitive than the qPCR assay, reflecting the technological differences (Ratcliff et al. 2007) between the 2 types of tests. Further improvements in qPCR-based tests for detection of CyHV-3 might be obtained by targeting higher copy number DNA or by targeting differentially expressed viral mRNA such as ORF6 mRNA, which is synthesized in IgM^+ B cells during CyHV-3 latent infections (Reed et al. 2014). This approach may increase the probability of detecting CyHV-3 nucleic acid in asymptomatic, apparently healthy fish with latent infections of the virus. The trade-off would be virus lifecycle-dependent variations in the test's analytical performance. Alternatively, different tissue types could be targeted for sampling depending on the population of fish to be tested. The possible neurotropic or leucotropic nature of CyHV-3 in the latent phase of its lifecycle suggests that blood or brain tissue may be better candidates than the kidney tissue used in this study for detection of CyHV-3 in apparently healthy fish. In the event that non-lethal sampling is a requisite, as in the case of high-value koi, then mucus or skin tissue should be explored as possible alternatives to kidney tissue in populations experiencing a suspected KHVD outbreak.

The ASp of the 2 PCR tests was inclusive of designated European and Asian lineages of CyHV-3 and exclusive of sympatric aquatic viruses. This means that the qORF89 and cTK tests were 100% inclusive, detecting all of the CyHV-3 isolates included in the study, and 100% exclusive, since they did not detect other members of the *Cyprinivirus* genus or *Alloherpesviridae* family or co-localizing viruses. In this context, it is still possible that a false negative test result could arise from a loss of target specificity via a mutation event. However, the qORF89 and cTK tests target highly conserved regions of the CyHV-3 genome evidenced by 100% identity of the nucleic acid sequences found in the National Center for Biotechnology Information database corresponding to the primer/probe binding sites. Genetic changes in CyHV-3 PCR target sequences have been noted by Bergmann et al. (2010), who reported that the cTK assay was unable to detect newly adapted CyHV-3 variants. However, nucleic acid sequences from these CyHV-3 isolates have not yet been reported. The TK gene, a common diagnostic target, encodes a protein that in other herpesviruses has been shown to be involved in virulence and may play a role in reactivation of latent herpesvirus infections (Lycke 1990, Tenser 1991). Frameshift mutations caused by insertions or deletions within long homopolymer nucleotide stretches in the TK gene occur

in drug-resistant herpesviruses (Kit et al. 1987, Sasadeusz et al. 1997, Yamada et al. 2005). The function of the ORF89 protein remains unknown, making it impossible to predict what, if any, genetic changes may occur in the region targeted by the primers. The risk of false negative results due to loss of specificity can be addressed through recursive evaluation of the tests' ASp. The likelihood of detecting changes in ASp is increased using 2 tests like qORF89 and cTK which target distinct genes.

The qORF89 test performed consistently within a laboratory, and proficiency testing showed that it was transferable to the 3 laboratories participating in the diagnostic precision study. These laboratories could detect CyHV-3 in experimentally exposed koi displaying clinical signs of infection within 3 d following virus challenge with isolate U-NY1999. The virus load ($10^{6.81} \pm 10^{7.07}$ equivalent plasmid copies per µg DNA) and the wide range of virus quantities (range $10^{1.95}$ to $10^{7.87}$ equivalent plasmid copies per µg DNA) found in the experimentally infected koi were similar to values reported for CyHV-3 in wild carp undergoing an outbreak of KHVD (Garver et al. 2010). In the absence of available samples from a natural KHVD outbreak, tissues from the experimentally infected koi served as suitable proxies for a natural acute infection. However, the laboratories were able to detect CyHV-3 MB2010 in apparently healthy and asymptomatic wild carp collected in 2010 from Delta Marsh on Lake Manitoba. In this case, CyHV-3 MB2010 nucleic acid was present at a low prevalence (1 to 2%) and low virus load (≤5 plasmid copies), suggesting some fish were latently infected with CyHV-3.

The diagnostic performance characteristics of the qORF89 and cTK tests were affected by their analytical performance. The lower ASe of cTK relative to qORF89 likely explained the lower number of positive cTK test pair results in the precision study. Similar differences in the relative ASe of the 2 assays have been reported by others (Bergmann et al. 2010, Monaghan et al. 2015a). Thus, during routine surveillance when the pathogen load of apparently healthy individuals is likely to be low, the more analytically sensitive qORF89 test would be the preferred screening test. The cTK test is recommended as the second assay conducted sequentially to the qORF89 test. When the pathogen load in samples is low, the cTK test may or may not be able confirm a positive test result obtained with the qORF89 test. In our study, this grey zone was represented by the virus load of less than 50 equivalent plasmid copies per µg DNA (mean Cq ≥ 31.15 ± 1.14; Fig. 5). Increasing the number of replicates of the cTK assay or lowering the Cq threshold value selected for a sample to be considered positive may increase the probability of confirming the initial positive qORF89 result. The trade-off associated with lowering the Cq cutoff below 40 would be a potential reduction in the DSe of qORF89. An alternative approach would be to use a second qPCR test targeting a nucleic acid sequence other than ORF89.

The precision study provided evidence of false positive results with the qORF89 and cTK tests. The most likely explanation is cross-contamination (Wilson 1997). No evidence of this error was present in the qPCR test results reported by laboratories B and C1 or the cTK test results from laboratories A2 and B. Results from these contamination-free sample sets were considered valid and were used to generate estimates of repeatability and reproducibility. The cross-contamination events evident in the precision study were traced to the nucleic acid extraction step. Theoretically, the probability of contamination is associated with the prevalence of infection and the pathogen load in each batch of samples tested. In our study, the level for both of these variables was high in tissue samples from the positive experimental reference population which had an estimated CyHV-3 U-NY1999 prevalence of 98 to 100% and virus loads of $10^{6.81} \pm 10^{7.07}$ equivalent plasmid copies per µg DNA. These results illustrate the risk from increased contamination pressure that exists for a laboratory even when experienced analysts are processing tissues from heavily infected populations.

The source of the contamination was identified by processing in parallel a combination of highly posi-

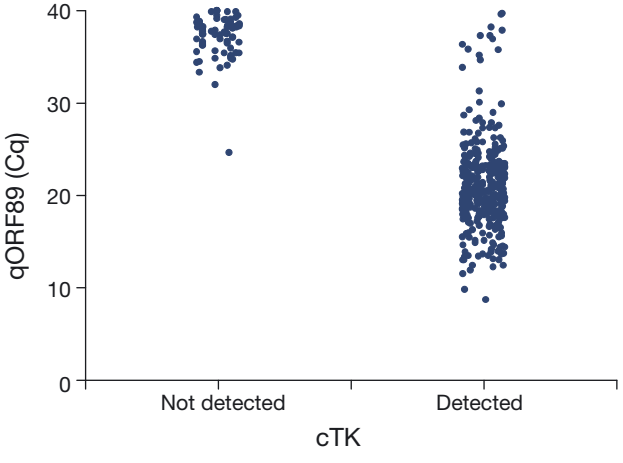


Fig. 5. qORF89 and cTK test results using duplicate samples. Abbreviations as in Fig. 1

tive and negative samples under variable conditions. All 3 laboratories were using the spin column-based DNeasy Blood & Tissue Kit (Qiagen). We found that separating the columns in the microfuge by 1 or 2 spaces during centrifugation steps reduced the risk of cross-contamination, whereas 3 spaces eliminated the risk altogether. We also initiated the practice of sub-dividing kit reagents in volumes sufficient to process each batch of samples. Any sub-divided reagent remaining after a batch extraction was discarded to prevent cross-contamination between batches. These precautions were implemented by laboratory A for the subsequent accuracy study, and no further evidence of cross-contamination was observed. Proper technology transfer of the qORF89 and cTK tests to other laboratories must account for this risk to ensure consistent test performance across all laboratories.

Control samples were added at each step of the 2-stage process for the qORF89 and cTK tests to ensure the validity of the results. The positive controls for this test were unlikely to be the source of cross-contamination observed in the precision study given that their pathogen loads were at or below the levels present in the qORF89 proficiency panel samples. However, the plasmid-based positive control samples for the qPCR test were not designed to include an artificial probe-binding site (Snow et al. 2009). The next generation of control material for the qORF89 test will include this design feature to facilitate the identification of potential false positive results.

The precision study should periodically be repeated to ensure consistent repeatability and reproducibility of the qORF89 and cTK tests. We can make some initial statements regarding the performance consistency of these tests using select data sets from the precision study. Repeatability of the qORF89 and cTK tests was high within laboratory B. These tests were more repeatable than the virus isolation method. The concordance correlation plots for duplicate sample sets tested within laboratory B provided further evidence for high repeatability, as the data points were clustered close to the best-fit line and little deviation of the best-fit line from the 45° line was observed. Reproducibility estimates generated with binary paired test results from the contamination-free sample sets (qORF89: laboratory B1, B2, C1; cTK: laboratory A2, B1, B2) provided evidence that the qORF89 and cTK tests performed consistently. Inter-laboratory estimates based on the Cq values were high, but the values may have been affected by the consistently lower average Cq values reported for samples by laboratory C1 relative to laboratory B. These results support the recommendation that qORF89 and cTK are fit for use as diagnostic assays.

The variability in test operating characteristics observed in our study was likely multi-factorial. Factors influencing test precision include its analytical performance, the targeted population (i.e. virus load and prevalence), differences in laboratory practices (technician experience and proficiency, laboratory training program, equipment, facility design) and sample preparation (collection method, storage, tissue preparation) (Caraguel et al. 2009, 2012, OIE 2015b). The cross-contamination issue outlined in the previous paragraphs likely contributed to a subportion of the non-repeatable and unacceptable agreement results observed for some of the inter-laboratory test pairs. The low virus load in tissue samples from apparently healthy carp from Lake Manitoba could have been another contributing factor. If the concentration of target DNA was at the tests' LOD, then it would not always be detected, and test precision estimates would be adversely affected. A correlate for this was evident in the qORF89 test pair results for laboratories B and C1, as the non-repeatable pairs were exclusively from same-fish samples from the lightly infected carp population. Low pathogen loads may also confound electrophoretic gel reading for the cTK test. Gel reading subjectivity has been presented as a possible explanation for the low precision reported for cPCR tests in other validation studies (Nérette et al. 2005, Caraguel et al. 2009). Uneven distribution of CyHV-3 MB2010 virions in the kidney may also have affected test precision estimates since tissues were not homogenized prior to sample distribution into tubes for our study. It was expected that this impact would be more evident with same-fish samples with a low virus load (Nérette et al. 2005, Caraguel et al. 2009, 2012). In our study, the concordance correlation plots for duplicate sample sets tested by 2 laboratories revealed greater variation in results reported for samples with low Cq values, suggesting that the precision of the test was reduced for samples with a high virus load (data not shown). This finding may indicate unequal distribution of the virus in heavily infected kidney tissue. The cumulative impact of these factors may correlate qORF89 and cTK test precision performance with virus load in the target population. Variation of repeatability and reproducibility across virus infection prevalences and infection stages has been reported by other groups (Caraguel et al. 2009, 2012). Future modelling of the

impact of these factors will provide a better understanding of the performance of qORF89 and cTK under different circumstances and would facilitate interpretation of test results.

In the absence of a perfect reference test, estimates of DSp and DSe for qORF89, cTK and VI/qPCR were generated in this study using gold standard negative and positive reference populations as well as LCMs. The precision study results provided evidence that the VI CPE-based results were not reliable, so in the accuracy study, the CyHV-3 status of samples displaying suspect CPE was confirmed using the qORF89 test. The combined results of the 2 tests conducted in series were used as the third test in the models. The LCMs were run without success using results from the qORF89 and cTK tests. When all 3 tests were used in parallel, conventional 2-class LCMs generated estimates but only with the Bayesian approach. The ML approach was not successful likely due to violation of one or more of the assumptions (Hui & Walter 1980) associated with these models. For example, the prevalence of the negative reference population (i.e. 0 to 1%) was not significantly different from that of the positive reference population with low prevalence (i.e. 1 to 2%), and these 2 populations essentially condensed into 1 population. This situation violates the assumption that the prevalence of the 3 populations differ. The Bayesian test estimates generated in this study were similar across the 8 models tested, suggesting that the estimates were equivalent, stable and robust. Test performance characteristics from LCM analyses were similar to those predicted using gold standard reference populations. The accuracy study provided evidence that the qORF89 and cTK tests were able to identify CyHV-3 positive fish even in the population with a low virus prevalence and low virus load.

LCMs such as those run in this study assume that the accuracy of diagnostic tests is consistent across all fish populations (Hui & Walter 1980). In some cases, the assumption is not valid as shown by Caraguel et al. (2012), who demonstrated that test performance may actually differ depending on the infection status of the host. In our study, estimates of test performance reflect an average across the 3 populations. Additional research would be required to evaluate the validity of the assumption inherent in LCMs that qORF89, cTK and VI/qPCR perform consistently across populations of fish in different stages of infection with CyHV-3. We suspect that the diagnostic operating characteristics of the tests would change with the health status of the population such that their DSe and DSp would be high with naïve fish and with fish during an outbreak and potentially lower with convalescent fish following an outbreak of KHVD.

Canada's NAAHP has selected qPCR as the choice methodology for testing apparently healthy or clinically diseased populations of fish for the presence of aquatic pathogens including CyHV-3. The primary advantages of molecular diagnostic assays over culture- or antibody detection-based methods is their relatively higher ASe and ASp as well as their faster turn-around times and greater ease of automation (Cobo 2012, Gullett & Nolte 2015). These features are important especially in cases of rapid onset of mortality in populations of cultured fish. Rapid detection of viruses facilitates management decisions and reduces the risk of disease transmission to neighboring tanks, ponds, cages or farm sites. Molecular techniques such as the CyHV-3 qORF89 test detect DNA from intact virions and nascent genomes as well as residual nucleic acid fragments. A positive result therefore does not imply active replication but can inform stakeholders that an apparently healthy fish population has been exposed to the virus. Fish that have been exposed to CyHV-3 can become persistently infected and thus pose a fish health risk, as the latent virus can be reactivated, propagated and shed into the water (St-Hilaire et al. 2005). qPCR tests also provide information such as virus load, which can be used to predict disease progression, distinguish symptomatic from asymptomatic infection and assess the efficacy of a vaccination or antiviral treatment program (Gullett & Nolte 2015). Diagnosis of human herpesviruses such as cytomegalovirus and Epstein-Barr virus by qPCR tests includes an estimate of the viral load to distinguish between asymptomatic infections and actively replicating virus (Pavšič et al. 2015). Cutoff values to discriminate between colonization (i.e. asymptomatic) and disease causation (i.e. actively replicating virus) have been established for these viruses (Kraft et al. 2012, Pavšič et al. 2015). The challenges associated with the application of qPCR Cq results to clinical management decisions are significant (Kraft et al. 2012, Pillet et al. 2014). The Fisheries and Oceans Canada (DFO) laboratories, which are members of Canada's National Aquatic Animal Health Laboratory System, report qORF89 diagnostic test results as detected (Cq <40), not detected (no detectable amplification in all technical replicates) or inconclusive (amplification is detected in some but not all technical replicates). In the latter case, additional diagnostic testing would be required. The laboratories do not report the Cq values.

Acknowledgements. This study was supported by the Centre of Expertise for Aquatic Animal Health Research and Diagnostics, DFO. We thank Ron Hedrick and Tomo Kurobe for providing the CyHV-1, CyHV-2, CyHV-3, SalHV-2, AciHV-1, AciHV-2 and IctHV-1 genomic DNA; Mohamed Faisal for providing CyHV-3 U-MI2011; and Evi Emennegger and David Stone for providing isolates of SVCV and PFRV. Special thanks are extended to Dale Wrubleski, who coordinated the collection of wild common carp from Delta Marsh.

LITERATURE CITED

Avarre JC, Madeira JP, Santika A, Zainun Z and others (2011) Investigation of cyprinid herpesvirus-3 genetic diversity by a multi-locus variable number of tandem repeats analysis. J Virol Methods 173:320–327

Barnhart HX, Haber MJ, Lin LI (2007) An overview on assessing agreement with continuous measurements. J Biopharm Stat 17:529–569

Bercovier H, Fishman Y, Nahary R, Sinai S and others (2005) Cloning of the koi herpesvirus (KHV) gene encoding thymidine kinase and its use for a hig hly sensitive PCR based diagnosis. BMC Microbiol 5:13–21

Bergmann SM, Riechardt M, Fichtner D, Lee P, Kempter J (2010) Investigation on the diagnostic sensitivity of molecular tool s used for detection of koi herpesvirus. J Virol Methods 163:229–233

Bigarré L, Baud M, Cabon J, Antychowicz J and others (2009) Differentiation between cyprinid herpesvirus type-3 lineages using duplex PCR. J Virol Methods 158: 51–57

Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327:307–310

Boutier M, Ronsmans M, Rakus K, Jazowiecka-Rakus J and others (2015) Cyprinid herpesvirus 3: an archetype of fish alloherpesviruses. Adv Virus Res 93:161–256

Bretziner A, Fischer-Scherl T, Oumouna M, Hoffmann R, Truyen U (1999) Mass mortalities in koi carp, *Cyprinus carpio*, associated with gill and skin disease. Bull Eur Assoc Fish Pathol 19:182–185

Bustin SA, Benes V, Garson JA, Hellemans J and others (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622

Caraguel C, Stryhn H, Gagne N, Dohoo I, Hammell L (2009) Traditional descriptive analysis and novel visual representation of diagnostic repeatability and reproducibility: application to an infectious salmon anaemia virus RT-PCR assay. Prev Vet Med 92:9–19

Caraguel C, Stryhn H, Gagne N, Dohoo I, Hammell L (2012) Use of a third class in latent class modelling for the diagnostic evaluation of five infectious salmon anaemia virus detection tests. Prev Vet Med 104:165–173

CFIA (Canadian Food Inspection Agency) (2016) Federally reportable aquatic animal diseases in Canada — 2016. Available at www.inspection.gc.ca/animals/aquatic-animals/diseases/reportable/2016/eng/1339174937153/1339175227861 (accessed on 12 May 2016)

Cobo F (2012) Application of molecular diagnostic techniques for viral testing. Open Virol J 6:104–114

Cohen JA (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20:37–46

Costes B, Raj VS, Michel B, Fournier G and others (2009) The major portal of entry of koi herpesvirus in *Cyprinus carpio* is the skin. J Virol 83:2819–2830

Davison AJ (2010) Herpesvirus systematics. Vet Microbiol 143:52–69

Davison AJ, Eberle R, Ehlers B, Hayward GS and others (2009) The order *Herpesvirales.* Arch Virol 154:171–177

Dishon A, Perelberg A, Bishara-Shieban J, Ilouze M, Davidovich M, Werker S, Kotler M (2005) Detection of carp interstitial nephritis and gill necrosis virus in fish droppings. Appl Environ Microbiol 71:7285–7291

Dohoo IR, Martin W, Stryhn H (2009) Veterinary epidemiologic research, 2nd edn. VER Inc, Charlottetown

Eide KE, Miller-Morgan T, Heidel JR, Kent ML and others (2011) Investigation of koi herpesvirus latency in koi. J Virol 85:4954–4962

Enøe C, Georgiadis MP, Johnson WO (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev Vet Med 45:61–81

Fijan NN, Wellborne TL Jr, Naftel JP (1970) An acute viral disease of channel catfish. US Fish Wildl Serv Tech Pap 43, Washington, DC

Fijan N, Sulimanovi D, Bearzotti M, Mužini D and others (1983) Some properties of the *Epithelioma papulosum cyprini* (EPC) cell line from carp *Cyprinus carpio*. Ann Inst Pasteur Virol 134:207–220

Fournier G, Boutier M, Raj VS, Mast J and others (2012) Feeding *Cyprinus carpio* with infectious materials mediates cyprinid herpesvirus 3 entry through infection of pharyngeal periodontal mucosa. Vet Res 43:6–15

Gardner IA, Stryhn H, Lind P, Collins MT (2000) Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. Prev Vet Med 45:107–122

Gardner IA, Whittington RJ, Caraguel CGB, Hick P and others (2016) Recommended reporting standards for test accuracy studies of infectious diseases of finfish, amphibians, molluscs and crustaceans: the STRADAS-aquatic checklist. Dis Aquat Org 118:91–111

Garver KA, Al-Hussinee L, Hawley LM, Schroeder T and others (2010) Mass mortality associated with koi herpesvirus in wild common carp in Canada. J Wildl Dis 46: 1242–1251

Gelman AE, Carlin J, Stern H, Rubin D (2003) Bayesian data analysis, 2nd edn. Chapman & Hall/CRC, Boca Raton, FL

Gilad O, Yun S, Adkison MA, Way K, Willits NH, Bercovier H, Hedrick RP (2003) Molecular comparison of isolates of an emerging fish pathogen, koi herpesvirus, and the effect of water temperature on mortality of experimentally infected koi. J Gen Virol 84:2661–2668

Gilad O, Yun S, Zagmutt-Vergara FJ, Leutenegger CM, Bercovier H, Hedrick RP (2004) Concentrations of a Koi herpesvirus (KHV) in tissues of experimentally-infected *Cyprinus carpio koi* as assessed by real-time TaqMan PCR. Dis Aquat Org 60:179–187

Gullett JC, Nolte FS (2015) Quantitative nucleic acid amplification methods for viral infections. Clin Chem 61:72–78

Haenen OLM, Way K, Bergmann SM, Ariel E (2004) The emergence of koi herpesvirus and its significance to European aquaculture. Bull Eur Assoc Fish Pathol 24: 293–307

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hanson L, Dishon A, Kotler M (2011) Herpesviruses that

infect fish. Viruses 3:2160–2191
Hedrick RP, McDowell TS, Groff JM, Yun S, Wingfield WH (1991) Isolation of an epitheliotropic herpesvirus from white sturgeon *Acipenser transmontanus*. Dis Aquat Org 11:49–56
Hedrick RP, Gilad O, Yun S, Spangenberg JV and others (2000) A herpesvirus associated with mass mortality of juvenile and adult koi, a strain of common carp. J Aquat Anim Health 12:44–55
Hui SL, Walter SD (1980) Estimating the error rates of diagnostic tests. Biometrics 36:167–171
Ilouze M, Dishon A, Kotler M (2006) Characterization of a novel virus causing a lethal disease in carp and koi. Microbiol Mol Biol Rev 70:147–156
Jung SJ, Miyazaki T (1995) Herpesviral haematopoietic necrosis of goldfish, *Carassius auratus* (L.). J Fish Dis 18: 211–220
Karber G (1931) Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. Naunyn Schmiedebergs Arch Exp Pathol Pharmakol 162:480–483
Kit S, Sheppard M, Ichimura H, Nusinoff-Lehrman S, Ellis MN, Fyfe JA, Otsuka H (1987) Nucleotide sequence changes in thymidine kinase genes of herpes simplex virus type 2 clones from an isolate of a patient treated with Acyclovir. Antimicrob Agents Chemother 31: 1483–1490
Kraft CS, Armstrong WS, Caliendo AM (2012) Interpreting quantitative cytomegalovirus DNA testing: understanding the laboratory perspective. Clin Infect Dis 54: 1793–1797
Kurita J, Yuasa K, Ito T, Sano M and others (2009) Molecular epidemiology of koi herpesvirus. Fish Pathol 44:59–66
Kurobe T, Kelley GO, Waltzek TB, Hedrick RP (2008) Revised phylogenetic relationships among herpesviruses isolated from sturgeons. J Aquat Anim Health 20:96–102
Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
Lin LI (1989) A concordance correlation coefficient to evaluate reproducibility. Biometrics 45:255–268
Lin LI (2000) A note on the concordance correlation coefficient. Biometrics 56:324–325
Lycke E (1990) Biological and molecular aspects on herpes simplex virus latency. Scand J Infect Dis Suppl 69: 113–119
McClure CA, Hammell KL, Stryhn H, Dohoo IR, Hawkins LJ (2005) Application of surveillance data in evaluation of diagnostic tests for infectious salmon anemia. Dis Aquat Org 63:119–127
Miller O, Fuller FJ, Gebreyes WA, Lewbart GA and others (2007) Phylogenetic analysis of spring viremia of carp virus reveals distinct subgroups with common origins for recent isolates in North America and the UK. Dis Aquat Org 76:193–204
Miyazaki T, Kuzuya Y, Yasumoto S, Yasuda M, Kobayashi T (2008) Histopathological and ultrastructural features of Koi herpesvirus (KHV)-infected carp *Cyprinus carpio*, and the morphology and morphogenesis of KHV. Dis Aquat Org 80:1–11
Monaghan SJ, Thompson KD, Adams A, Bergmann SM (2015a) Sensitivity of seven PCRs for early detection of koi herpesvirus in experimentally infected carp, *Cyprinus carpio* L., by lethal and non-lethal sampling methods. J Fish Dis 38:303–319
Monaghan SJ, Thompson KD, Adams A, Kempter J, Bergmann SM (2015b) Examination of the early infection stages of koi herpesvirus (KHV) in experimentally infected carp, *Cyprinus carpio* L. using *in situ* hybridization. J Fish Dis 38:477–489
Nérette 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. J Fish Dis 28:89–99
Nérette P, Stryhn H, Dohoo I, Hammell L (2008) Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests. Prev Vet Med 85:207–225
Neukirch M, Böttcher K, Bunnajjirakul S (1999) Isolation of a virus from koi with altered gills. Bull Eur Assoc Fish Pathol 19:221–224
OIE (World Organisation for Animal Health) (2015a) Aquatic animal health code 2015, Chap 1.3. Diseases listed by the OIE. Available at www.oie.int/en/international-standard-setting/aquatic-code/access-online/ (accessed on 12 May 2016)
OIE (2015b) Manual of diagnostic tests for aquatic animals 2015, Chap 1.1.2. Principles and methods of validation of diagnostic assays for infectious diseases. Available at www.oie.int/international-standard-setting/aquatic-manual/access-online/ (accessed on 11 May 2016)
OIE (2015c) Manual of diagnostic tests for aquatic animals 2015, Chap 2.3.7. Koi herpesvirus disease. Available at www.oie.int/international-standard-setting/aquatic-manual/access-online/ (accessed on 12 May 2016)
Pavšič J, Devonshire AS, Parkes H, Schimmel H and others (2015) Standardization of nucleic acid tests for clinical measurements of bacteria and viruses. J Clin Microbiol 53:2008–2014
Perelberg A, Smirnov M, Hutoran M, Diamant A, Bejerano Y, Kotler M (2003) Epidemiological description of a new viral disease afflicting cultured *Cyprinus carpio* in Israel. Isr J Aquacult Bamidgeh 55:5–12
Pillet S, Roblin X, Cornillon J, Mariat C, Pozzetto B (2014) Quantification of cytomegalovirus viral load. Expert Rev Anti Infect Ther 12:193–210
Pouillot R, Guillaume G, Gardner IA (2002) 'TAGS', a program for the evaluation of test accuracy in the absence of a gold standard. Prev Vet Med 53:67–81
Prescott MA, Reed AN, Jin L, Pastey MK (2016) Rapid detection of cyprinid herpesvirus 3 in latently infected koi by recombinase polymerase amplification. J Aquat Anim Health 28:173–180
Raj VS, Fournier G, Rakus K, Ronsman M and others (2011) Skin mucus of *Cyprinus carpio* inhibits cyprinid herpesvirus 3 binding to epidermal cells. Vet Res 42:92–100
Ratcliff RM, Chang G, Kok T, Sloots TP (2007) Molecular diagnosis of medical viruses. Curr Issues Mol Biol 9: 87–102
Reed AN, Izume S, Dolan BP, LaPatra S, Kent M, Dong J, Jin L (2014) Identification of B cells as a major site for cyprinid herpesvirus 3 latency. J Virol 88:9297–9309
Reichenheim ME (2004) Confidence intervals for the kappa statistic. Stata J 4:421–428
Ronsmans M, Boutier M, Rakus K, Farnir F and others (2014) Sensitivity and permissivity of *Cyprinus carpio* to cyprinid herpesvirus 3 during the early stages of its development: importance of the epidermal mucus as an innate immune barrier. Vet Res 45:100–111
Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sano T, Fukuda H, Okamoto N, Kaneko F (1983) Yamame tumor virus: lethality and oncogenicity. Bull Jpn Soc Sci Fish 49:1159–1163

Sano T, Fukuda H, Furukawa M, Hosoya H, Moriya Y (1985) A herpesvirus isolated from carp papilloma in Japan. Fish Shell Pathol 32:307–311

Sasadeusz JJ, Tufaro F, Safrin S, Schubert K, Hubinette MM, Cheung PK, Sacks SL (1997) Homopolymer mutational hot spots mediate herpes simplex virus resistance to Acyclovir. J Virol 71:3872–3878

Snow M, McKay P, Matejusova I (2009) Development of a widely applicable positive control strategy to support detection of infectious salmon anaemia virus (ISAV) using Taqman real-time PCR. J Fish Dis 32:151–156

Spearman C (1908) The method of 'right and wrong cases' ('constant stimuli') without Gauss's formulae. Br J Psychol 2:227–242

Spiegelhalter DJ, Best NG, Carlin BP, van der Lind A (2002) Bayesian measures of model complexity and fit. J R Stat Soc B 64:583–639

Spiegelhalter DJ, Thomas A, Best NG, Lunn D (2003) WinBUGS version 1.4 user manual. MRC Biostatistic Unit, www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

St-Hilaire S, Beevers N, Way K, Le Deuff RM, Martin P, Joiner C (2005) Reactivation of koi herpesvirus infections in common carp *Cyprinus carpio*. Dis Aquat Org 67: 15–23

Stone DM, Ahne W, Denham KL, Dixon PF and others (2003) Nucleotide sequence analysis of the glycoprotein gene of putative spring viraemia of carp virus and pike fry rhabdovirus isolates reveals four genogroups. Dis Aquat Org 53:203–210

Sunarto A, McColl KA, Crane MStJ, Schat KA, Slobedman B, Barnes AC, Walker PJ (2014) Characteristics of cyprinid herpesvirus 3 in different phases of infection: implications for disease transmission and control. Virus Res 188:45–53

Tenser RB (1991) Role of herpes simplex virus thymidine kinase expression in viral pathogenesis and latency. Intervirology 32:76–92

Uchii K, Matsui K, Iida T, Kawabata Z (2009) Distribution of the introduced cyprinid herpesvirus 3 in a wild population of common carp, *Cyprinus carpio* L. J Fish Dis 32:857–864

Uchii K, Telschow A, Minamoto T, Yamanaka H, Honjo MN, Matsui K, Kawabata Z (2011) Transmission dynamics of an emerging infectious disease in wildlife through host reproductive cycles. ISME J 5:244–251

Uchii K, Minamoto T, Honjo MN, Kawabata Z (2014) Seasonal reactivation enables *Cyprinid herpesvirus 3* to persist in a wild host population. FEMS Microbiol Ecol 87: 536–542

Walster C (1999) Clinical observations of severe mortalities in koi carp, *Cyprinus carpio*, with gill disease. Fish Vet J 3:54–58

Waltzek TB, Kelley GO, Stone DM, Way K and others (2005) Koi herpesvirus represents a third cyprinid herpesvirus (CyHV-3) in the family *Herpesviridae*. J Gen Virol 86: 1659–1667

Waltzek TB, Kelley GO, Alfaro ME, Kurobe T, Davison AJ, Hedrick RP (2009) Phylogenetic relationships in the family *Alloherpesviridae*. Dis Aquat Org 84:179–194

Way K (2008) Report of the 2008 international KHV PCR methods ring-trial. Centre for Environment, Fisheries and Aquaculture Science, Weymouth

Wilson IG (1997) Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63:3741–3751

Winton J, Batts W, deKinkelin P, LeBerre M, Bremont M, Fijan N (2010) Current lineages of the epithelioma papulosum cyprinid (EPC) cell line are contaminated with fathead minnow, *Pimephales promelas* cells. J Fish Dis 33: 701–704

Yamada S, Matsumoto Y, Takashima Y, Otsuka H (2005) Mutation hot spots in the canine herpesvirus thymidine kinase gene. Virus Genes 31:107–111

Editorial responsibility: James Jancovich, San Marcos, California, USA

Submitted: August 23, 2016; Accepted: December 2, 2016
Proofs received from author(s): February 23, 2017