Wǔhàn sharpbelly bornavirus (genus *Cultervirus*) can infect a wide range of host cells and establishes persistent infection in diverse cypriniform cells

Mirette I. Y. Eshak[1], Angele Breithaupt[2], Birke A. Tews[3], Christine Luttermann[3], Kati Franzke[3], Marion Scheibe[4], Sören Woelke[3], Dennis Rubbenstroth[1], Martin Beer[1] and Florian Pfaff[1,*]

1: Friedrich-Loeffler-Institut, Institute of Diagnostic Virology, Greifswald - Insel Riems, Germany

2: Friedrich-Loeffler-Institut, Department of Experimental Animal Facilities and Biorisk Management, Greifswald - Insel Riems, Germany

3: Friedrich-Loeffler-Institut, Institute of Infectology, Greifswald - Insel Riems, Germany

4: Friedrich-Loeffler-Institut, Institute of Molecular Virology and Cell Biology, Greifswald - Insel Riems, Germany

*Corresponding author: Florian Pfaff

E-mail: florian.pfaff@fli.de

Tel: +49 38351 7 1385

Fax: +49 38351 7 1226

Abstract word count: 249

Importance word count: 148

Main text word count: 8945

Figure count: 8

Reference count: 41

Keywords: *Bornaviridae*, *Cultervirus*, fish, Cypriniformes, metagenomics, infection, persistence

ABSTRACT

Our recent study using *in silico* data mining identified novel culterviruses (*Bornaviridae*) in fish, including a variant of Wǔhàn sharpbelly bornavirus (WhSBV) in grass carp kidney and liver cell lines. Here, metagenomic sequencing of different fish cell lines revealed WhSBV in two additional grass carp swim bladder cell lines, GSB and CiSB. Using these cell lines, we investigated the ability of WhSBV to infect and establish persistent infection in other cell lines from bony fish (Cypriniformes, Chichliformes, Salmoniformes, Centrarchiformes and Spariformes), reptiles (Testudines and Squamata), birds (Galliformes) and mammals (Primates and Rodentia). WhSBV showed efficient replication and a time-dependent increase in viral RNA levels in cypriniform cells, whereas replication was limited, confined to single cells, and lacked a clear time-dependent increase in certain bony fish and reptilian cells. No persistent replication was detected in avian or mammalian cells. *In situ* hybridisation and electron microscopy confirmed the presence of viral RNA and particles in infected fish cells. Transcriptomic sequencing revealed minimal innate immune activation during early stages of infection and antiviral response only at later stages, suggesting that WhSBV establishes persistence by evading early immune recognition. In addition, we identified polycistronic viral mRNAs regulated by specific transcriptional start and termination sites and RNA splicing. Viral proteins were detected, confirming previous *in silico* predictions. These findings provide insights into the potential infectivity, persistence mechanisms and transcriptional strategies of WhSBV. This study validates previous findings from *in silico* data mining, further reinforcing its effectiveness as a powerful tool for discovering hidden viruses.

IMPORTANCE

Understanding the diversity and host range of viruses is crucial for assessing their ecological role, associated diseases and zoonotic potential. However, many newly discovered viruses are characterised using sequence data alone because isolates are often difficult to obtain. Using cell culture models, this study characterises Wǔhàn sharpbelly bornavirus (WhSBV), a member of the genus *Cultervirus*. Here we demonstrate its ability to establish persistent infections in cypriniform fish cell lines, while exhibiting restricted replication in non-cypriniform fish and reptilian hosts. The identification of polycistronic transcription, splicing events and immune evasion mechanisms advances our understanding of the molecular biology of WhSBV and

culterviruses in general. By validating *in silico* predictions, this study highlights the power of computational approaches in uncovering viral diversity. As cypriniform fish include economically important species such as carp, understanding the dynamics of WhSBV host range and infection biology may be crucial for future aquaculture health management.

INTRODUCTION

The family *Bornaviridae*, within the order *Mononegavirales*, contains four genera: *Orthobornavirus*, *Carbovirus*, *Cultervirus*, and *Cartilovirus* (Kuhn et al. 2015; Eshak et al. 2023; Rubbenstroth et al. 2021). The genus *Orthobornavirus* has the broadest known host range, infecting avian, reptilian, and mammalian species (Rubbenstroth et al. 2021). Furthermore, orthobornaviruses include zoonotic viruses, such as Borna disease virus 1 (BoDV-1) and variegated squirrel bornavirus 1 (VSB-1), both of which can cause fatal encephalitis in humans (Niller et al. 2020; Hoffmann et al. 2015). Carboviruses and culterviruses, on the other hand, have only been found in reptiles and fish, respectively, and no zoonotic spillover has been reported (Rubbenstroth et al. 2021; Shi et al. 2018; Eshak et al. 2023; Hyndman et al. 2018). The genus *Cartilovirus* currently only includes the little skate bornavirus (LSBV), which was discovered using public datasets from the Sequence Read Archive (SRA) (Eshak et al. 2023).

Bornavirids possess enveloped virions that enclose a linear, negative-sense, monopartite RNA molecule of approximately 9,000 nucleotides (nt) (Briese et al. 1994; Rubbenstroth et al. 2021). Their genomes encode for at least six open reading frames (ORFs) arranged in the order 3'-N-X/P-M-G-L-5' for orthobornaviruses or 3'-N-X/P-G-M-L-5' for carboviruses and culterviruses. These ORFs encode for six viral proteins; nucleoprotein (N), accessory protein (X), phosphoprotein (P), matrix protein (M), glycoprotein (G), and the large protein (L) which contains a RNA-directed RNA polymerase (RdRP) (Briese et al. 1994; Rubbenstroth et al. 2021). Recently, the largest bornavirid genome, comprising 11,090 nt, was identified for the cartilovirus LSBV. The LSBV genome encodes two additional possible ORFs, viral proteins 1 and 2 (vp1 and -2), and represents a unique genomic architecture 3'-N-vp1-vp2-X/P-G-M-L-5' (Eshak et al. 2023). Studies of BoDV-1 have demonstrated that its RNA and genome are transcribed and replicated in the host cell nucleus, respectively (Briese et al. 1992). In bornavirids, the generation of a diverse array of mRNAs is regulated by differential use of transcription start and termination sites (Schneemann et al. 1994), as well as alternative splicing of polycistronic primary transcripts to increase and control their transcriptional complexity (Schneider et al. 1994; Tomonaga et al. 2000).

63 Wǔhàn sharpbelly bornavirus (WhSBV, NC_055169; species *Cultervirus hemicultri*), a member
64 of the genus *Cultervirus*, was initially identified by RNA sequencing of the gut, liver, and gill
65 tissues from a freshwater sharpbelly (*Hemiculter leucisculus* [Basilewsky, 1855], family
66 Xenocyprididae, order Cypriniformes) from China (Shi et al. 2018; Rubbenstroth et al. 2021).
67 Our recent study using *in silico* data mining of publicly available sequencing datasets led to the
68 identification of novel fish culterviruses, including the first complete genome of the Murray-
69 Darling carp bornavirus (MDCBV, BK063521) from a goldfish (*Carassius auratus* [Linnaeus,
70 1758], family Cyprinidae, order Cypriniformes) tissue pool data set and a WhSBV variant
71 (BK063520) detected in grass carp (*Ctenopharyngodon idella* [Valenciennes, 1844], family
72 Xenocyprididae, order Cypriniformes) CIK (kidney) and L8824 (liver) cell lines (Eshak et al.
73 2023). The two viruses belong to the same bornavirid species (*Cultervirus hemicultri*) and
74 share 78% nt overall identity. However, current knowledge of these culterviruses remains
75 limited to *in silico* genetic discoveries and bioinformatic predictions, with no virus isolates
76 available for experimental research.

77 To deepen our understanding of the biology and abundance of culterviruses, we screened 48
78 fish-derived cell lines for the presence of known or novel culterviruses using metagenomic
79 RNA sequencing. This revealed the presence of WhSBV in two grass carp swim bladder cell
80 lines, GSB (CCLV no. 1441) and CiSB (1542), with genomes nearly identical to the WhSBV
81 variant from the liver and kidney cell lines (BK063520). This study represents the first
82 investigation into the biology of WhSBV and lays the groundwork for future research to further
83 explore culterviruses.

84 # MATERIALS AND METHODS

85 ## Nucleic acids extraction and metagenomic sequencing

86 A total of 48 fish cell lines were provided by the Collection of Cell Lines in Veterinary Medicine
87 (CCLV) (see Supplementary Table S1 for a complete list of cell lines) at the Friedrich-Loeffler-
88 Institut, Riems, Germany. Fresh-frozen cell lines were lysed in 1 ml TRIzol LS reagent (Life
89 Technologies) and shaken vigorously for 10 min at room temperature. After addition of 200 µl
90 chloroform (Carl Roth GmbH & Co. KG) and centrifuging at 13,000 × g, 10 min, 4°C, 200 µl of
91 the upper aqueous phase was used for RNA extraction. Total RNA was extracted using the
92 Agencourt RNAdvance Tissue Kit (Beckman Coulter) with a KingFisher Flex Purification
93 System (Thermo Fisher Scientific) according to the manufacturer's instructions. DNase I
94 digestion (Qiagen) was carried out prior to the final elution to remove genomic DNA residues.
95 The quantity of total RNA was determined using the NanoDrop Lite Spectrophotometer
96 (Thermo Fisher Scientific). Poly(A) RNA was isolated using the Dynabeads mRNA DIRECT
97 Purification Kit (Invitrogen) following the manufacturer's instructions. Fragmentation and
98 strand-specific construction of whole transcriptome libraries was performed using the Collibri
99 Stranded RNA Library Prep Kit for Illumina Systems (Invitrogen). The quality and integrity of
100 the RNA and the final library were verified using the Agilent TapeStation 4150 (Agilent
101 Technologies) with appropriate chips and reagents. The final libraries were quantified using
102 the Qubit dsDNA HS Assay Kit (Invitrogen) in combination with the Qubit 2.0 Fluorometer
103 (Invitrogen). Equimolar amounts of the library were then sequenced using Novaseq 6000 with
104 SP flow cell, 1 x 101 bp (CeGat GmbH) according to the manufacturer's instructions.

105 ## Metagenomic analysis

106 The raw reads obtained from the 48 fish cell lines were trimmed using *TrimGalore!* (v0.6.10)
107 and *cutadapt* (v4.0). The metagenomic pipeline *SqueezeMeta* (v1.6.4) was then used to
108 individually assemble the trimmed reads from each dataset and then sequentially merge them
109 into a single transcriptome ("seqmerge" mode). The assembled and merged transcripts were
110 then filtered for poly(A/T) sequences at their termini using *cutadapt* (v4.0). The filtered
111 transcriptome was then used as input for taxonomic assignment and abundance estimation

112 using *SqueezeMeta* (v1.6.4, database version Sep-3-2023). The results were further analysed
113 using the R (v4.3.1) package “SQMtools” (v1.6.3).

114 **Validation of WhSBV genome termini**

115 Persistently WhSBV-infected GSB and CiSB cell cultures (see Supplementary Table S1) were
116 grown in non-vented T25 tissue culture flasks (Corning, Inc.) in a suitable growth medium
117 supplemented with 10% fetal bovine serum (FBS) to 95-100% confluence. Total RNA was
118 extracted using the TRIzol-chloroform method, followed by purification using the RNeasy Mini
119 Kit (Qiagen) and DNase I digestion prior to elution. The quality and quantity of the RNA were
120 assessed using the Agilent TapeStation 4150 in combination with appropriate chips and
121 reagents, and the NanoDrop Lite spectrophotometer, respectively.

122 The 5'-Rapid Amplification of cDNA Ends (RACE) kit, version 2.0 (Invitrogen), was used to
123 determine the 3' end of the WhSBV genome. Reverse transcription polymerase chain reaction
124 (RT-PCR) was performed on the extracted RNA using a gene-specific primer WhSBV-945-R
125 (0.4 µM) (see Supplementary Table S2 for a complete list of primers). The resulting cDNA was
126 then digested with RNase cocktail (Invitrogen) at 37°C for 20 min and subsequently purified
127 using AMPure XP beads (Beckman Coulter). A tailing reaction was then performed using dCTP
128 and dATP (New England Biolabs). A second round of amplification was performed using gene-
129 specific primers WhSBV-777-R (0.2 µM) and the AAP primer (0.2 µM) for purified C-tailed
130 cDNA or the adapter primer (AP) (0.2 µM) for purified A-tailed cDNA (see Supplementary Table
131 S2). The hemi-nested PCR was performed with gene-specific primers WhSBV-545-R (0.2 µm)
132 and the AUAP primer (0.2 µm) (see Supplementary Table S2). To identify the 5' end, total RNA
133 samples were treated with the *E. coli* Poly(A) Polymerase Kit (New England Biolabs) according
134 to the manufacturer's instructions in order to add poly(A) tails. The poly(A) tailed RNA were
135 then purified using AMPure XP beads. This purified poly(A) tailed RNA was used as input to
136 the 3'-RACE system for rapid amplification of cDNA ends (Invitrogen), using the AP primer (see
137 Supplementary Table S2) for RT-PCR. The cDNA was then digested with an RNase cocktail
138 and purified using AMPure XP beads. For the second round of amplification, six different gene-
139 specific primers (WhSBV_7388F, WhSBV_7528F, WhSBV_7911F, WhSBV_8215F,
140 WhSBV_8737F, WhSBV_8901F; 0.2 µm) were combined with the AUAP primer (see
141 Supplementary Table S2). The resulting PCR amplicons were separated by 2% agarose gel
142 electrophoresis. The remaining product was purified with AMPure XP beads and sequenced
143 using the BigDye Terminator v1.1 Cycle Sequencing Kit on a 3500 Genetic Analyzer (Applied

144 Biosystems). The final WhSBV genome sequences were uploaded to Genbank under
145 accession number PV171101.

146 **RT-qPCR and qPCR assays for specific detection of WhSBV**

147 For the following assays, persistently WhSBV-infected GSB and CiSB cells and the WhSBV-
148 uninfected GK (0119) cell line (see Supplementary Table S1) were grown to 95-100%
149 confluence in non-vented T25 tissue culture flasks using appropriate growth medium
150 supplemented with 10% FBS. Cell pellets were harvested using Alsever's trypsin-versus-
151 versus-solution (ATV; FLI, Riems) after discarding the supernatant, and RNA/DNA was
152 extracted using the NucleoMag VET kit (Macherey-Nagel) according to the manufacturer's
153 instructions.

154 WhSBV-specific reverse transcription quantitative polymerase chain reaction (RT-qPCR)
155 assays were established by designing primers and probe sets targeting the viral N, P, G, M and
156 L genes (see Supplementary Table S2, for a complete list of primers and probes). WhSBV-
157 specific RNA (and potentially DNA) was detected using the AgPath-ID one-step RT-PCR
158 reagents (Thermo Fisher Scientific). Briefly, 5 µl of extracted RNA was reverse transcribed and
159 amplified in a reaction mix of 25 µl total volume containing the specific primer/probe mix (final
160 primer and probe concertation: 0.4 and 0.2 µM, respectively). For the internal control, primers
161 ACT-1030-F (final concentration: 0.2 µM) and ACT-1135-R (0.2 µM) along with probe ACT-
162 1081-HEX (0.2 µM) were included in each reaction to target the cellular actin beta (*ACTB*)
163 gene. The reaction was performed on a Bio-Rad CFX96 qPCR cycler (Bio-Rad) using the
164 following cycling conditions: 45°C for 10 min, 95°C for 10 min, 42 cycles of 95°C for 15 s, 57°C
165 for 20 s and 72°C for 30 s.

166 To exclude the presence of WhSBV sequences as endogenous viral elements (EVE) integrated
167 into the cellular genome of GSB, CiSB, and GK, we specifically tested DNA extracted from
168 these cells using a non-reverse transcription qPCR system. WhSBV cDNA from the 3'-RACE
169 experiment, as previously described, was used as a positive control. The Maxima Probe/ROX
170 qPCR Master Mix (Thermo Fisher Scientific) was used according to the manufacturer's
171 protocol, targeting the viral genes N, P, G, M and L (final primer and probe concertation: 0.4
172 and 0.2 µM, respectively) and the cellular *ACTB* gene (final primer and probe concertation: 0.2
173 and 0.2 µM, respectively). The reaction was performed on a Bio-Rad CFX96 qPCR cycler using
174 the following cycling conditions: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s,
175 60°C for 30 s and 72°C for 30 s.

176 **Experimental infection of different cell cultures with WhSBV**

177 Bony fish cells RT/F (CCLV no. 0088), CaPi (0112), EPC (0173), RTG-2/f (0686), CCB (0816),
178 SAF-1 (0826), CHSE-214 (1104), GTS-9 (1388), ZF4 (1492), and TiB (1550), the reptilian cells
179 SKL-R (0484), VH-2 (1092), and CDSK (1536), the avian cells QM7 (0466) and DF-1 (1529),
180 and the mammalian cells Vero (0015) and C6 (1452) were obtained from CCLV (see
181 Supplementary Table S3 for a complete list of cell lines). All fish and reptile cell lines were
182 grown in non-vented T25 flasks (Corning) to 95-100% confluence in an appropriate 10 ml
183 growth medium supplemented with 10% FBS. Incubation temperatures were set at 20°C for
184 the fish cells CHSE-214 and RTG-2/f, 26°C for all other fish cells and reptilian cell SKL-R, and
185 28°C for the reptile cells VH-2 and CDSK. Mammalian and avian cells were grown in vented
186 T25 flasks at 37°C in a humidified atmosphere containing 5% CO_2.

187 For inoculum preparation, several non-vented T75 (75 cm²) culture flasks (Corning) containing
188 persistently WhSBV-infected GSB cells with 5 ml of supernatant (equivalent to 1 ml of inoculum
189 per 15 m² of infected cell layer) were frozen at -80°C and thawed three times. All supernatants
190 were then combined into a single solution. To initiate the infection experiment, cells were grown
191 to confluence in T25 flasks (25 cm²) as described above. Prior to inoculation (T0), 1 ml of
192 supernatant was collected from each cell line. An equal volume (1 ml) of WhSBV inoculum was
193 then added to the cultures (corresponding to an infection ratio of 1:1.7 between the area of the
194 cell layer from which the inoculum originated and the area of the cell layer inoculated). Cultures
195 were then incubated for one hour under the above conditions. After this incubation, a further 1
196 ml of supernatant was collected (T1). Cells were then incubated and split at regular intervals
197 of 2-3 days at a 2-fold splitting factor for a total of 4-6 passages (P1-6). The splitting process
198 involved collecting the supernatant before each passage, washing and detachment of the cells
199 with ATV, and collecting half of the detached cells for analysis. The remaining cells were used
200 for further culture by adding fresh growth medium supplemented with 10% FBS. The collected
201 supernatants and cell pellets were stored at -80°C for subsequent RNA extraction and RT-
202 qPCR analysis. Negative controls were treated in the same way as infected samples, except
203 that the WhSBV inoculum was replaced by 1 ml RNase-free water (Qiagen). The infection
204 experiment was performed in two independent runs for each cell line. As the infection of reptile,
205 avian and mammalian cell lines was performed separately from the comparison of different fish
206 cell lines, the cypriniform cell line GTS-9 was used as a positive control to validate the infectivity
207 of the inoculum.

208 Nucleic acid extraction from supernatant and cell pellet samples was performed using the
209 NucleoMag VET kit according to the manufacturer's instructions. The extracted nucleic acids

210 were used for RT-qPCR analysis as described above, using the G gene-specific RT-qPCR
211 primer/probe mix to monitor viral load and the ACTB primer/probe mix as an internal control
212 for normalisation and to verify RNA integrity.

213 Electron microscopy

214 WhSBV-infected and uninfected ZF4 cells (see Supplementary Table S3) from the previous
215 infection experiment were transferred to non-vented T75 flasks for electron microscopy
216 analysis, along with the persistently WhSBV-infected GSB cells grown in non-vented T75 flasks.
217 Cells were fixed in 2.5% glutaraldehyde buffered in 0.1 M sodium cacodylate pH 7.2 (SERVA
218 Electrophoresis). 1% aqueous OsO_4 was used for post-fixation and 2.5% uranyl acetate in
219 ethanol for *en bloc* staining (SERVA Electrophoresis). After a graded dehydration in ethanol,
220 the samples were cleared in propylene oxide and infiltrated with Glycid Ether 100 (SERVA
221 Electrophoresis). For polymerisation, samples were incubated at 60°C for 3 days. Ultrathin
222 sections were transferred to formvar-coated nickel grids (slot grids; Plano). All grids were
223 counterstained with uranyl acetate and lead citrate before examination on a Talos F200i
224 transmission electron microscope (FEI) at an accelerating voltage of 80 kV.

225 Specific detection of viral RNA using RNA *in situ* hybridization

226 The cypriniform cell lines EPC and CCB and the non-cypriniform fish cells RTG-2/f and TiB
227 (see Supplementary Table S3) were maintained in non-vented T75 flasks in a suitable medium
228 supplemented with 10% FBS until they reached 95-100% confluence. Cultures were
229 maintained under the conditions described in the previous infection experiment. For infection,
230 3 ml of medium was replaced with an equal volume of inoculum to achieve an infection ratio of
231 1:1.7, as previously described. Cells were harvested from two independent experiments at
232 early and late time points: 0, 4, 8, 24, 48 hours post infection (hpi) for the early time points and
233 0, 72, 96, 120 and 240 hpi for the late time points. Harvesting was performed with a cell scraper
234 (Sarstedt). The harvested cells were collected in 50 ml Falcon tubes (Sarstedt), centrifuged at
235 3500 rpm for 15 minutes and the supernatants discarded. The cell pellets were then fixed in
236 10% neutral buffered formalin (Carl Roth) for at least 24 hours. After paraffin embedding, 4-µm
237 thick sections were processed for subsequent RNA *in situ* hybridisation (RNA ISH) using
238 RNAScope 2-5 HD Reagent Kit Red and custom-designed probes targeting WhSBV (-)RNA
239 (genomic RNA; WhSBV-L-O2, catalogue number 1567161-C1, Advanced Cell Diagnostics,
240 Hayward, USA) on all samples according to the manufacturer's instructions. In addition,
241 custom-designed probes targeting WhSBV (+)RNA (mRNA and antigenomic RNA; WhSBV-G-

242 O1, catalogue number 1567171-C1) were used on CCB samples. As technical controls, probes
243 against the genes for peptidylprolyl isomerase B (PPIB) and dihydrodipicolinate reductase
244 (DapB) were included in each run. All pages were scanned using a Hamamatsu S60 scanner
245 and evaluated using NDPview.2 plus software (version 2.8.24, Hamamatsu Photonics, K.K.
246 Japan) by a board-certified pathologist (AB, DiplECVP). The abundance of infected cells was
247 recorded as a percentage in a 10 × 10 grid (size 100 × 100 µm per grid cell). Thus, a positive
248 signal within a grid cell resulted in one percentage point. The distribution pattern of the viral
249 RNA was also scored using a "pattern score": 0 = none; 1 = chromogenic labelling in single
250 cells; 2 = labelling in cell clusters; 3 = labelling in confluent cell clusters; 4 = diffuse labelling.
251 Finally, the size of the chromogenic signal was determined as "signal score", 1 = fine-granular
252 or 2 = globular. The total pattern/signal score was calculated as the sum of the signal and
253 pattern scores.

254 In a separate kinetics experiment to correlate with the RNA ISH results, the same cell lines as
255 above were seeded in 12-well plates (3.8 cm^2 cell layer per well) (Corning) and maintained to
256 95-100% confluence in appropriate medium supplemented with 10% FBS. Cultures were
257 maintained at 26 or 20°C (for RTG-2/f) and 5% CO_2. Inoculation was performed at a 1:0.8
258 infection ratio as described above, with 300 µl of medium replaced by 300 µl of WhSBV
259 inoculum. Infection of each cell line was performed in three independent replicate wells for
260 each time point. Cell pellets were harvested at 0, 4, 8, 24, 48, 72, 96, 120, 144, 168, 192, 216
261 and 240 hpi. At each time point, supernatants were discarded and cell pellets were collected
262 by trypsinisation using ATV. RNA was extracted using the NucleoMag VET kit as described
263 previously. RT-qPCR was then performed using G gene and ACTB specific RT-qPCR
264 primer/probe mixes.

265 **Gene expression analysis of early and late WhSBV infection**

266 The CCB cells (see Supplementary Table S3) were used to study the effect of WhSBV on
267 cellular gene expression. Cells were seeded in 6-well plates (9.6 cm^2 cell layer per well)
268 (Corning) and cultured for 24 hours to 95-100% confluence in 10% FBS supplemented
269 medium. They were maintained under the conditions described in the previous kinetics
270 experiment. Infection with WhSBV was performed at a 1:1.3 ratio as described above by
271 replacing 500 µl of culture medium in each well with 500 µl of WhSBV inoculum. Negative
272 controls were inoculated with 500 µl of the same medium used for persistently WhSBV-infected
273 GSB cultivation, replacing 500 µl of the cell culture medium. Cell pellets were collected in two
274 independent experiments to assess early and late infection stages. In the early infection

275 experiment, mock-infected and WhSBV-infected cell pellets were collected at 4, 8, 24 and 48
276 hpi. For the late-stage experiment, samples were collected at 72, 96, 120 and 240 hpi. In
277 addition, pre-infection cell pellets were collected at 7 and 24 before inoculation, the latter
278 representing the 0 hpi time point. Infection was performed in three independent replicate wells
279 for each time point. Cell pellets were collected by trypsinisation after supernatants were
280 discarded. Total RNA was extracted using Trizol, chloroform and the Agencourt RNAdvance
281 Tissue Kit as previously described. RT-qPCR was performed on all samples using G-gene and
282 ACTB-specific RT-qPCR primer/probe mixes to validate samples prior to sequencing.
283 Transcriptomic libraries were prepared as previously described, but including ERCC spike-in
284 mix (Invitrogen) as an internal control. Subsequently, the libraries were pooled and sent for
285 sequencing on an Illumina NovaSeq 6000 system (Novogene GmbH) running in paired-end
286 150 bp mode. Raw reads were subsequently trimmed using *TrimGalore!* (v0.6.10) and then
287 analysed using the nf-core RNAseq pipeline (version 3.14.0; (Ewels et al. 2020)). In detail, the
288 common carp reference genome GCF_018340385.1 was used for splice-aware mapping with
289 STAR and subsequent quantification with salmon. The quantification data was then further
290 analysed in R (v4.3.1) and differentially expressed genes (DEGs) were deduced using the
291 DESeq2 package (v1.40). Log2-fold changes were shrinked using the function "lfcShrink". The
292 cut-off values for DEGs were set to absolute shrinked log2-fold change > 1 and adjusted *p-*
293 *value* < 0.05. To explore global gene expression patterns, principal component analysis (PCA)
294 was performed using R. Transcriptomic data have been deposited in ArrayExpress under E-
295 MTAB-14894.

296 **Validation of predicted splice sites in WhSBV genome**

297 The trimmed raw reads from the metagenomic RNA datasets from persistently WhSBV-
298 infected GSB and CiSB cells (see Supplementary Table S1) were mapped back to the
299 respective WhSBV genomes using the generic mapper Geneious Prime with medium
300 sensitivity settings (options: find structural variants, short insertions and deletions of any size)
301 to visualise the transcriptional profile of WhSBV and predict potential splice sites and introns.
302 To experimentally validate each predicted splice site, primers were designed to amplify the
303 specific genomic regions by RT-PCR. MEME (v5.5.2) was used to predict the conserved motifs
304 of potential transcription start and termination sites.

305 Purified RNA was extracted by RNeasy Mini kit as described above. cDNA from poly(A) RNA
306 was synthesized using the Protoscript II First Strand cDNA Synthesis Kit and oligo-dT primers
307 (New England Biolabs). The cDNA products were then digested with an RNase cocktail and

308 purified using AMPure XP beads. Conventional Taq-PCR was performed using the Accuprime
309 *Taq* DNA Polymerase System (Invitrogen) with primers specific for the predicted splice sites
310 within the viral M and L genes (see Supplementary Table S2). RNase-free water was used
311 instead of cDNA as a negative control. The amplified DNA products were run on a 2% agarose
312 gel and bands were excised. The DNA was extracted using a QIAquick Gel Extraction Kit
313 (Qiagen) and then sequenced using the BigDye Terminator v1.1 on a 3500 Genetic Analyzer.

314 **Specific detection of viral RNA using Northern blot**

315 RNA was extracted from the persistently WhSBV-infected GSB, WhSBV-infected and
316 uninfected EPC cells (see Supplementary Tables S1 and S3) from the previous infection
317 experiment using the TRIzol-chloroform method and the RNeasy Mini Kit with DNase I as
318 previously described. RNA quality and quantity were assessed using the Agilent Tapestation
319 4150 and NanoDrop Lite spectrophotometer as described previously. Subsequently, 2-4 µg of
320 total RNA was glyoxylated (Glyoxal, Sigma-Aldrich) at 56°C for 45 min before separation on a
321 0.9% denaturing agarose gel (4.7% formaldehyde, Carl Roth) in a phosphate buffer system.
322 RNA was transferred overnight to Hybond-N membranes (Cytiva Amersham) using SSC
323 transfer buffer. RNA was cross-linked to the membrane and viral RNA (mRNA and genomic
324 RNA) was detected using DNA probes. Probes were generated by RT-PCR using specific
325 forward and reverse primers against each gene of WhSBV (see Supplementary Table S2) using
326 the Platinum Script III One-step RT-PCR Kit (Thermo Fisher Scientific) according to the
327 manufacturer's instructions. The amplified cDNA products were then run on a 2% agarose gel.
328 After amplification, the products were purified using AMPure XP beads according to the
329 manufacturer's protocol. The quantity of purified products was measured using a NanoDrop
330 Lite spectrophotometer and labelled with ^{32}P-CTP (Hartmann-analytic) using a Nick translation
331 kit (GE Healthcare). Hybridisation was performed overnight at 60-65°C and the signal was
332 analysed after washing using the CR35 phosphoimager system (Dürr medical).

333 **Specific detection of WhSBV proteins**

334 Approximately 3×10^5 persistently WhSBV-infected GSB cells were harvested, resuspended in
335 1× LDS Buffer (Thermo Fisher Scientific) containing 100 mM DTT (Sigma-Aldrich) and heated
336 for 10 min at 70°C shaking at 1400 rpm in a Thermomixer (Eppendorf). Reconstitution buffer
337 (50 mM HEPES/KOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, pH 8.0, 0.5% [w/v] Sodium
338 deoxycholate, 1% [v/v] Triton X-100) was added to a final volume of 100 µl. Samples were
339 reduced in 10 mM DTT for 30 min at 60°C followed by alkylation in 25 mM iodoacetamide

340 (Sigma-Aldrich) for 25 min at RT in the dark with orbital shaking at 800 rpm in a Thermomixer
341 (Eppendorf). The reaction was quenched by incubation with 10 mM DTT f.c. for 5 min at RT.
342 For protein binding 125 µg hydrophilic and 125 µg hydrophobic Sera-Mag Carboxylate-
343 Modified Magnetic Particles (GE Healthcare) together with 150 µl absolute ethanol (Carl Roth)
344 were added to each sample and incubated for 15 min at RT at 1000 rpm in a Thermomixer
345 (Eppendorf). Beads were washed three times with 80% LC/MS grade ethanol (Supelco, Inc.).
346 Residual ethanol was removed prior to adding 1 µg LC-grade trypsin (Serva) in 100 µl of 50
347 mM ammonium bicarbonate buffer pH 8. After incubation for 4 h at 37°C and 1000 rpm in a
348 Thermomixer (Eppendorf) peptides were collected and desalted on StageTips (Rappsilber et
349 al. 2007). The amount of 400 ng digested peptides were analyzed with a nanoElute2 HPLC
350 system (Bruker) coupled to a TimsTOF HT mass spectrometer (Bruker). Peptides were
351 separated on an Ultimate CSI 25×75 C18 UHPLC column (Ionopticks) by running an optimized
352 100 min gradient of 2 to 95% MS-grade acetonitrile with 0.1% formic acid at 250 nl/min at 50°C.
353 The mass spectrometer was operated with the manufacturer provided
354 DDA_PASEF_1.1sec_cycletime method. Mass spectrometry raw files were processed with
355 MaxQuant (version 2.4.13.0) (Cox and Mann 2008) using a custom virus database (6 entries)
356 and the *C. idella* database (C_idella_female_genemodels.v1; 32,811 entries) with standard
357 settings.

358 # RESULTS

359 ## Metagenomic screening of different fish cell lines reveals presence of WhSBV

360 During metagenomic sequencing of 48 fish cell lines, the cypriniform cell lines GSB and CiSB
361 were found to contain WhSBV sequences (Supplementary Figure S1). Complete viral genomes
362 were assembled from both cell lines, sharing 99.9% nt identity with each other and 99.9% and
363 87.8% nt identity with WhSBV variants CIK (BK063520) and DSYS4497 (NC_055169),
364 respectively. The initial assembly identified a WhSBV genome sequence of 8,985 nt in length.
365 However, 3' and 5' RACE extended the initial genome to 9,008 nt (Supplementary Figure S2).
366 Compared to the current WhSBV reference sequence in Genbank (RefSeq NC_055169), the
367 complete WhSBV genome sequence presented here is 13 and 10 nt longer at the 3' and 5'
368 ends, respectively.

369 Using the complete WhSBV genome sequences, gene-specific RT-qPCR assays targeting the
370 N, P, G, M, and L gene regions were established. Among these, the G gene-specific RT-qPCR
371 assay was selected for screening all the 48 fish cell lines, as it showed best sensitivity. The
372 assay yielded positive results for GSB and CiSB cell lines with Cq values of 18.9 and 19.4,
373 respectively, while all other tested cell lines tested negative for the WhSBV G gene (see
374 Supplementary Table S1). In GSB and CiSB, N, P, M, and L gene assays were also positive by
375 RT-qPCR (see Supplementary Figure S3), while no signals were detected in the negative
376 control cell line GK. Furthermore, qPCR assays targeting the WhSBV N, P, G, M and L gene
377 regions without prior reverse transcription were negative for GSB and CiSB cells, indicating
378 that the detected WhSBV sequences are unlikely to originate from EVEs integrated into the
379 host genome, but rather represent viral RNA (see Supplementary Figure S3).

380 ## WhSBV replicates in fish and reptilian but not in avian and mammalian cells

381 In order to test the infectivity of the identified WhSBV and its potential *in vitro* host range, we
382 used cell lysate from persistently WhSBV-infected GSB cells to inoculate various cell lines from
383 bony fish, reptiles, birds and mammals (see Supplementary Table S3). Following inoculation,
384 the cells were propagated for several passages, and both cell pellets and supernatants were
385 tested for presence of viral RNA using the established G gene-specific RT-qPCR assay. An RT-
386 qPCR assay targeting the cellular ACTB gene was used as a baseline and extraction control.

387 WhSBV RNA was detected in the supernatant and cell pellet of all fish cell lines tested during
388 the first one to two passages. In cells from the order Cypriniformes, including cyprinid (CaPi,
389 CCB and GTS-9), leuciscid (EPC) and danionid (ZF4) species, viral RNA was retained over all
390 six consecutive passages. The amount of viral RNA detected started to increase compared to
391 freshly inoculated cells (T1) and eventually reached a plateau at around Cq 25 in both
392 supernatant and cell pellet. This stabilisation occurred between passages 2 and 4 for the CaPi,
393 EPC, CCB and GTS-9 cell lines. For zebrafish cell line ZF4, the amount of viral RNA initially
394 decreased after inoculation and passaging until passage 2-3, but then recovered, reaching Cq
395 values of about 28 in supernatant and cell pellets (**Figure 1A**).

396 In contrast, the non-cypriniform fish, centrarchid (RT/F), and sparid (SAF-1) showed no
397 evidence of WhSBV replication after inoculation (**Figure 1B**). Immediately after inoculation
398 (T1), viral RNA levels in the supernatant were similar to those of the cypriniform cell lines
399 (**Figure 1A**), but dropped below the detection limit in cell pellets and supernatant at passage
400 2-3 (**Figure 1B**). The chichlid (TiB) and salmonid (RTG-2/f and CHSE-214) cell lines showed a
401 decrease in cell growth, which was also reflected in the measured ACTB RNA levels, leading
402 to a complete loss of the salmonid cultures after passage 4. Viral RNA was still detectable in
403 the cell pellets until the loss of the cultures or the end of the experiment, but without a
404 detectable increase in levels (**Figure 1B**).

405 In the reptilian cell lines SKL-R, VH-2, and CDSK, viral RNA levels in the supernatant decreased
406 after inoculation but stabilised at around Cq 32-38 and remained detectable up to passage 6
407 (**Figure 2**). In the SKL-R cell pellet, WhSBV RNA was detectable at similarly low levels up to
408 passage 6. In the CDSK and VH-2 cells, viral RNA levels increased during passages 5 and 6,
409 reaching Cq values of 29 and 30, respectively.

410 In the mammalian Vero and C6 cell lines and the avian QM7 and DF-1 cell lines, WhSBV RNA
411 levels in the cell pellets or supernatants dropped below or near the detection limit already after
412 passage 1 or 2, respectively (**Figure 2**).

413 All negative controls used in this experiment showed no Cq values for the viral glycoprotein.

414 No apparent cytopathic effect was observed in any of the cell cultures used throughout the
415 experiment.

416 Using transmission electron microscopy, we detected virion-like structures in WhSBV-
417 inoculated ZF4 cell lines, whereas similar structures were absent in non-inoculated controls.
418 The putative WhSBV virions are spherical to pleomorphic with a diameter of approximately

419 150-180 nm (**Figure 3A**). Virion budding was observed rarely at the plasma membrane of the
420 persistently WhSBV-infected GSB cells (**Figure 3B**).

421 **Progressive increase of viral RNA in cypriniform cell lines over time**

422 In order to gain deeper insights into viral propagation in different fish cell lines, a kinetic
423 experiment was performed over 10 days using both cypriniform (EPC and CCB) and non-
424 cypriniform fish cells (RTG-2/f and TiB). The level of viral RNA load was assessed by RT-qPCR
425 and the presence of viral RNA at the cellular level was assessed by RNA ISH at different time
426 points.

427 In the cypriniform EPC and CCB cell lines, the virus exhibited almost identical exponential
428 growth curves, with viral RNA levels increasing from 12 hpi and reaching Cq values of 18.6 for
429 EPC and 19.5 for CCB. In contrast, the salmonid RTG-2/f and cichlid TiB cell lines showed
430 stagnation or a gradual decrease in detectable viral RNA levels (**Figure 4A**).

431 The presence of viral RNA in the cells was detected by specific RNA ISH, which detects viral
432 (-)RNA in both cypriniform (EPC and CCB) and non-cypriniform fish cell lines (RTG-2/f and TiB)
433 from 4 hpi. In general, RNA ISH confirmed the results of the kinetic experiment based on the
434 RT-qPCR assay. In particular, initial visual inspection of the cypriniform cell lines revealed that
435 the number of infected cells was higher at later time points. At early time points (4 - 72 hpi),
436 infected single cells appeared, followed by foci of positive cells, later confluent and finally
437 diffuse labelling of the cell pellet (**Figure 4B**). Percentage quantification of individual cell pellets
438 for EPC showed widespread infection, with 95% of the fields analysed showing positive
439 detection as early as 4 hpi. Infection remained widespread (63-100% of infected test fields)
440 throughout the course of the experiment until 240 hpi (see Supplementary Figure S4). In CCB
441 cells, only 29% of the test fields were infected at 4 hpi and longer cultivation correlated with a
442 higher proportion of positive test fields, eventually reaching 100% at 96, 120 and 240 hpi (see
443 Supplementary Figure S4). Similar observations were made using a custom-designed probe
444 against WhSBV (+)RNA on CCB cell lines (Supplementary Figure S5). In terms of labelling
445 signal, infected cells harvested at early time points were more likely to show fine-granular
446 cytoplasmic labelling, and larger, globular signals were consistently found at 48 hpi and later
447 (**Figure 4B**). In clear contrast, the infection pattern in the non-cypriniform fish cell lines were
448 essentially confined to single cells and showed no evidence of increasing infection abundance
449 over time (**Figure 4B**). Only the intensity of the chromogen signal within the cell increased in
450 TiB from 72 hpi, but in RTG-2/f this phenomenon was limited to 72 and 96 hpi. RTG-2/f and TiB

451 cells ultimately yielded 28% and 16% infected test fields at 240 hpi, respectively (see
452 Supplementary Figure S4).

453 **Cellular response to early and late WhSBV infection is limited in cypriniform cells**

454 To investigate cellular transcriptional changes associated with WhSBV infection, we performed
455 RNA sequencing (RNAseq) of RNA transcripts at early and late time points following viral
456 inoculation in CCB cells (**Figure 5A**). Quantitative RT-qPCR and sequencing confirmed the
457 presence of WhSBV RNA and the onset of infection at 4 hpi, with viral RNA levels progressively
458 increasing at all subsequent time points (**Figure 5B**). This trend was consistent with the RNA
459 ISH data, which showed a gradual increase in the percentage of infected CCB cells over time,
460 reaching 100% infection at 96, 120, and 240 hpi (**Figure 5B**).

461 The number of DEGs remained minimal at early and late time points but became more
462 pronounced at 240 hpi, indicating a delayed transcriptional response to infection (**Figure 5C**).
463 PCA of early time points revealed clear transcriptional differences between WhSBV-infected
464 and control cells at 4 and 8 hpi, which diminished at 24 and 48 hpi, with both groups becoming
465 more similar (**Figure 5C**). In contrast, the PCA of late time points experiment showed no
466 significant transcriptional changes at 72 and 96 hpi, However, minimal transcriptional
467 alterations were observed at 120 hpi, followed by distinct clustering of infected and control
468 samples at 240 hpi, suggesting a change in the host transcriptional response at later stages of
469 infection (**Figure 5D**).

470 Differential gene expression analysis of early time points identified three DEGs at 4 or 8 hpi.
471 Specifically, RASAL2 (RAS GTPase-activating protein 2-like [LOC109087896]) and IRF1B
472 (interferon regulatory factor 1) were upregulated at 4 hpi and 8 hpi (**Figure 5E**), whereas
473 EDEM3 (ER degradation enhancing alpha-mannosidase-like protein 3 [LOC109095466]) was
474 upregulated exclusively at 8 hpi. Conversely, ABCC1 (multidrug resistance-associated protein
475 1-like [LOC109090229]) was downregulated at 4 hpi. The expression of these genes returned
476 to baseline levels by 24 and 48 hpi. At later stages of infection (72 and 96 hpi), no DEGs were
477 detected between infected and control cells. At 120 hpi, eight DEGs were upregulated,
478 including IRF7 (interferon regulatory factor 7) and MxB (interferon-induced GTP-binding
479 protein MxB-like) (**Figure 5E**). At 240 hpi, 113 DEGs were upregulated, including IRF7, MxB,
480 and several other key antiviral genes involved in early innate immune responses in fish, such
481 as IRF3 (interferon regulatory factor 3), ISG15 (interferon-stimulated gene 15), and RSDA2
482 (viperin, radical S-adenosyl methionine domain containing 2), along with other genes involved

483 in interferon pathways, RIG-I (antiviral innate immune response receptor RIG-I
484 [LOC109081892]) and GVINP1 (interferon-induced very large GTPase 1-like [LOC109058642])
485 (**Figure 5E** and Supplementary Table S4 for complete list of DEGs).

486 **Detection of multicistronic and spliced WhSBV gene transcripts**

487 To investigate the transcriptional profile and splice sites of the WhSBV RNA, we mapped raw
488 sequence data from the persistently WhSBV-infected GSB and CiSB cells to the de novo
489 assembled WhSBV genome (**Figure 6A**). The observed sequence coverage across the
490 genomes was uneven, with abrupt increases and decreases observed in certain potential
491 intergenic regions. These coverage variations corresponded to predicted transcription start
492 and stop motifs. Three transcriptional start sites (S1-3) were identified upstream of the N, X/P
493 and M ORFs. Start sites were marked by a significant increase in read coverage, indicating
494 transcription initiation. Four transcription termination sites (T1-T4) were detected, T1, T2 and
495 T4 were located downstream of the N, G and L ORFs, respectively, while T3 was found within
496 the L ORF. The termination sites were characterised by a decrease in read coverage and the
497 presence of reads transitioning to poly(A) tails, consistent with transcriptional termination
498 (**Figure 6A**). S2 and S3 are observed immediately adjacent to T1 and T2, respectively. The
499 predicted start and termination motifs showed a conserved sequence pattern (see
500 Supplementary Table S5). As the 5' RACE analysis was performed on total RNA including
501 mRNA, Sanger sequencing of some 5' RACE products identified the transcription start site S1
502 at nucleotide position 33 to 35 in the viral genome, marked by the sequence “GAA”, in
503 agreement with the prediction results.

504 Regions of the genome with comparable levels of sequence coverage between adjacent start
505 and termination sites were interpreted as belonging to the same viral RNA transcript or mRNA.
506 Briefly, the N gene appeared to be expressed from a monocistronic mRNA, while X/P and G
507 genes were expressed from polycistronic mRNA. The mRNAs containing the M and L genes
508 shared a common start site (S3); however, their transcript levels differed with the M gene being
509 more highly expressed than the L gene (**Figure 6A**).

510 In addition, mapping revealed the presence of three potential introns (In1-3) located at the
511 beginning of the L ORF. Conventional RT-PCR and Sanger sequencing confirmed the presence
512 of In1-3 within the L ORF and also revealed an additional, unpredicted splice site, In4, within
513 the L ORF (**Figure 6B**).

514 Based on the above findings, we predicted several potential transcripts for the M and L genes,
515 which share the same start site S3 and terminate at either T3 or T4. The M gene could produce
516 a short transcript encoding the M ORF, which can either remain unspliced (M t1; size: 741 nt)
517 or undergo splicing at In1 to produce a shorter transcript (M t2; 565 nt), both terminating at T3.
518 However, the unspliced transcript appears to be in the minority of transcripts or non-existent,
519 as mapping suggests truncation at the In1 region. In addition, four transcripts were predicted,
520 termed M-L t1-4, all sharing the same start site S3 and termination site T4, but differing in their
521 splicing patterns. The longest unspliced transcript, M-L t1, contains the full-length M and L
522 ORFs (5,680 nt), whereas the other transcripts (M-L t2-4) are progressively truncated by
523 splicing at different introns, resulting in transcripts of size 5,218 nt, 5,150 nt, and 4,945 nt,
524 respectively (**Figure 6C**). These transcripts may encode either truncated versions of the L ORF
525 or only the M ORF.

526 A complementary northern blot analysis of RNA from the persistently WhSBV-infected GSB
527 and the uninfected and experimentally WhSBV-infected EPC cells, supported these findings
528 (**Figure 7A-C**). All probes (N, X, P, G, M, or L ORF specific) detected the full genome of WhSBV
529 from the GSB and EPC cells at the expected size of 9,008 nt. X, P and G are encoded by the
530 same RNA, as all three probes detected the same RNA species at approximately the predicted
531 size of 2,118 nt. This is consistent with their expression from polycistronic mRNA. The M and
532 L probes both detected a band between 6 and 8 kb, likely representing the M-L t1 transcript.
533 A faint band at 500 nt was detected with the M probe, but not with the L probe, which may
534 represent the short transcripts of the M ORF. The N probe identified an RNA species that was
535 consistently higher than expected and was also recognised by the X, P and G probes. This
536 observation suggests the presence of a long RNA spanning all four ORFs, rather than the
537 expected monocistronic N ORF mRNA.

538 High-resolution mass spectrometry was performed on the GSB cell suspension persistently
539 infected with WhSBV to provide an overview of the viral proteins. The analysis revealed
540 peptides that matched all the predicted proteins of WhSBV, except for the L ORF where no
541 matching peptides were detected. No fused peptides were detected either, suggesting no
542 additional splicing (**Figure 8**).

543 # DISCUSSION

544 This study represents the first *in vitro* molecular characterisation of Wǔhàn sharpbelly
545 bornavirus (WhSBV), a member of the family *Bornaviridae*, genus *Cultervirus* (Rubbenstroth et
546 al. 2021; Shi et al. 2018), and thus the first biological characterisation of a bornavirus not
547 belonging to the well-studied genus *Orthobornavirus*. Initially, identified by RNA sequencing of
548 sharpbelly tissues (gut, liver and gill) (Eshak et al. 2023; Shi et al. 2018), WhSBV was later
549 detected by data mining in grass carp cell lines (CIK and L8824) (Eshak et al. 2023; Shi et al.
550 2018; Costa et al. 2021). Here, we further identified WhSBV in two grass carp swim bladder-
551 derived cell lines by metagenomic RNA sequencing. Both cell lines were originally established
552 by the Pearl River Fisheries Research Institute (PRFRI), Guangzhou, China. Despite their
553 extensive use in grass carp reovirus research (Liang et al. 2014; Yang et al. 2021; Ma et al.
554 2011), these cell lines had not previously been associated with WhSBV. Whether the cells used
555 to establish the cell lines were harvested from an WhSBV-infected animal or were
556 contaminated during handling from an unknown source remains unknown. While
557 contamination can't be completely ruled out, the fact that WhSBV replicates and persistently
558 infects cypriniform cell lines suggests that the virus is more likely to have originated from the
559 grass carp used for the cell culture. This suggests that WhSBV may be endemic in cypriniform
560 fish, such as grass carp and sharpbelly fish.

561 The identification of WhSBV in well-established cell lines allowed us to study the virus *in vitro*,
562 offering insights into WhSBV persistence, potential *in vitro* host range, and transcriptional
563 strategies. By comparing its ability to replicate in cypriniform cell lines, we demonstrated that
564 WhSBV establishes a persistent infection without inducing cytopathic effects and maintains
565 stable viral RNA levels over multiple passages (**Figure 1A**). Similar behaviour has been
566 observed for BoDV-1, BoDV-2, VSBV-1 and several avian orthobornaviruses, where some
567 inoculated cell lines became persistently infected without cytopathic effect and could be sub-
568 cultivated without loss of infectivity (Herzog and Rott 1980; Rinder et al. 2009; Pham et al. 2020;
569 Rubbenstroth et al. 2012; Rubbenstroth et al. 2014; Rubbenstroth et al. 2013; Schlottau et al.
570 2020; Nowotny et al. 2000). The efficient replication in cypriniform cells and the ability to
571 establish persistent infections may indicate that fish of this order are the preferred host of
572 WhSBV in natural environments. However, the natural host range of WhSBV and other
573 culterviruses remains unknown.

574 To assess the infectivity of WhSBV in different cell lines, we inoculated cell lines from different
575 taxonomic groups, including non-cypriniform fish, reptiles, birds and mammals. In some non-
576 cypriniform fish cell lines, susceptibility to WhSBV infection was difficult to assess because a
577 decline in ACTB levels over several passages, with eventual loss of cultures, suggested
578 suboptimal growth conditions, possibly due to excessive passaging or the use of non-vented
579 flasks. In others, however, viral RNA was lost after a few passages despite stable ACTB levels,
580 suggesting that they did not facilitate persistent WhSBV infection (**Figure 1B**). Despite these
581 limitations, RNA ISH confirmed the presence of viral RNA within cells for RTG-2/f and TiB,
582 indicating that these cells were indeed infected by WhSBV. However, infection remained
583 confined to individual cells and there was no progressive increase in the percentage of infected
584 cells, unlike in cypriniform cells where infection spread over time. Surprisingly, WhSBV was
585 able to infect, replicate and establish persistence in all reptilian cell lines tested, albeit with
586 lower viral RNA levels compared to the positive control, cypriniform GTS-9 fish cells (**Figure 2**).
587 However, similar declines in ACTB levels were observed as in non-cypriniform fish. No
588 sustained replication of WhSBV was detected in either avian or mammalian cell lines, as viral
589 RNA levels declined rapidly and were undetectable after several passages (**Figure 2**). This
590 may be due either to the absence of essential host factors in avian or mammalian cells required
591 for WhSBV replication, or to an effective antiviral response.

592 The limited ability of WhSBV to infect and replicate in non-cypriniform fish and reptilian cells
593 may indicate an evolutionary history of adaptation to these hosts, and therefore the potential
594 existence of other related culterviruses with a greater affinity for these hosts. An *in silico*
595 screening identified five novel fish bornavirids in bony and cartilaginous fish hosts that showed
596 significant genomic divergence, suggesting that the diversity of culterviruses is currently
597 underestimated (Eshak et al. 2023). Future studies will be essential to determine host range of
598 WhSBV and other *Cultervirus* members in natural settings.

599 In order to study how WhSBV is able to establish a persistent infection without inducing
600 cytopathic effects, we analysed the host gene expression at different time points of infection.
601 Transcriptomic analysis of the experimentally WhSBV-infected CCB cells indicated a limited
602 initial activation of the innate immune response shortly after inoculation at 4 and 8 hpi by
603 upregulation of the genes IRF1B and RASAL2 (**Figure 5E**). Interferon regulatory factors, such
604 as IRF1B, are key regulators of the innate antiviral response in vertebrates, controlling the
605 expression of interferons (IFNs) and interferon-stimulated genes (ISGs) (Nehyba et al. 2009;
606 Negishi et al. 2018). In zebrafish, IRF1B alone was shown to induce the expression of interferon
607 and ISGs, resulting in effective protection against viral infection (Feng et al. 2015). While the

608 role of RASAL2 in fish is currently unknown, the gene may function as tumour and metastasis
609 suppressor in humans (McLaughlin et al. 2013). Despite the initial expression of IRF1B, no
610 significant host cell transcriptional changes were observed between 24 and 96 hpi (**Figure**
611 **5C**), possibly allowing for efficient viral replication and spread (**Figure 5B**). The lack of
612 significant host cell transcriptional changes between 24 and 96 hpi is in stark contrast to the
613 observed rising levels of viral RNA and high percentage of infected cells.

614 In comparison, orthobornaviruses have been shown to evade IFN induction by possessing
615 monophosphorylated genome ends (Habjan et al. 2008; Reuter et al. 2010). BoDV-1 further
616 suppresses the interferon response via P-protein-mediated inhibition of TANK-binding kinase
617 1 (TBK-1) (Unterstab et al. 2005) and N-protein interference with IRF7 activation (Song et al.
618 2013).

619 Only at 72 hpi we observed the activation of IRFs other than IRF1B, that resulted in a delayed
620 innate immune response at 120 and 240 hpi, characterised by the significant upregulation of
621 many ISGs, such as MxB, GVINP1, ISG15, and RSDA2, which play a role in the innate antiviral
622 response in fish (Del Ortega-Villaizan et al. 2022). This suggests that the host's innate immune
623 system did not mount a response until the infection had spread to all cells. However, this
624 response was ultimately ineffective, as the cells remained persistently infected, possibly
625 indicating the presence of a secondary regulatory mechanism controlling viral load. Despite
626 the upregulation of ISGs, no significant induction of apoptosis- or necrosis-related genes was
627 observed throughout infection, suggesting that WhSBV establishes persistence without
628 inducing programmed cell death. This allows for continued viral replication and long-term
629 infection. In contrast, common carp hepatocytes exposed to ammonia stress undergo
630 apoptosis via the P53-BAC/BCL-2 pathway, demonstrating the species' ability to induce cell
631 death under certain conditions (Xue et al. 2021). The absence of such responses in WhSBV-
632 infected cells raised the possibility that the virus actively suppresses apoptosis, further
633 contributing to its persistence. In fish, there is ongoing debate regarding the classification of
634 their IFNs as type I, II, or III due to the presence of introns in fish IFN genes. However,
635 phylogenetic and structural analyses suggest that fish IFN genes primarily encode type I IFNs
636 (Qi et al. 2010; Hamming et al. 2011; Chang et al. 2013). Type I IFNs are known to trigger a
637 cascade of immune responses upon viral infection, leading to the induction of antiviral genes
638 (Huang et al. 2015; Chang et al. 2013). Interestingly, despite the induction of ISGs, type I IFNs
639 (e.g. IFNA3) were not among the upregulated DEGs (**Figure 5E**). This suggests the involvement
640 of alternative signalling pathways in common carp or very subtle IFN expression. Additionally,
641 certain viral pathogens have evolved mechanism to suppress host interferon signalling, which

642 could explain the absence of type I IFN upregulation (Swaraj and Tripathi 2024). Further studies
643 are needed to elucidate the molecular mechanisms underlying WhSBV-mediated immune
644 evasion and/or modulation and to identify the specific viral genes, proteins or genomic
645 structures involved in these processes. The establishment of persistence without cytopathic
646 effects is a common feature of orthobornaviruses (Rubbenstroth et al. 2021) and has also been
647 observed in other mononegaviruses, such as arenavirids (family *Arenaviridae*), which suppress
648 host antiviral responses, particularly by inhibiting type I interferon production, to maintain
649 chronic infections (Meyer and Ly 2016; Kane et al. 2016).

650 In order to regulate its own gene expression, WhSBV shares conserved transcriptional features
651 with other *Cultervirus* genomes, notably the presence of an intron at the boundary between
652 the M and L genes, suggesting a common transcriptional mechanism. Additionally,
653 culterviruses have a conserved nucleotide sequence pattern for termination sites (T1-T4),
654 "AYUUWAKAAAAACAU", while start sites (S1-3) follow the motif "GAM" (Eshak et al. 2023).
655 However, experimental validation of these transcriptional elements is still required. In contrast
656 to the predominantly monocistronic transcription seen in other mononegaviruses
657 (Amarasinghe et al. 2019), such as paramyxovirids (family *Paramyxoviridae*), where the viral
658 genome is transcribed into monocistronic mRNAs, each encoding a single protein (Wignall-
659 Fleming et al. 2019). Orthobornaviruses use polycistronic mRNAs with extensive splicing of the
660 overlapping M, G, and L ORFs to regulate their gene expression (Cubitt et al. 1994; Cubitt et
661 al. 2001; Schneider et al. 1994; Tomonaga et al. 2000), along with multiple transcriptional start
662 and termination sites, as observed in BoDV-1 (Tomonaga et al. 2002; Schneider et al. 1994).
663 In WhSBV, the M and G ORFs have been repositioned during evolution and splicing events
664 occur at the M-L boundary. Understanding the differences in splicing patterns may provide
665 insight into the evolutionary forces shaping genome organisation and how each genus
666 regulates transcription in different host species. There is currently no data is available on the
667 splicing patterns of carboviruses, which share the same genome organisation as culterviruses.
668 Determining whether they exhibit similar or distinct transcriptional mechanisms is essential to
669 understanding their different regulatory strategies (Hyndman et al. 2018). Notably, bornavirids
670 are not the only mononegaviruses that use polycistronic transcription; for example, Ebola virus
671 (EBOV; family *Filoviridae*) uses RNA editing of its glycoprotein (GP) gene to generate multiple
672 GP variants, optimising replication and pathogenicity without increasing genome size (Yu et al.
673 2017).

674 Northern blot analysis suggested that the N proteins can be expressed not only from a
675 monocistronic mRNA, but also co-expressed with the X/P and G from a polycistronic mRNA.

676 However, it remains to be clarified whether the detected RNA of approximately 3000-4000 nt
677 in length is actually mRNA or a subgenomic RNA. For BoDV-1, the non-coding region between
678 the N and X genes shows structural similarities to that of WhSBV, where T1 is located
679 downstream of the transcription start site S2 (Poenisch et al. 2008). The consensus motif of
680 the BoDV-1 termination sites, $U(A)_{6-7}$ (Schneemann et al. 1995), serves as a core signal
681 sequence but is insufficient for efficient transcription termination on its own, as the surrounding
682 upstream and downstream nucleotides significantly influence T1 utilization. Consequently,
683 when transcription is initiated at S1 and terminated at T1, these regulatory elements contribute
684 to the production of a monocistronic (1,200 nt) mRNA encoding the N protein (Poenisch et al.
685 2008). However, studies have demonstrated that the viral transcriptase frequently bypasses
686 the T1 transcription termination site, resulting in the production of a longer (1,900 nt)
687 polycistronic mRNA containing the N, X, and P ORFs (Poenisch et al. 2008). The presence of
688 such transcripts in BoDV-1-infected cell cultures and rat brains suggests that they are bona
689 fide mRNAs rather than aberrant replication intermediates, and their production may be related
690 to viral fitness (Poenisch et al. 2008). Conversely, other studies have suggested that these long
691 mRNAs may represent aborted replication products or an extended form of “leader RNA”,
692 similar to those described in certain other negative-strand RNA viruses (Schneemann et al.
693 1994; Banerjee et al. 1991). In addition, although L transcripts were detected by Northern blot,
694 proteomic analysis did not identify L protein, suggesting low levels of expression, possibly
695 regulated by splicing events that modulate its protein level, similar to BoDV-1 (Schneider et al.
696 1994). A similar mechanism has been observed for vesicular stomatitis virus (VSV; family
697 *Rhabdoviridae*), which controls L gene expression by polycistronic transcription, co-expressing
698 it with other viral genes to balance protein levels and minimise host cell stress while optimising
699 viral replication and assembly (Barr et al. 2002). This specific transcriptional regulation of
700 WhSBV may balance viral protein levels but requires further investigation.

701 CONCLUSION

702 This study provides a molecular characterisation of WhSBV, suggesting an adaptation to
703 predominantly cypriniform hosts and its ability to establish a persistent infection without
704 inducing detectable cytopathic effects. The efficient replication of the virus in cypriniform cell
705 cultures and the suppressed induction of an early host innate immune response underline its
706 evolutionary co-adaptation to these fish hosts. Our findings highlight the importance of moving
707 from *in silico* screening to *in vitro* validation, a critical step in virological research. While
708 metagenomic approaches facilitate virus discovery, the lack of direct sequence validation from

real samples often remains a key limitation. Furthermore, the ability of WhSBV to establish persistent infections in cypriniform cells, while showing limited replication in non-cypriniform cultures, highlights the need for *in vivo* experimental validation in defining viral host ranges. In addition, this study provides valuable insights into the transcriptional mechanisms of WhSBV, highlighting its regulatory strategies to modulate gene expression. Further research is essential to unravel the molecular mechanisms underlying the persistence of WhSBV and to assess its ecological impact, particularly in aquaculture where persistent viral infections could have significant implications for fish health and disease management.

DATA AVAILABILITY

All sequencing data associated with this study have been deposited in public repositories and are freely accessible. The genome sequence is available in GenBank under the accession number PV171101. Additional metadata can be found under BioProject PRJNA1226401 and BioSample SAMN46928747. Raw sequencing reads of WhSBV persistently infected GSB are available in the Sequence Read Archive (SRA) under SRR32424572, and RNAseq data have been deposited at ArrayExpress under E-MTAB-14894.

ACKNOWLEGEMENTS

The investigations were supported by a Friedrich-Loeffler-Institut internal PhD program, grant number FLI-IVD-XX-2021-83, granted to F.P.

We gratefully acknowledge the assistance of Fermin Georgio Lorenzen-Schmidt, Mathias Lenk from the Collection of Cell Lines in Veterinary Medicine (CCLV), and the contributions of Robin Brandt, and Kristin Vorpahl.

CONFLICT OF INTEREST

The authors declare no conflicts of interest. The funders had no role in study design, data collection, analysis, interpretation, manuscript writing, or decision to publish the findings.

REFERENCES

Amarasinghe, Gaya K.; Ayllón, María A.; Bào, Yīmíng; Basler, Christopher F.; Bavari, Sina; Blasdell, Kim R. et al. (2019): Taxonomy of the order Mononegavirales: update 2019. In *Archives of virology* 164 (7), pp. 1967–1980. DOI: 10.1007/s00705-019-04247-4.

Banerjee, A. K.; Barik, S.; De, B. P. (1991): Gene expression of nonsegmented negative strand RNA viruses. In *Pharmacology & therapeutics* 51 (1), pp. 47–70. DOI: 10.1016/0163-7258(91)90041-j.

Barr, John N.; Whelan, Sean P. J.; Wertz, Gail W. (2002): Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. In *Biochimica et biophysica acta* 1577 (2), pp. 337–353. DOI: 10.1016/S0167-4781(02)00462-1.

Briese, T.; La Torre, J. C. de; Lewis, A.; Ludwig, H.; Lipkin, W. I. (1992): Borna disease virus, a negative-strand RNA virus, transcribes in the nucleus of infected cells. In *Proceedings of the National Academy of Sciences of the United States of America* 89 (23), pp. 11486–11489. DOI: 10.1073/pnas.89.23.11486.

Briese, T.; Schneemann, A.; Lewis, A. J.; Park, Y. S.; Kim, S.; Ludwig, H.; Lipkin, W. I. (1994): Genomic organization of Borna disease virus. In *Proceedings of the National Academy of Sciences of the United States of America* 91 (10), pp. 4362–4366. DOI: 10.1073/pnas.91.10.4362.

Chang, Ming-Xian; Zou, Jun; Nie, Pin; Huang, Bei; Yu, Zhanglong; Collet, Bertrand; Secombes, Chris J. (2013): Intracellular interferons in fish: a unique means to combat viral infection. In *PLoS Pathogens* 9 (11), e1003736. DOI: 10.1371/journal.ppat.1003736.

Costa, Vincenzo A.; Mifsud, Jonathon C. O.; Gilligan, Dean; Williamson, Jane E.; Holmes, Edward C.; Geoghegan, Jemma L. (2021): Metagenomic sequencing reveals a lack of virus exchange between native and invasive freshwater fish across the Murray-Darling Basin, Australia. In *Virus evolution* 7 (1), veab034. DOI: 10.1093/ve/veab034.

759 Cox, Jürgen; Mann, Matthias (2008): MaxQuant enables high peptide identification rates,
760 individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. In
761 *Nature biotechnology* 26 (12), pp. 1367–1372. DOI: 10.1038/nbt.1511.

762 Cubitt, B.; Oldstone, C.; Valcarcel, J.; La Carlos de Torre, J. (1994): RNA splicing contributes
763 to the generation of mature mRNAs of Borna disease virus, a non-segmented negative strand
764 RNA virus. In *Virus research* 34 (1), pp. 69–79. DOI: 10.1016/0168-1702(94)90120-1.

765 Cubitt, Beatrice; Ly, Calvin; La Torre, Juan Carlos de (2001): Identification and characterization
766 of a new intron in Borna disease virus. In *The Journal of general virology* 82 (Pt 3), pp. 641–
767 646. DOI: 10.1099/0022-1317-82-3-641.

768 Del Ortega-Villaizan, Maria Mar; Chico, Veronica; Perez, Luis (2022): Fish Innate Immune
769 Response to Viral Infection-An Overview of Five Major Antiviral Genes. In *Viruses* 14 (7). DOI:
770 10.3390/v14071546.

771 Eshak, Mirette I. Y.; Rubbenstroth, Dennis; Beer, Martin; Pfaff, Florian (2023): Diving deep into
772 fish bornaviruses: Uncovering hidden diversity and transcriptional strategies through
773 comprehensive data mining. In *Virus evolution* 9 (2), vead062. DOI: 10.1093/ve/vead062.

774 Ewels, Philip A.; Peltzer, Alexander; Fillinger, Sven; Patel, Harshil; Alneberg, Johannes; Wilm,
775 Andreas et al. (2020): The nf-core framework for community-curated bioinformatics pipelines.
776 In *Nature biotechnology* 38 (3), pp. 276–278. DOI: 10.1038/s41587-020-0439-x.

777 Feng, Hui; Zhang, Yi-Bing; Zhang, Qi-Min; Li, Zhi; Zhang, Qi-Ya; Gui, Jian-Fang (2015):
778 Zebrafish IRF1 regulates IFN antiviral response through binding to IFNφ1 and IFNφ3 promoters
779 downstream of MyD88 signaling. In *Journal of immunology (Baltimore, Md. : 1950)* 194 (3),
780 pp. 1225–1238. DOI: 10.4049/jimmunol.1402415.

781 Habjan, Matthias; Andersson, Ida; Klingström, Jonas; Schümann, Michael; Martin, Arnold;
782 Zimmermann, Petra et al. (2008): Processing of genome 5' termini as a strategy of negative-
783 strand RNA viruses to avoid RIG-I-dependent interferon induction. In *PloS one* 3 (4), e2032.
784 DOI: 10.1371/journal.pone.0002032.

785 Hamming, Ole Jensen; Lutfalla, Georges; Levraud, Jean-Pierre; Hartmann, Rune (2011):
786 Crystal structure of Zebrafish interferons I and II reveals conservation of type I interferon
787 structure in vertebrates. In *Journal of virology* 85 (16), pp. 8181–8187. DOI:
788 10.1128/JVI.00521-11.

789 Herzog, S.; Rott, R. (1980): Replication of Borna disease virus in cell cultures. In *Medical*
790 *microbiology and immunology* 168 (3), pp. 153–158. DOI: 10.1007/BF02122849.

791 Hoffmann, Bernd; Tappe, Dennis; Höper, Dirk; Herden, Christiane; Boldt, Annemarie; Mawrin,
792 Christian et al. (2015): A Variegated Squirrel Bornavirus Associated with Fatal Human
793 Encephalitis. In *The New England journal of medicine* 373 (2), pp. 154–162. DOI:
794 10.1056/nejmoa1415627.

795 Huang, Zhilin; Chen, Song; Liu, Jiachen; Xiao, Jun; Yan, Jun; Feng, Hao (2015): IFNa of black
796 carp is an antiviral cytokine modified with N-linked glycosylation. In *Fish & shellfish immunology*
797 46 (2), pp. 477–485. DOI: 10.1016/j.fsi.2015.07.020.

798 Hyndman, Timothy H.; Shilton, Catherine M.; Stenglein, Mark D.; Wellehan, James F. X. (2018):
799 Divergent bornaviruses from Australian carpet pythons with neurological disease date the
800 origin of extant Bornaviridae prior to the end-Cretaceous extinction. In *PLoS Pathogens* 14 (2).
801 DOI: 10.1371/journal.ppat.1006881.

802 Kane, Melissa; Zang, Trinity M.; Rihn, Suzannah J.; Zhang, Fengwen; Kueck, Tonya; Alim,
803 Mudathir et al. (2016): Identification of Interferon-Stimulated Genes with Antiretroviral Activity.
804 In *Cell host & microbe* 20 (3), pp. 392–405. DOI: 10.1016/j.chom.2016.08.005.

805 Kuhn, Jens H.; Dürrwald, Ralf; Bào, Yīmíng; Briese, Thomas; Carbone, Kathryn; Clawson, Anna
806 N. et al. (2015): Taxonomic reorganization of the family Bornaviridae. In *Archives of virology*
807 160 (2), pp. 621–632. DOI: 10.1007/s00705-014-2276-z.

808 Liang, Hong-Ru; Li, Yong-Gang; Zeng, Wei-Wei; Wang, Ying-Ying; Wang, Qing; Wu, Shu-Qin
809 (2014): Pathogenicity and tissue distribution of grass carp reovirus after intraperitoneal
810 administration. In *Virology journal* 11, p. 178. DOI: 10.1186/1743-422X-11-178.

811 Ma, Jie; Wang, Weiming; Zeng, Lingbing; Fan, Yuding; Xu, Jin; Zhou, Yong (2011): Inhibition of
812 the replication of grass carp reovirus in CIK cells with plasmid-transcribed shRNAs. In *Journal*
813 *of virological methods* 175 (2), pp. 182–187. DOI: 10.1016/j.jviromet.2011.05.008.

814 McLaughlin, Sara Koenig; Olsen, Sarah Naomi; Dake, Benjamin; Raedt, Thomas de; Lim,
815 Elgene; Bronson, Roderick Terry et al. (2013): The RasGAP gene, RASAL2, is a tumor and
816 metastasis suppressor. In *Cancer cell* 24 (3), pp. 365–378. DOI: 10.1016/j.ccr.2013.08.004.

817 Meyer, Bjoern; Ly, Hinh (2016): Inhibition of Innate Immune Responses Is Key to Pathogenesis
818 by Arenaviruses. In *Journal of virology* 90 (8), pp. 3810–3818. DOI: 10.1128/JVI.03049-15.

Negishi, Hideo; Taniguchi, Tadatsugu; Yanai, Hideyuki (2018): The Interferon (IFN) Class of Cytokines and the IFN Regulatory Factor (IRF) Transcription Factor Family. In *Cold Spring Harbor perspectives in biology* 10 (11). DOI: 10.1101/cshperspect.a028423.

Nehyba, Jirí; Hrdlicková, Radmila; Bose, Henry R. (2009): Dynamic evolution of immune system regulators: the history of the interferon regulatory factor family. In *Molecular biology and evolution* 26 (11), pp. 2539–2550. DOI: 10.1093/molbev/msp167.

Niller, Hans Helmut; Angstwurm, Klemens; Rubbenstroth, Dennis; Schlottau, Kore; Ebinger, Arnt; Giese, Sebastian et al. (2020): Zoonotic spillover infections with Borna disease virus 1 leading to fatal human encephalitis, 1999-2019: an epidemiological investigation. In *The Lancet. Infectious diseases* 20 (4), pp. 467–477. DOI: 10.1016/s1473-3099(19)30546-8.

Nowotny, N.; Kolodziejek, J.; Jehle, C. O.; Suchy, A.; Staeheli, P.; Schwemmle, M. (2000): Isolation and characterization of a new subtype of Borna disease virus. In *Journal of virology* 74 (12), pp. 5655–5658. DOI: 10.1128/jvi.74.12.5655-5658.2000.

Pham, Phuc H.; Leacy, Alexander; Deng, Li; Nagy, Éva; Susta, Leonardo (2020): Isolation of Ontario aquatic bird bornavirus 1 and characterization of its replication in immortalized avian cell lines. In *Virology journal* 17 (1), p. 16. DOI: 10.1186/s12985-020-1286-6.

Poenisch, Marion; Wille, Sandra; Staeheli, Peter; Schneider, Urs (2008): Polymerase read-through at the first transcription termination site contributes to regulation of borna disease virus gene expression. In *Journal of virology* 82 (19), pp. 9537–9545. DOI: 10.1128/jvi.00639-08.

Qi, Zhitao; Nie, Pin; Secombes, Chris J.; Zou, Jun (2010): Intron-containing type I and type III IFN coexist in amphibians: refuting the concept that a retroposition event gave rise to type I IFNs. In *Journal of immunology (Baltimore, Md. : 1950)* 184 (9), pp. 5038–5046. DOI: 10.4049/jimmunol.0903374.

Rappsilber, Juri; Mann, Matthias; Ishihama, Yasushi (2007): Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. In *Nature protocols* 2 (8), pp. 1896–1906. DOI: 10.1038/nprot.2007.261.

Reuter, Antje; Ackermann, Andreas; Kothlow, Sonja; Rinder, Monika; Kaspers, Bernd; Staeheli, Peter (2010): Avian bornaviruses escape recognition by the innate immune system. In *Viruses* 2 (4), pp. 927–938. DOI: 10.3390/v2040927.

Rinder, Monika; Ackermann, Andreas; Kempf, Hermann; Kaspers, Bernd; Korbel, Rüdiger; Staeheli, Peter (2009): Broad tissue and cell tropism of avian bornavirus in parrots with proventricular dilatation disease. In *Journal of virology* 83 (11), pp. 5401–5407. DOI: 10.1128/JVI.00133-09.

Rubbenstroth, D.; Rinder, M.; Kaspers, B.; Staeheli, P. (2012): Efficient isolation of avian bornaviruses (ABV) from naturally infected psittacine birds and identification of a new ABV genotype from a salmon-crested cockatoo (Cacatua moluccensis). In *Veterinary microbiology* 161 (1-2), pp. 36–42. DOI: 10.1016/j.vetmic.2012.07.004.

Rubbenstroth, Dennis; Briese, Thomas; Dürrwald, Ralf; Horie 堀江真行, Masayuki; Hyndman, Timothy H.; Kuhn, Jens H. et al. (2021): ICTV Virus Taxonomy Profile: Bornaviridae. In *The Journal of general virology* 102 (7). DOI: 10.1099/jgv.0.001613.

Rubbenstroth, Dennis; Rinder, Monika; Stein, Malte; Höper, Dirk; Kaspers, Bernd; Brosinski, Katrin et al. (2013): Avian bornaviruses are widely distributed in canary birds (Serinus canaria f. domestica). In *Veterinary microbiology* 165 (3-4), pp. 287–295. DOI: 10.1016/j.vetmic.2013.03.024.

Rubbenstroth, Dennis; Schmidt, Volker; Rinder, Monika; Legler, Marko; Corman, Victor Max; Staeheli, Peter (2014): Discovery of a new avian bornavirus genotype in estrildid finches (Estrildidae) in Germany. In *Veterinary microbiology* 168 (2-4), pp. 318–323. DOI: 10.1016/j.vetmic.2013.11.032.

Schlottau, Kore; Nobach, Daniel; Herden, Christiane; Finke, Stefan; Beer, Martin; Hoffmann, Donata (2020): First isolation, in-vivo and genomic characterization of zoonotic variegated squirrel Bornavirus 1 (VSBV-1) isolates. In *Emerging microbes & infections* 9 (1), pp. 2474–2484. DOI: 10.1080/22221751.2020.1847604.

Schneemann, A.; Schneider, P. A.; Kim, S.; Lipkin, W. I. (1994): Identification of signal sequences that control transcription of borna disease virus, a nonsegmented, negative-strand RNA virus. In *Journal of virology* 68 (10), pp. 6514–6522. DOI: 10.1128/JVI.68.10.6514-6522.1994.

Schneemann, A.; Schneider, P. A.; Lamb, R. A.; Lipkin, W. I. (1995): The remarkable coding strategy of borna disease virus: a new member of the nonsegmented negative strand RNA viruses. In *Virology* 210 (1), pp. 1–8. DOI: 10.1006/viro.1995.1311.

878 Schneider, P. A.; Schneemann, A.; Lipkin, W. I. (1994): RNA splicing in Borna disease virus, a
879 nonsegmented, negative-strand RNA virus. In *Journal of virology* 68 (8), pp. 5007–5012. DOI:
880 10.1128/JVI.68.8.5007-5012.1994.

881 Shi, Mang; Lin, Xian-Dan; Chen, Xiao; Tian, Jun-Hua; Chen, Liang-Jun; Li, Kun et al. (2018):
882 The evolutionary history of vertebrate RNA viruses. In *Nature* 556 (7700), pp. 197–202. DOI:
883 10.1038/s41586-018-0012-7.

884 Song, Wuqi; Kao, Wenping; Zhai, Aixia; Qian, Jun; Li, Yujun; Zhang, Qingmeng et al. (2013):
885 Borna disease virus nucleoprotein inhibits type I interferon induction through the interferon
886 regulatory factor 7 pathway. In *Biochemical and biophysical research communications* 438 (4),
887 pp. 619–623. DOI: 10.1016/j.bbrc.2013.08.006.

888 Tomonaga, K.; Kobayashi, T.; Lee, B. J.; Watanabe, M.; Kamitani, W.; Ikuta, K. (2000):
889 Identification of alternative splicing and negative splicing activity of a nonsegmented negative-
890 strand RNA virus, Borna disease virus. In *Proceedings of the National Academy of Sciences of*
891 *the United States of America* 97 (23), pp. 12788–12793. DOI: 10.1073/pnas.97.23.12788.

892 Tomonaga, Keizo; Kobayashi, Takeshi; Ikuta, Kazuyoshi (2002): Molecular and cellular biology
893 of Borna disease virus infection. In *Microbes and infection* 4 (4), pp. 491–500. DOI:
894 10.1016/s1286-4579(02)01564-2.

895 Unterstab, Gunhild; Ludwig, Stephan; Anton, Aline; Planz, Oliver; Dauber, Bianca; Krappmann,
896 Daniel et al. (2005): Viral targeting of the interferon-{beta}-inducing Traf family member-
897 associated NF-{kappa}B activator (TANK)-binding kinase-1. In *Proceedings of the National*
898 *Academy of Sciences of the United States of America* 102 (38), pp. 13640–13645. DOI:
899 10.1073/pnas.0502883102.

900 Wignall-Fleming, Elizabeth B.; Hughes, David J.; Vattipally, Sreenu; Modha, Sejal; Goodbourn,
901 Steve; Davison, Andrew J.; Randall, Richard E. (2019): Analysis of Paramyxovirus Transcription
902 and Replication by High-Throughput Sequencing. In *Journal of virology* 93 (17). DOI:
903 10.1128/JVI.00571-19.

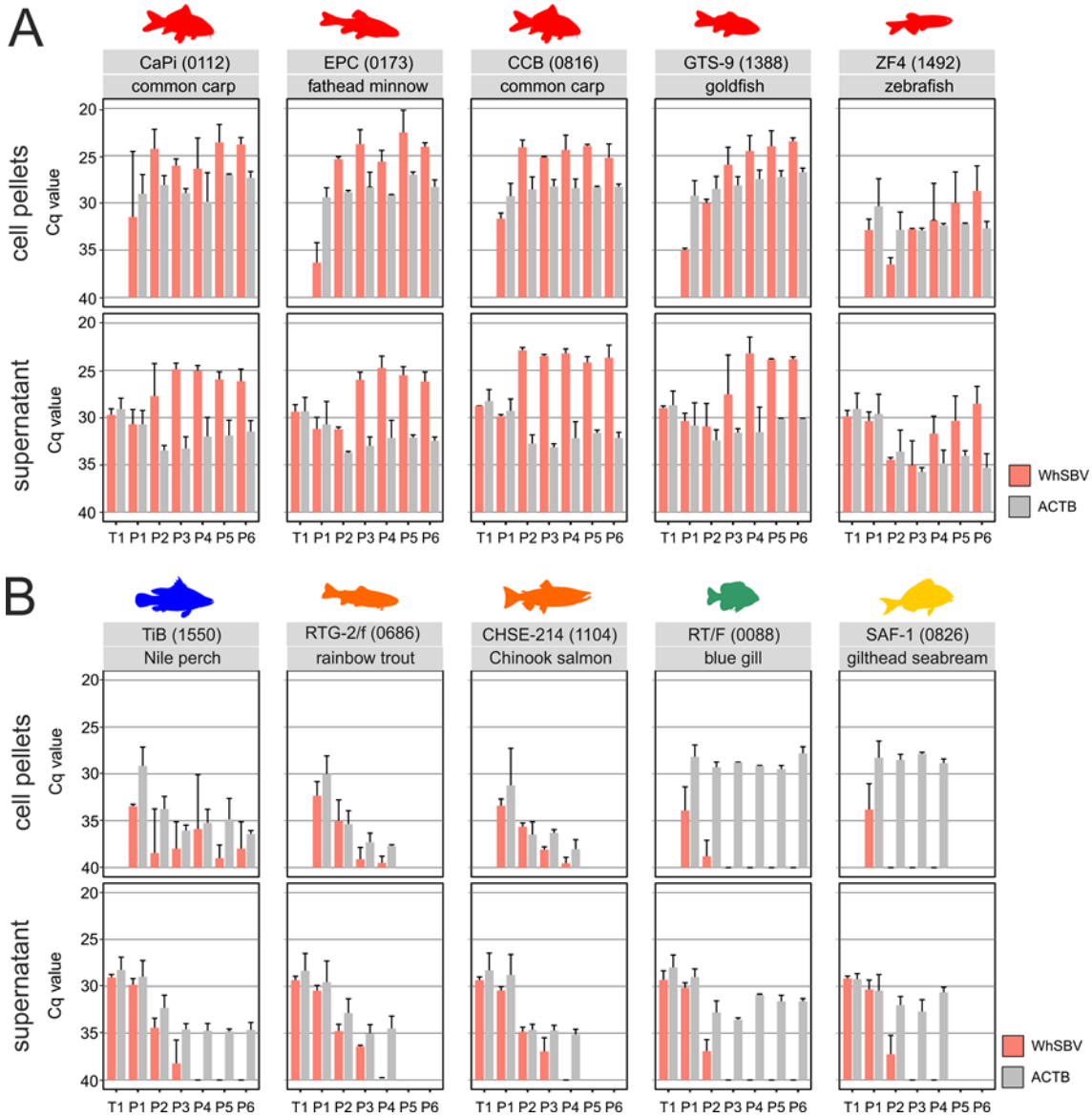
904 Xue, Shuqun; Lin, Jiawen; Han, Yue; Han, Ying (2021): Ammonia stress–induced apoptosis by
905 p53–BAX/BCL-2 signal pathway in hepatopancreas of common carp (Cyprinus carpio). In
906 *Aquacult Int* 29 (4), pp. 1895–1907. DOI: 10.1007/s10499-021-00724-3.

907 Yang, Yuru; Wang, Yingying; Wang, Qing; Zeng, Weiwei; Li, Yingying; Yin, Jiyuan et al. (2021):
908 Establishment of a cell line from swim bladder of the Grass carp (Ctenopharyngodon idellus)

for propagation of Grass Carp Reovirus Genotype II. In *Microbial pathogenesis* 151, p. 104739. DOI: 10.1016/j.micpath.2021.104739.

Yu, Dong-Shan; Weng, Tian-Hao; Wu, Xiao-Xin; Wang, Frederick X. C.; Lu, Xiang-Yun; Wu, Hai-Bo et al. (2017): The lifecycle of the Ebola virus in host cells. In *Oncotarget* 8 (33), pp. 55750–55759. DOI: 10.18632/oncotarget.18498.

914 **FIGURES**



915

916 **Figure 1: WhSBV RNA levels over successive passages in different fish cell lines.** Different
917 fish cell lines were inoculated with WhSBV. The cells were then passaged for 4-6 times by
918 splitting. (**A**) Cell lines from the taxonomic order of Cypriniformes (red); (**B**) Cell lines from the
919 taxonomic orders of Chichliformes (blue), Salmoniformes (orange), Centrarchiformes (green)
920 and Spariformes (yellow). RNA levels of viral glycoprotein (WhSBV; red bars) and cellular beta
921 actin (ACTB; grey bars) genes were detected by specific RT-qPCR in cell pellets (top row) and
922 cell culture supernatants (bottom row) from each passage. Results are presented as the
923 arithmetic mean (± standard deviation) of the cycle of quantification (Cq) values of two
924 independent experiments.

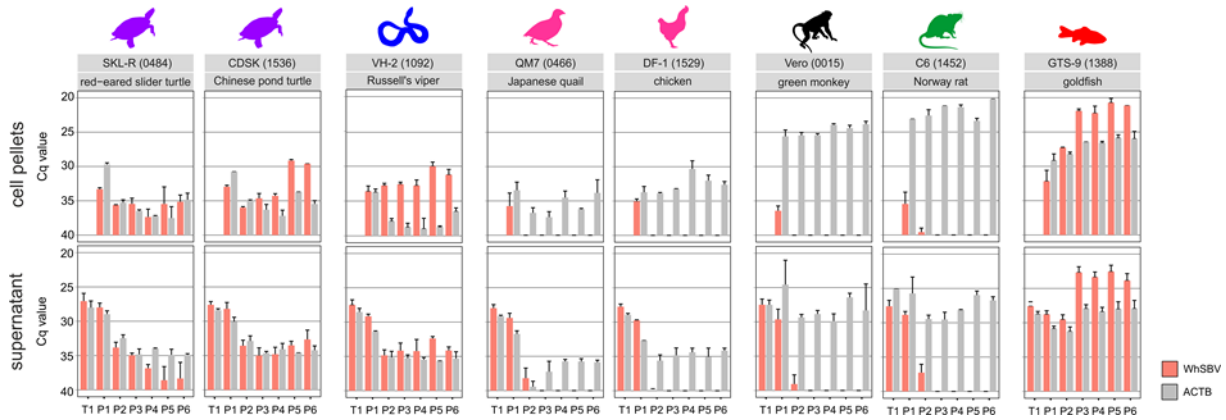


Figure 2: WhSBV RNA levels over successive passages in different non-fish cell lines. Different cell lines from the taxonomic orders of Testudines (purple), Squamata (blue), Galliformes (pink), Primates (black), and Rodentia (green) were inoculated with WhSBV and subsequently passaged for 6 times by splitting of the cells. The cypriniform fish cell line GTS-9 (red) was used as a positive control to validate the infection experiment. RNA levels of viral glycoprotein (WhSBV; red bars) and cellular beta actin (ACTB; grey bars) genes were detected by specific RT-qPCR in cell pellets (top row) and cell culture supernatants (bottom row) from each passage. Results are presented as the arithmetic mean (± standard deviation) of the cycle of quantification (Cq) values of two independent experiments.

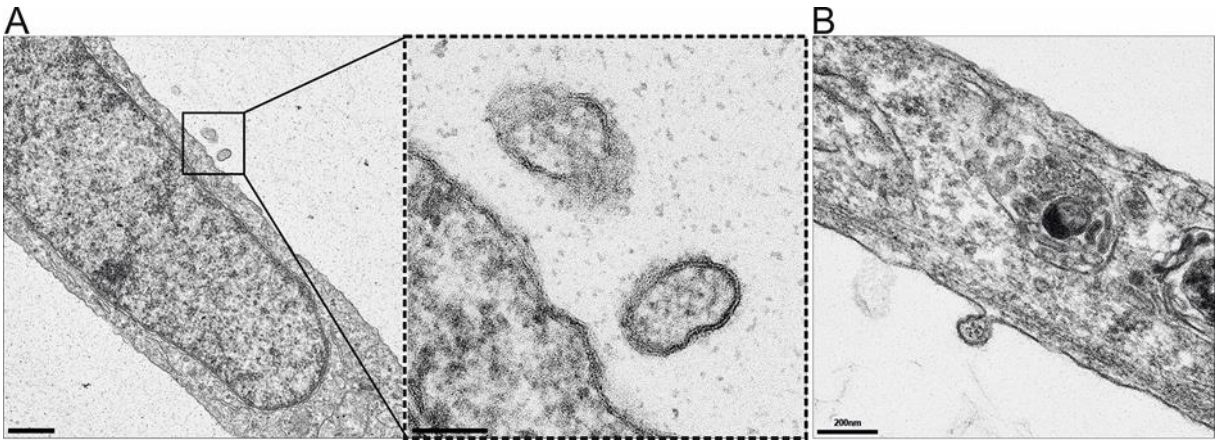


Figure 3: Detection of bornavirid-like particles by electron microscopy. (**A**) Bornavirid-like particles detected in the experimentally infected ZF4 cell line. The WhSBV virions are spherical to pleomorphic in shape, with a diameter of approximately 150-180 nm. (**B**) A potential WhSBV virion budding from the persistently WhSBV-infected GSB cell line.

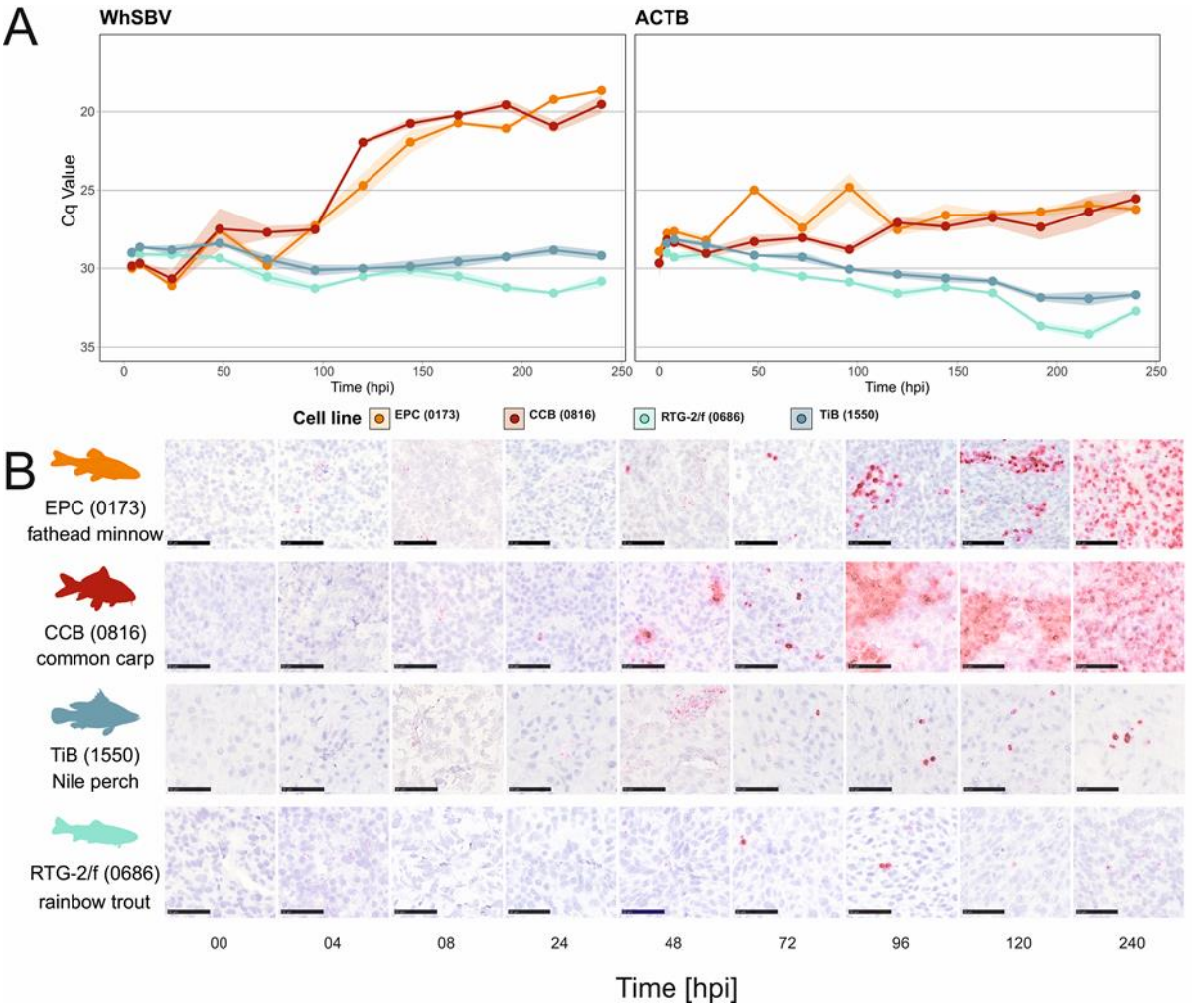
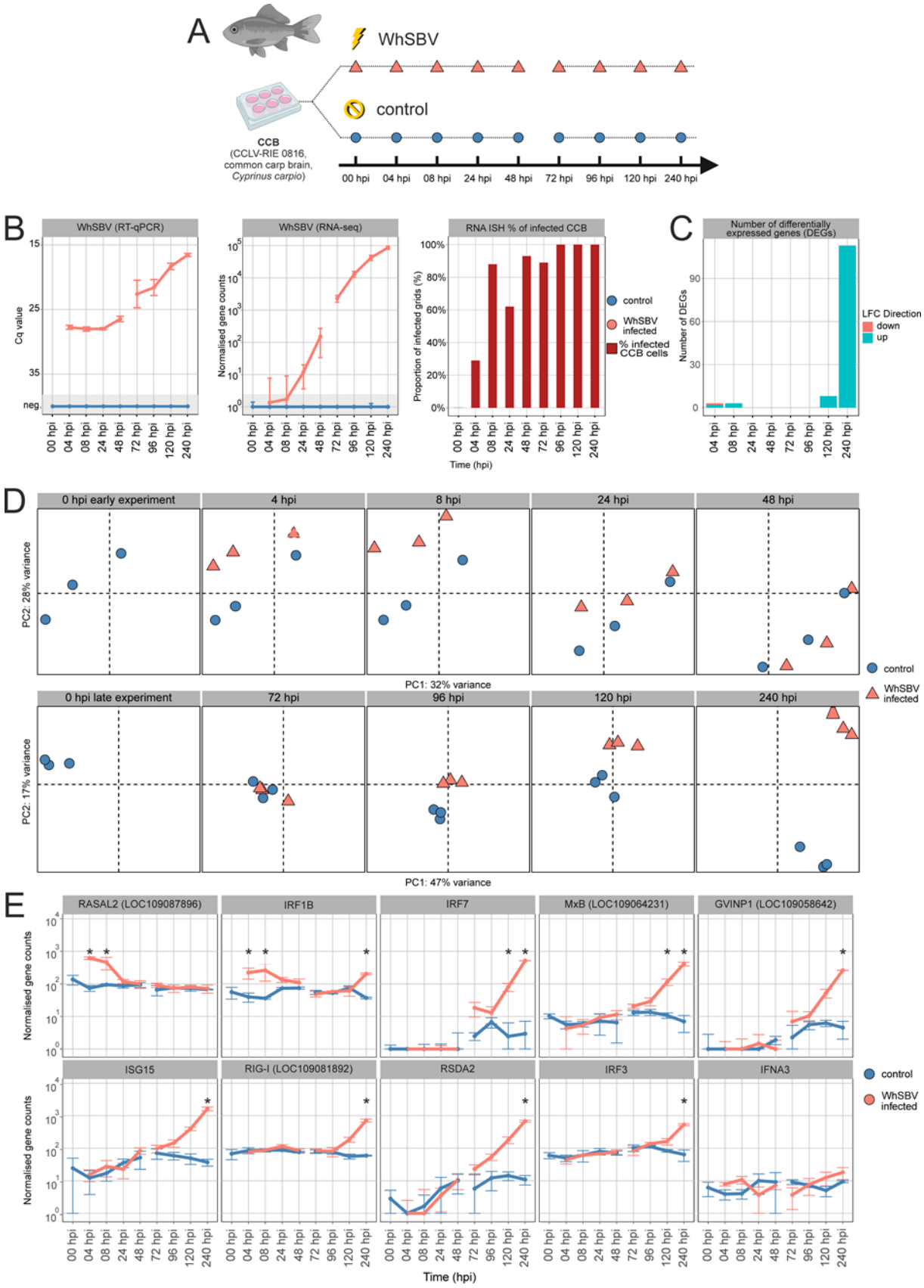


Figure 4: Detection of viral RNA by RT-qPCR and RNA *in situ* hybridisation. (**A**) RNA levels of the viral glycoprotein (WhSBV) and cellular beta actin (ACTB) genes were detected by reverse transcription quantitative PCR (RT-qPCR) and are presented as cycle of quantification (Cq) values for RNA extracted from cell pellets of EPC (orange), CCB (red), RTG-2/f (greenish-blue) and TiB (blue-grey) at different hours post inoculation (hpi). Results are presented as arithmetic means (± standard deviation) of three replicates. (**B**) Detection of viral RNA by RNA *in situ* hybridisation in formalin-fixed cell pellets from WhSBV-inoculated cypriniform (CCB, EPC) and non-cypriniform (TiB, RTG-2/f) cell lines. One representative image is shown for each cell line and time point. The black scale bar represents 50 µm.



953

954 **Figure 5: Cellular response to early and late WhSBV infection in the cypriniform brain**
955 **cells CCB.** (**A**) CCB cells were experimentally WhSBV-infected (orange triangles) and mock-
956 infected (blue dots) in two consecutive experiments, covering early (0, 4, 8, 24, and 48 hours

post infection (hpi)) and late (0, 72, 96, 120, and 240 hpi) infection stages, respectively. (**B**) Viral RNA was detected by reverse transcription quantitative PCR (RT-qPCR; left panel) and RNA sequencing (RNAseq; middle panel). The RT-qPCR targeted the viral glycoprotein and is presented as cycle of quantification values (Cq), while RNAseq quantified sequence reads corresponding to the WhSBV genome and is presented as normalised gene counts. Arithmetic means (± standard deviation, SD) of three replicate wells are shown for Cq values and normalised gene counts. WhSBV-specific RNA *in situ* hybridisation (RNA ISH) data from a separate experiment are shown for comparison (right panel; see Supplementary Figure S4). (**C**) Differentially expressed genes (DEGs) with an absolute log-fold change (LFC) of > 1 and an adjusted *p-value* < 0.05 were derived from the RNAseq data using the DESeq2 package (v1.40), and their number and LFC direction are shown for each time point. (**D**) Principal component analysis (PCA) was performed on the RNAseq gene expression data using R (v4.3.1). Orange triangles and blue dots represent replicates of WhSBV-infected and mock-infected CCB cells, respectively. (**E**) Trends in transcript levels of selected genes known to be key antiviral genes in the early innate immune response to viral infection in fish. These genes include RASAL2 (RAS GTPase-activating protein 2-like [LOC109087896]), IRF1B (interferon regulatory factor 1), IRF7 (interferon regulatory factor 7), IRF3 (interferon regulatory factor 3), MxB (interferon-induced GTP-binding protein MxB-like), GVINP1 (interferon-induced very large GTPase 1-like [LOC109058642]), ISG15 (interferon-stimulated gene 15), RIG-I (antiviral innate immune response receptor RIG-I [LOC109081892]), RSDA2 (viperin, radical S-adenosyl methionine domain containing 2) and IFNA3 (interferon-alpha 3). Asterisks indicate time points with significantly differentially gene expression as compared to uninfected controls. The complete list of DEGs can be found in Supplementary Table S4. Results are presented as the arithmetic/geometric mean (± standard deviation) of three replicate wells.

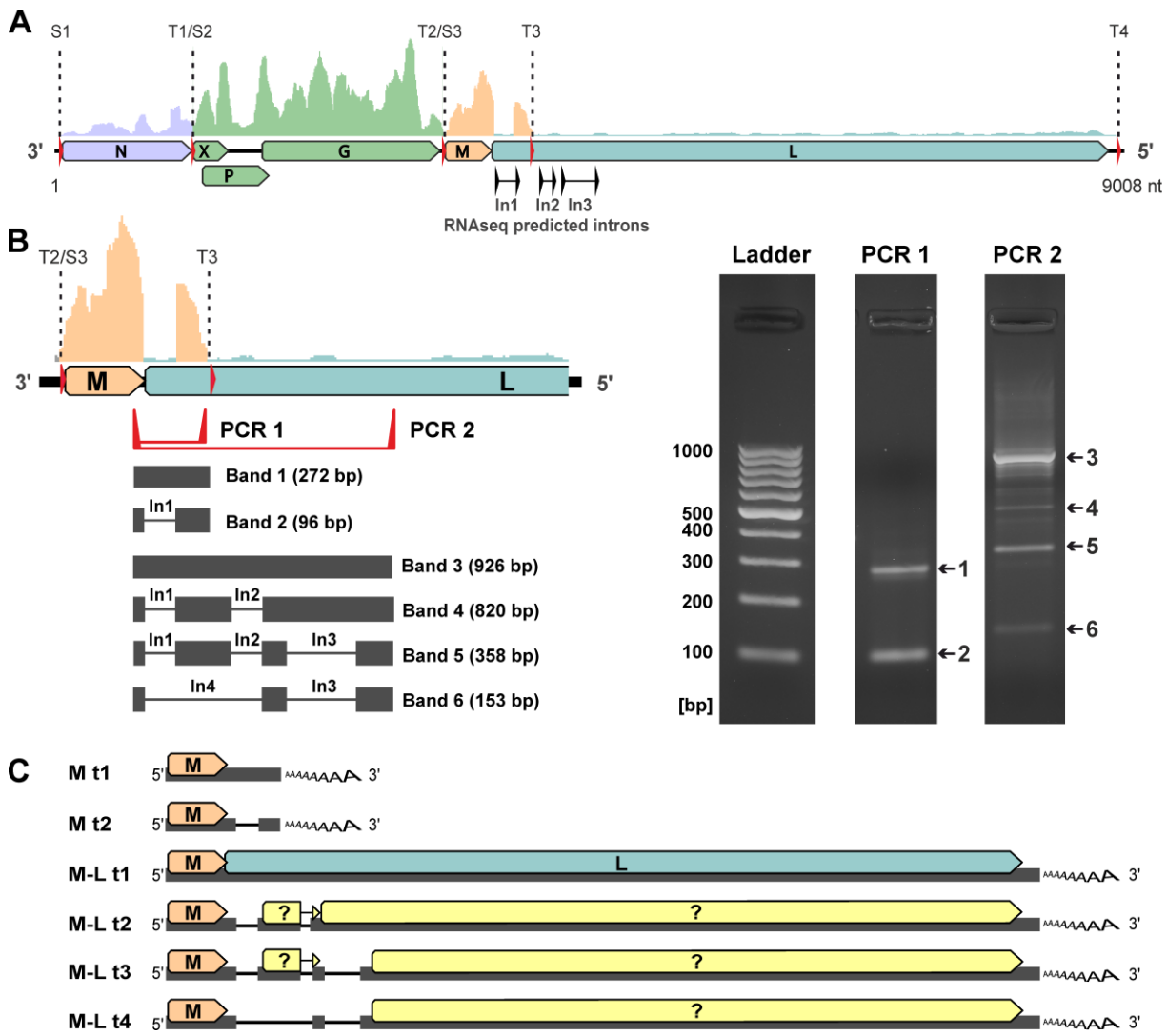


Figure 6: Detection of spliced WhSBV RNA in persistently WhSBV-infected CiSB cells. (**A**) Raw reads from the metagenomic RNA sequencing of the persistently WhSBV-infected CiSB cell line dataset were mapped to the corresponding WhSBV genome and the resulting coverage was plotted. Mapping was performed using the generic Geneious Prime mapper with medium sensitivity settings (options: find structural variants, short insertions, and deletions of any size). In the schematic WhSBV genome, open reading frames (ORFs) are shown as arrows and transcription start sites (S1-S3) and termination sites (T1-4) are shown as dashed lines. ORFs that are likely to be located on the same RNA transcript are shown in the same colour. Three splicing events (introns In1-3) were detected within the L ORF region using the Geneious Prime mapper. (**B**) RNA was extracted from the persistently WhSBV-infected CiSB cell line and cDNA was synthesised using oligo-dT primers. A PCR was performed using this cDNA and primers targeting the predicted M and L splice sites (PCR 1 and 2). The PCR products were run on a 2% agarose gel (right panel). Bands 1-6 were excised, purified and sequenced by Sanger sequencing. The sequences were mapped back to the respective WhSBV genome in

order to confirm the predicted introns In1-4 (black lines). (**C**) Alternative transcripts of WhSBV were inferred from the detected splice sites and transcription start and termination sites. The schematic shows six alternative transcripts, designated as M transcripts 1-2 (M t1-t2) or M-L transcripts 1-4 (t1-t4). Arrows indicate possible ORFs.

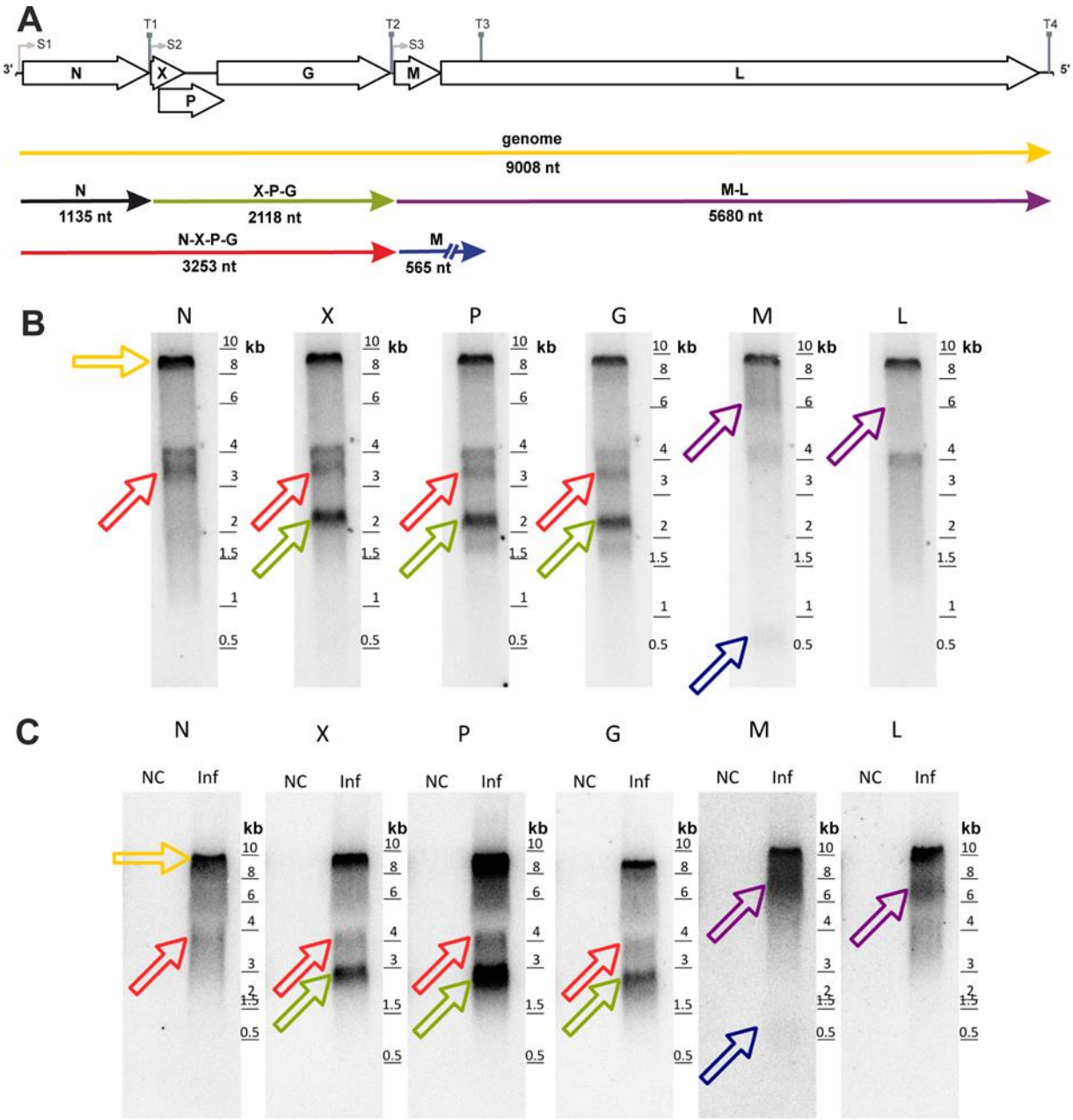


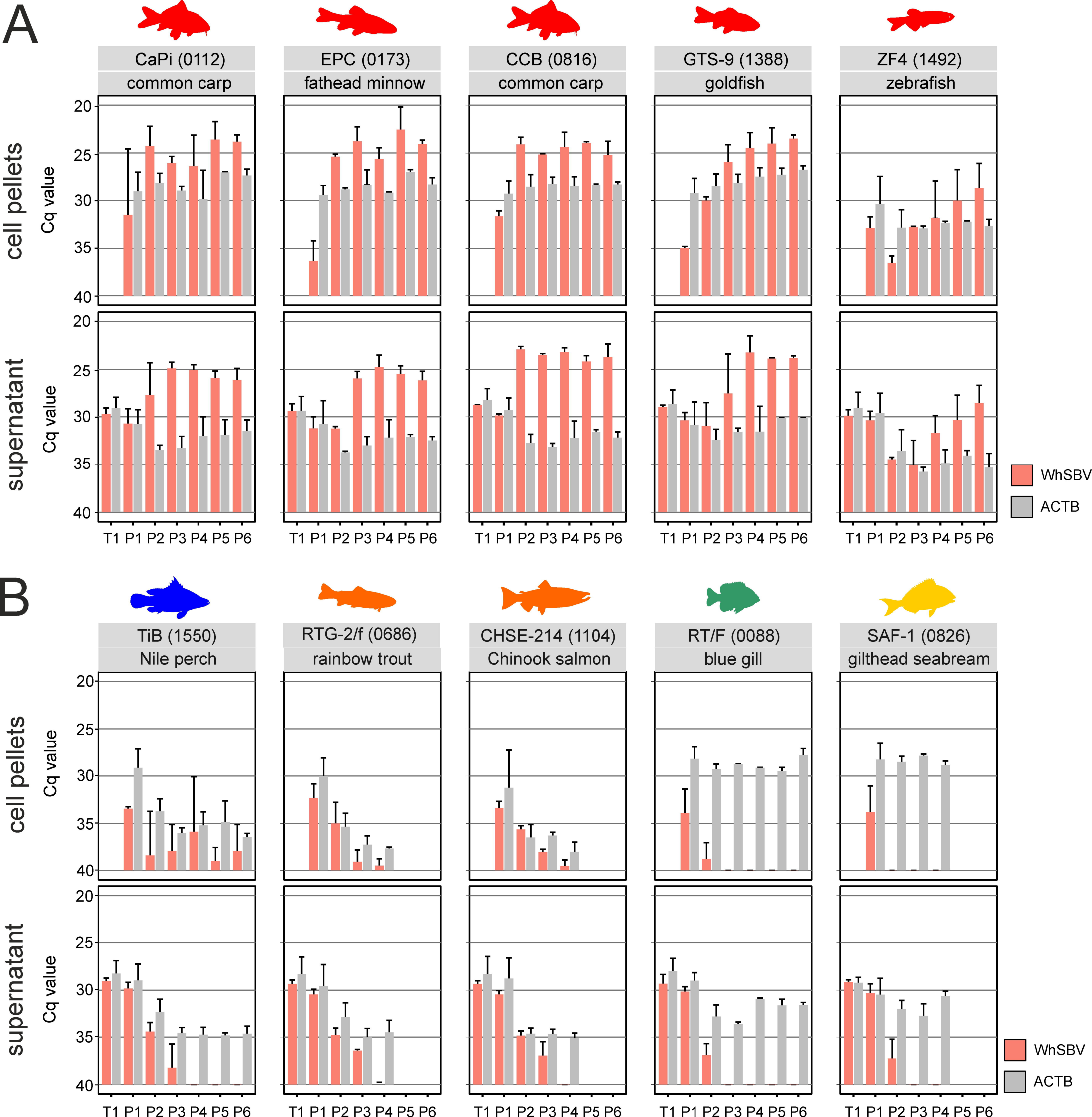
Figure 7: Detection of multicistronic WhSBV RNA in persistently WhSBV-infected and WhSBV-inoculated cells. WhSBV RNA transcripts were detected by Northern blot using ^{32}P-CTP-labelled probes specific for WhSBV genes (arrows) encoding for the nucleoprotein (N), accessory protein (X), phosphoprotein (P), matrix protein (M), glycoprotein (G), and the large

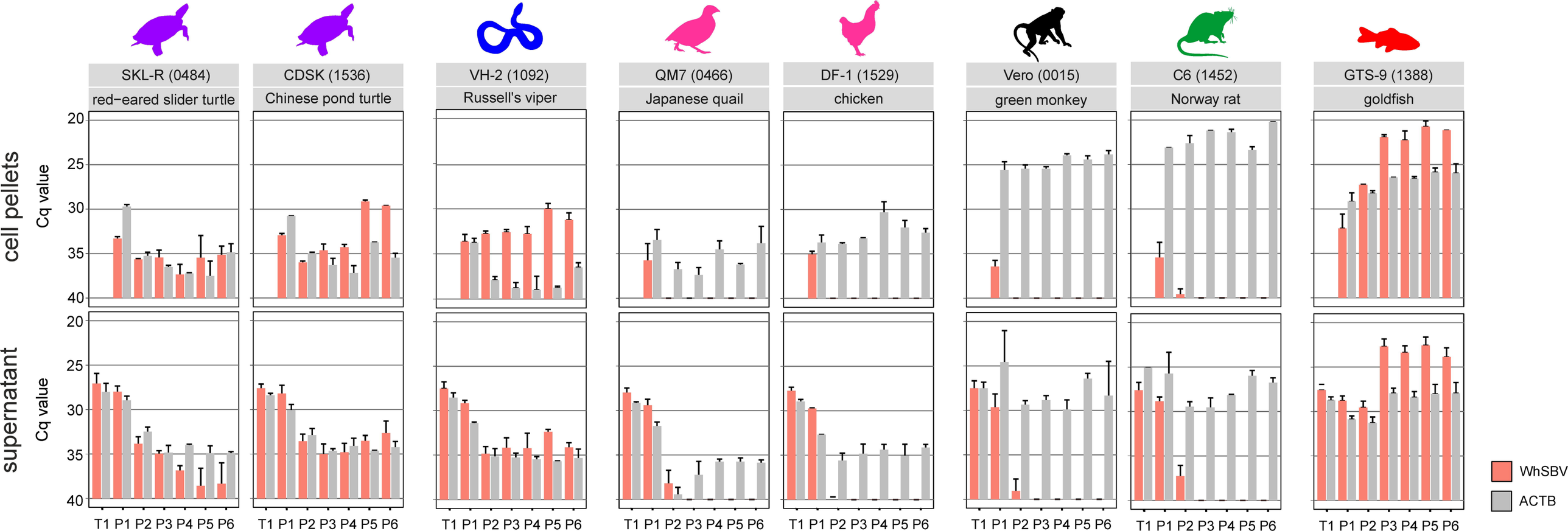
1007 protein (L). For reference, a schematic representation of the WhSBV genome, along with open
1008 reading frames (arrows) and possible RNA molecules (coloured lines) is shown in **(A).** Total
1009 RNA from the persistently WhSBV-infected GSB cell line **(B)** and WhSBV-inoculated (Inf) EPC
1010 cells compared to non-WhSBV-inoculated (NC) EPC cells **(C)**, was used for the Northern blot.
1011 Probes that detected RNA species of the same size were interpreted as corresponding to the
1012 same RNA molecule. The coloured arrows in **(B)** and **(C)** correspond to the respective WhSBV
1013 RNA molecules in the schematic. A ladder is provided in kilobase pairs (kb).

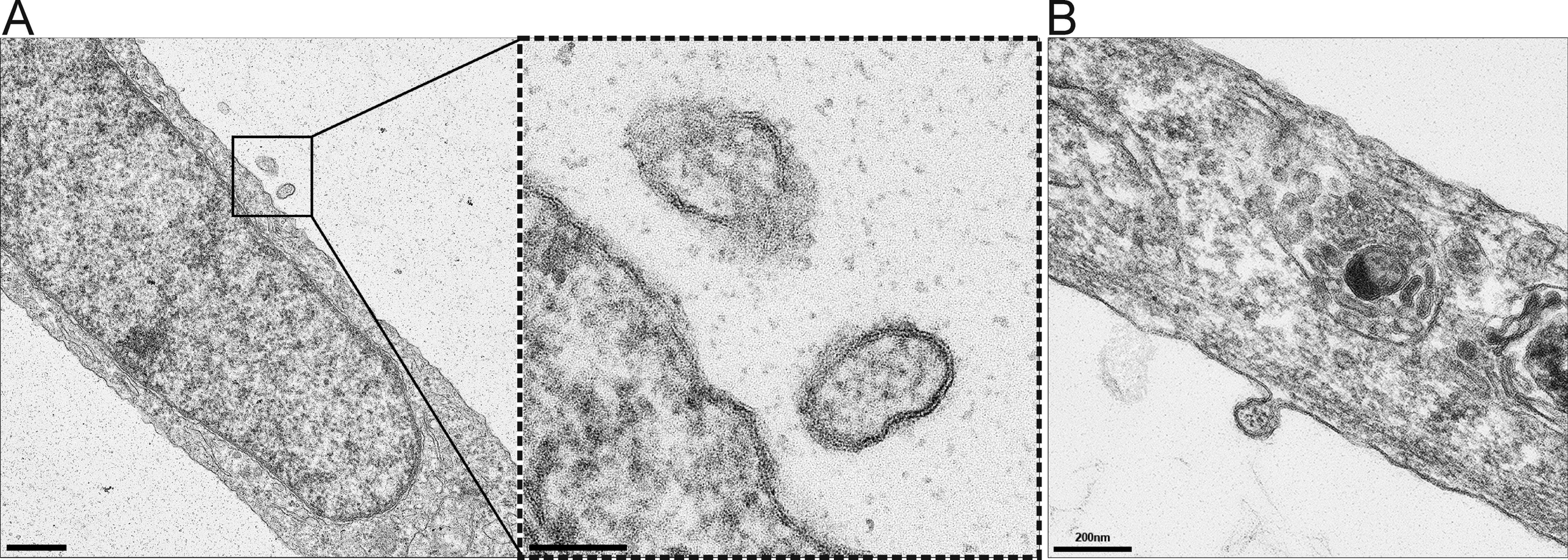
1014

1015

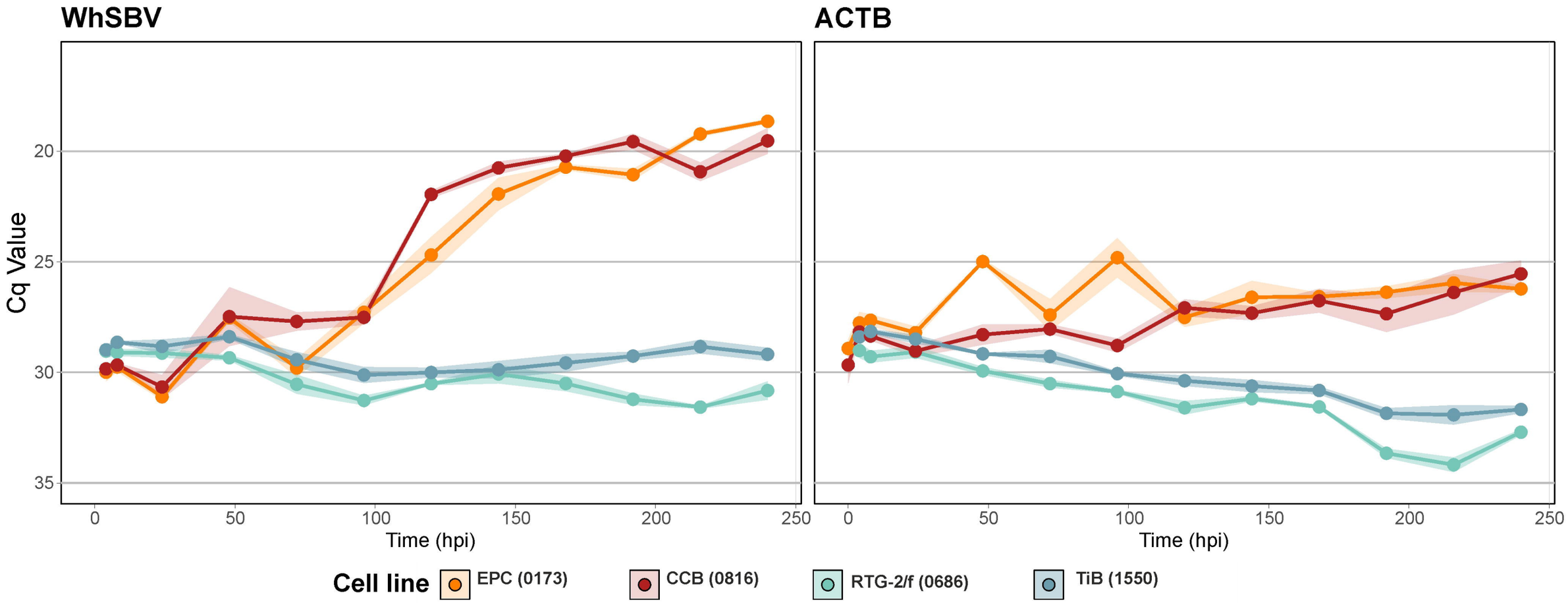
1016 **Figure 8: Detection of WhSBV viral peptides in persistently WhSBV-infected GSB cells.**
1017 High resolution mass spectrometry was performed on protein extracts from the persistently
1018 WhSBV-infected GSB cells. Mass spectrometry raw files were processed with MaxQuant
1019 (version 2.4.13.0) (Cox and Mann 2008) using a custom virus database (6 entries) and the *C.*
1020 *idella* database (C_idella_female_genemodels.v1; 32,811 entries) with default settings. The
1021 figure shows a schematic representation of the complete viral genome, with open reading
1022 frames (ORF) indicated by as arrows and transcription start (S1-S3) and termination sites (T1-
1023 4). Areas corresponding to detected peptides are colour coded for each associated ORF. No
1024 peptides matching the L protein or fused peptides were detected.


A
CaPi (0112)
common carp
EPC (0173)
fathead minnow
CCB (0816)
common carp
GTS-9 (1388)
goldfish
ZF4 (1492)
zebrafish
cell pellets
supernatant
Cq value
20
25
30
35
40
T1 P1 P2 P3 P4 P5 P6
WhSBV
ACTB
B
TiB (1550)
Nile perch
RTG-2/f (0686)
rainbow trout
CHSE-214 (1104)
Chinook salmon
RT/F (0088)
blue gill
SAF-1 (0826)
gilthead seabream
cell pellets
supernatant
Cq value
20
25
30
35
40
T1 P1 P2 P3 P4 P5 P6
WhSBV
ACTB

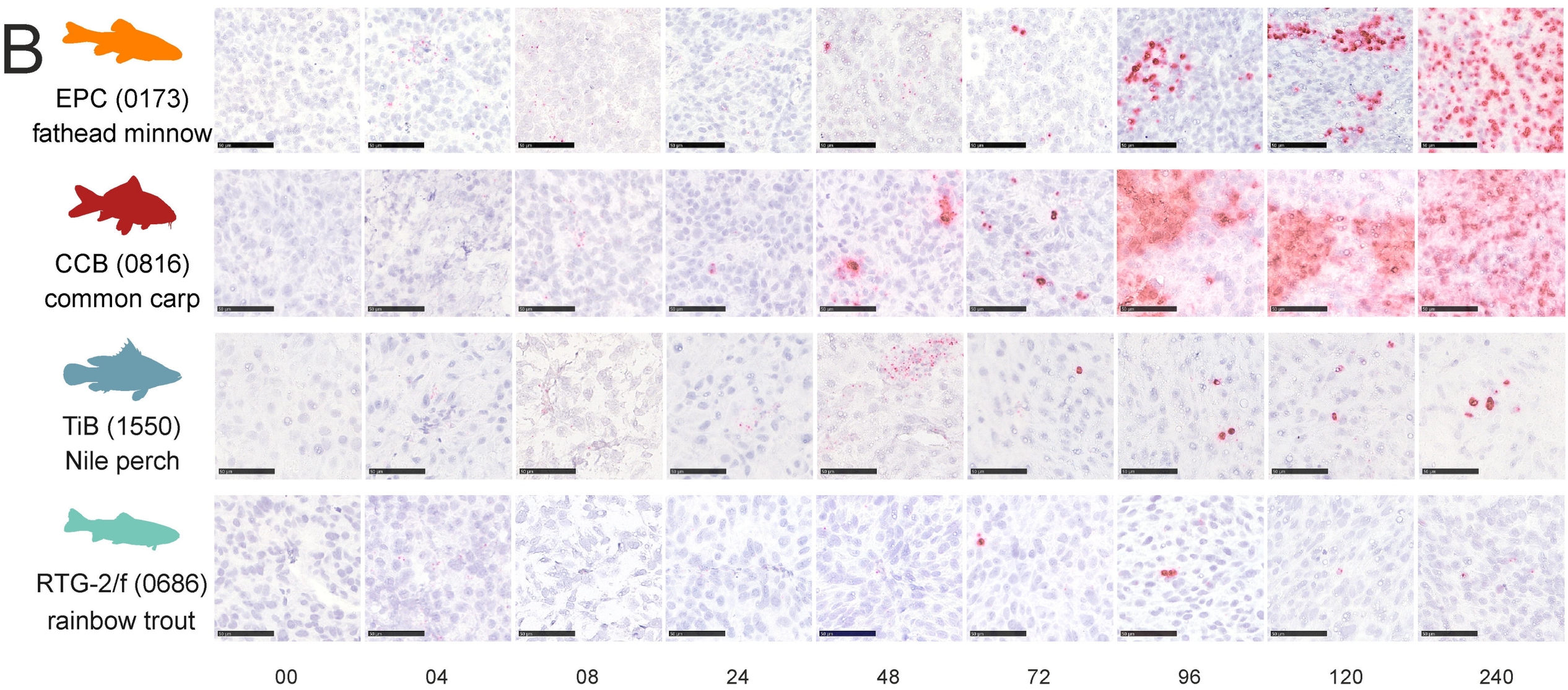

SKL-R (0484)
red−eared slider turtle
CDSK (1536)
Chinese pond turtle
VH-2 (1092)
Russell's viper
QM7 (0466)
Japanese quail
DF-1 (1529)
chicken
Vero (0015)
green monkey
C6 (1452)
Norway rat
GTS-9 (1388)
goldfish
cell pellets
supernatant
Cq value
20
25
30
35
40
T1 P1 P2 P3 P4 P5 P6
WhSBV
ACTB

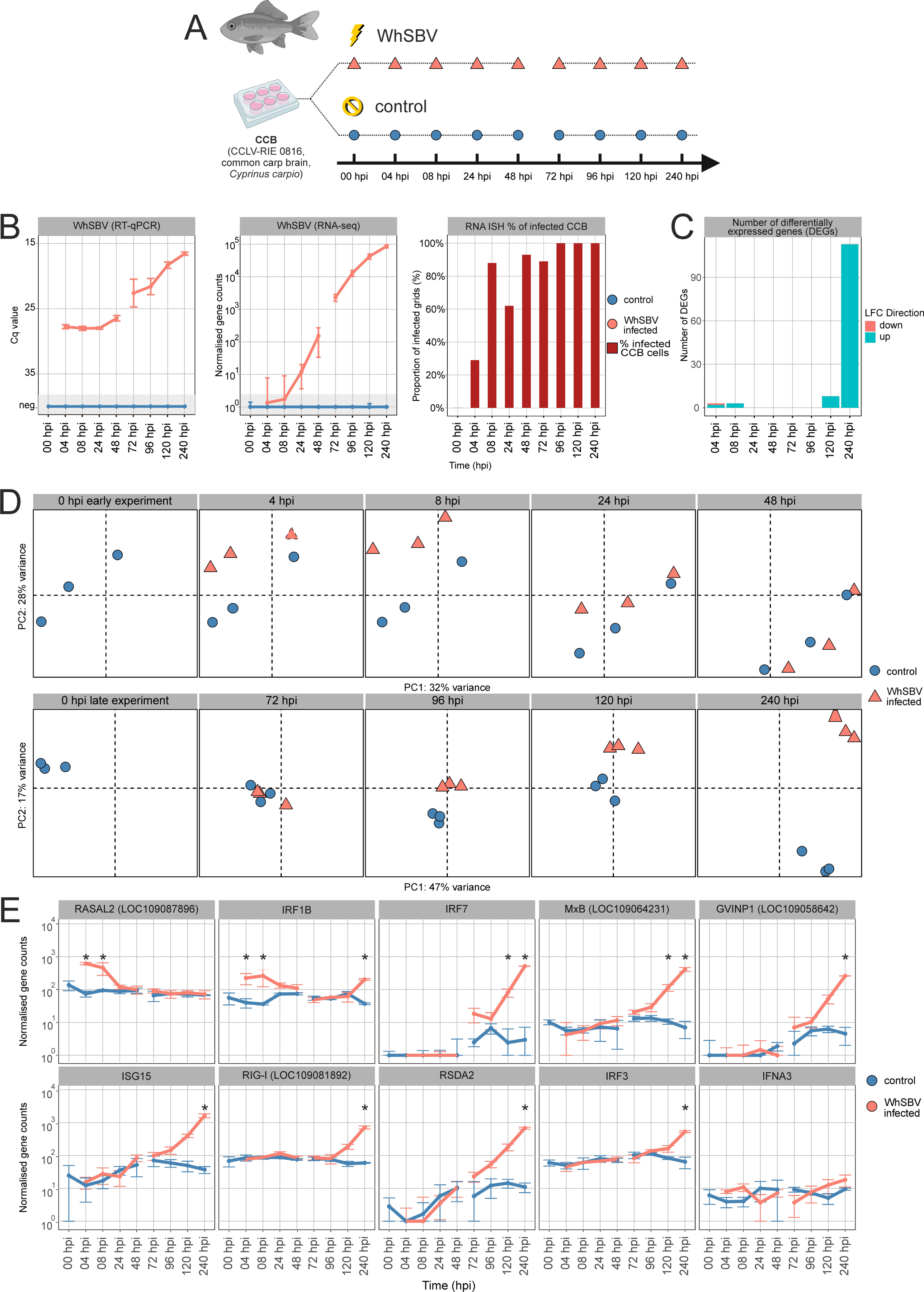

A
B
200nm

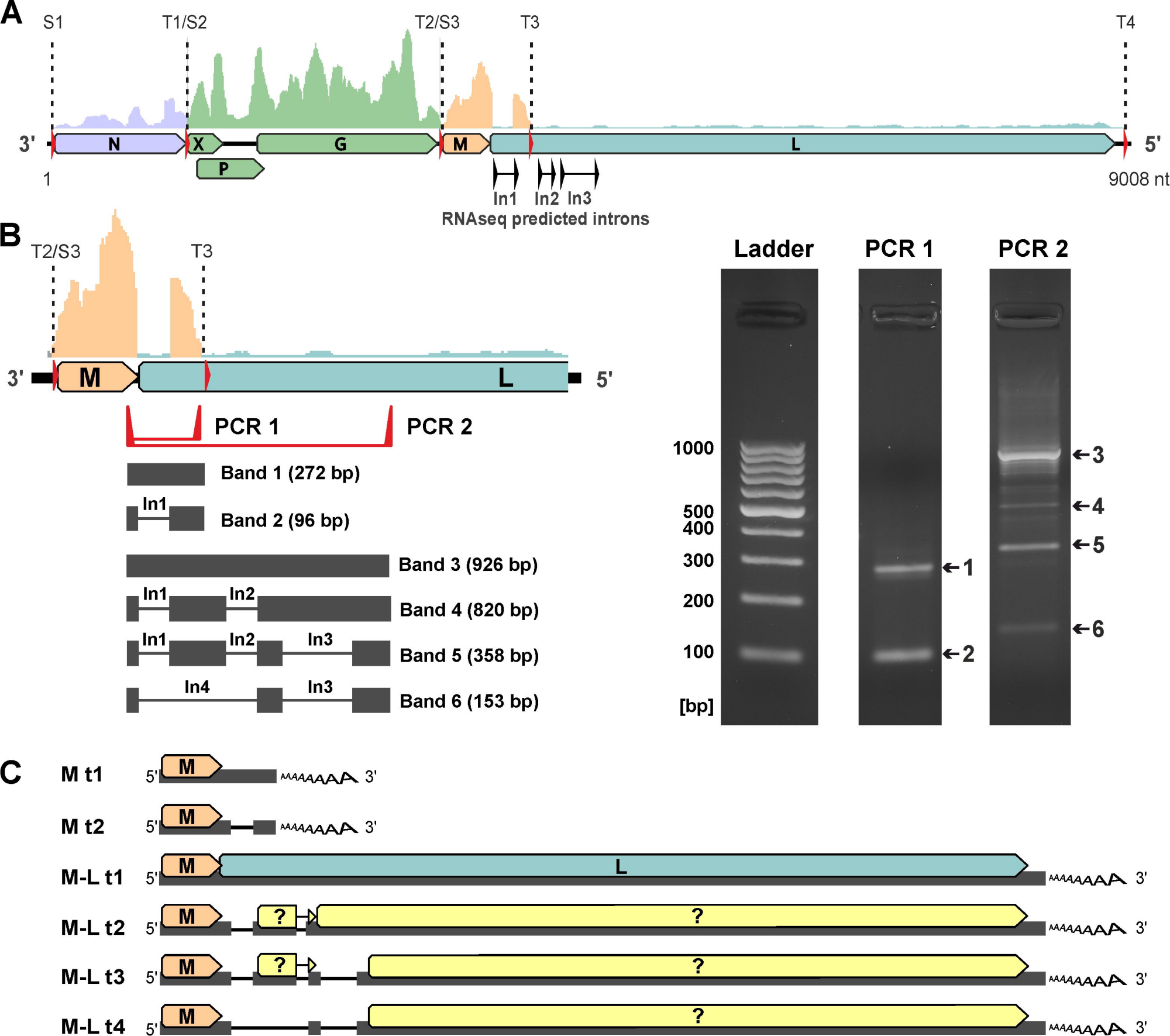
A

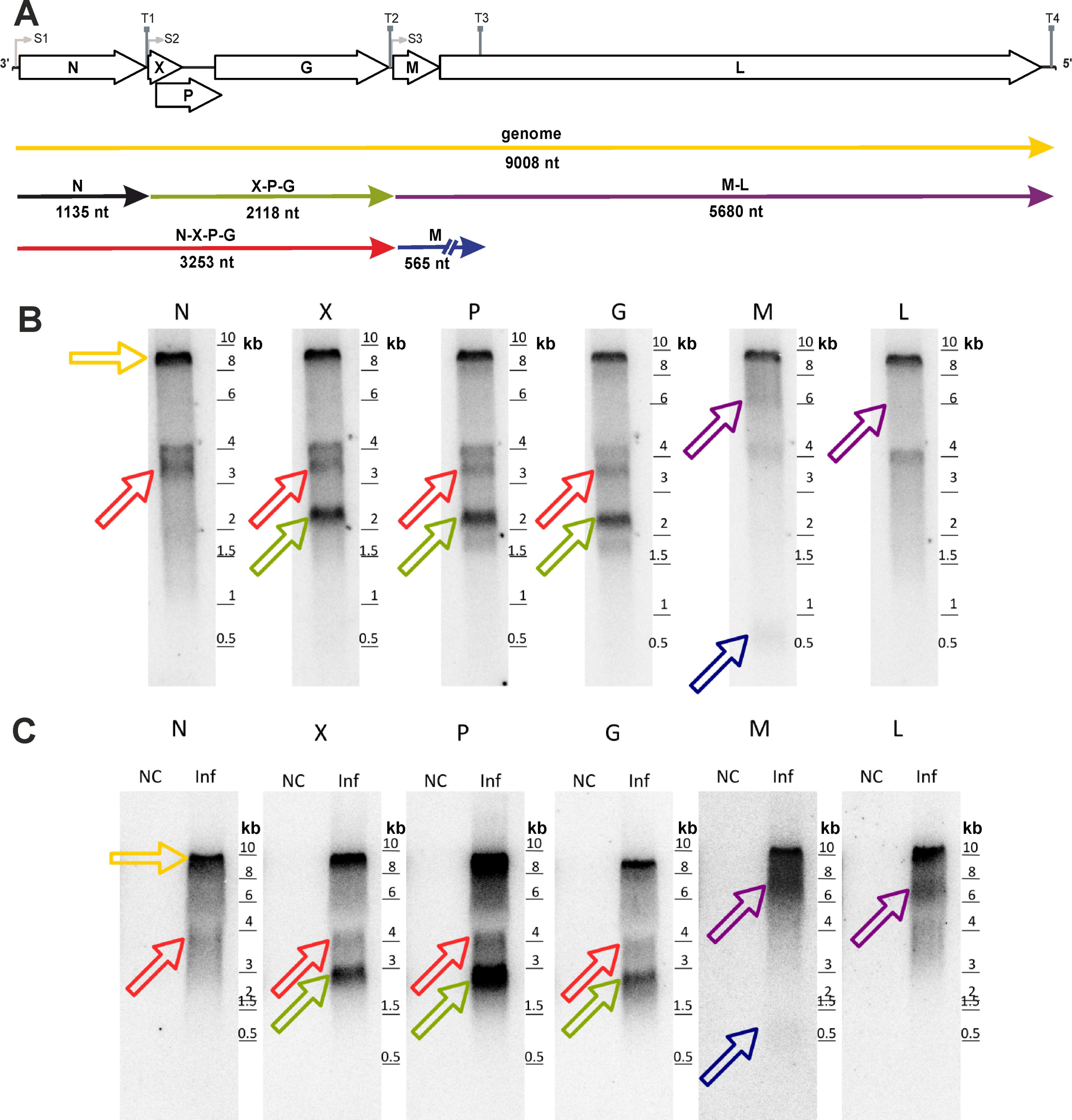

WhSBV
ACTB
Cq Value
20
25
30
35
0
50
100
150
200
250
Time (hpi)
Cell line
EPC (0173)
CCB (0816)
RTG-2/f (0686)
TiB (1550)

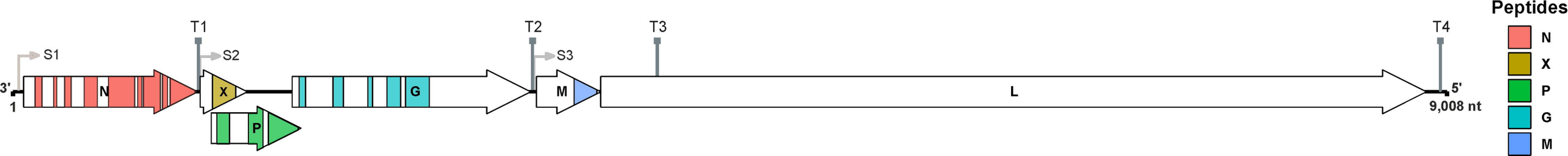
B


EPC (0173)
fathead minnow
CCB (0816)
common carp
TiB (1550)
Nile perch
RTG-2/f (0686)
rainbow trout
00
04
08
24
48
72
96
120
240
Time [hpi]


A
WhSBV
control
CCB
(CCLV-RIE 0816,
common carp brain,
Cyprinus carpio)
00 hpi
04 hpi
08 hpi
24 hpi
48 hpi
72 hpi
96 hpi
120 hpi
240 hpi
B
WhSBV (RT-qPCR)
Cq value
neg.
WhSBV (RNA-seq)
Normalised gene counts
RNA ISH % of infected CCB
Proportion of infected grids (%)
Time (hpi)
control
WhSBV infected
% infected CCB cells
C
Number of differentially expressed genes (DEGs)
Number of DEGs
LFC Direction
down
up
D
0 hpi early experiment
4 hpi
8 hpi
24 hpi
48 hpi
0 hpi late experiment
72 hpi
96 hpi
120 hpi
240 hpi
PC1: 32% variance
PC2: 28% variance
PC2: 17% variance
PC1: 47% variance
control
WhSBV infected
E
RASAL2 (LOC109087896)
IRF1B
IRF7
MxB (LOC109064231)
GVINP1 (LOC109058642)
ISG15
RIG-I (LOC109081892)
RSDA2
IRF3
IFNA3
Normalised gene counts
Time (hpi)


A
S1
T1/S2
T2/S3
T3
T4
3'
N
X
G
M
L
5'
P
1
9008 nt
In1
In2
In3
RNAseq predicted introns
B
T2/S3
T3
3'
M
L
5'
PCR 1
PCR 2
Band 1 (272 bp)
In1
Band 2 (96 bp)
Band 3 (926 bp)
In1
In2
Band 4 (820 bp)
In1
In2
In3
Band 5 (358 bp)
In4
In3
Band 6 (153 bp)
Ladder
PCR 1
PCR 2
1000
500
400
300
200
100
[bp]
1
2
3
4
5
6
C
M t1
M t2
M-L t1
M-L t2
M-L t3
M-L t4
5'
3'
M
L
?
AAAAAAAA


A
T1
T2
T3
T4
S1
S2
S3
3'
5'
N
X
P
G
M
L
genome
9008 nt
N
1135 nt
X-P-G
2118 nt
M-L
5680 nt
N-X-P-G
3253 nt
M
565 nt
B
N
X
P
G
M
L
kb
10
8
6
4
3
2
1.5
1
0.5
C
N
X
P
G
M
L
NC
Inf
kb
10
8
6
4
3
2
1.5
0.5


T1
T2
T3
T4
S1
S2
S3
3'
1
N
X
P
G
M
L
5'
9,008 nt
Peptides
N
X
P
G
M