Journal of Virology

Virology | Full-Length Text

Wǔhàn sharpbelly bornavirus infects and persists in 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] Martin Beer,[1] Dennis Rubbenstroth,[1] Florian Pfaff[1]

AUTHOR AFFILIATIONS See affiliation list on p. 23.

ABSTRACT Our recent study using *in silico* data mining identified novel culteriviruses (family: *Bornaviridae*) in fish, including a variant of Wuhan sharpbelly bornavirus (WhSBV) in grass carp kidney and liver cell lines. Here, metagenomic sequencing of different fish cell lines revealed WhSBV in two cell lines from grass carp (*Ctenopharyngodon idella;* order: Cypriniformes). 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 cells from other bony fish and reptiles. No replication was detected in avian and mammalian cells. *In situ* hybridization and electron microscopy confirmed the presence of viral RNA and particles in infected cypriniform 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 characterized using sequence data alone because isolates are often difficult to obtain. Using cell culture models, this study characterizes Wuhan sharpbelly bornavirus (WhSBV), a member of the genus *Cultervirus*. Here, we demonstrate its ability to establish persistent infection in cypriniform fish cell lines, while exhibiting restricted replication in certain non-cypriniform fish. The identification of polycistronic transcription, splicing events, and immune evasion mechanisms advances our understanding of the molecular biology of WhSBV and culteriviruses 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.

KEYWORDS persistence, transcription, metagenomics, Cypriniformes, fish, *Cultervirus*, *Bornaviridae*, splicing

Editor Colin R. Parrish, Cornell University Baker Institute for Animal Health, Ithaca, New York, USA

Address correspondence to Florian Pfaff, florian.pfaff@fli.de.

The authors declare no conflict of interest.

See the funding table on p. 23.

Received 7 August 2025
Accepted 2 October 2025
Published 30 October 2025

The family *Bornaviridae*, within the order *Mononegavirales*, contains four genera: *Orthobornavirus*, *Carbovirus*, *Cultervirus*, and *Cartilovirus* (1–3). The genus *Orthobornavirus* has a remarkably broad host range, infecting avian, reptilian, and mammalian species (3). Furthermore, orthobornaviruses include zoonotic viruses, such as Borna disease virus 1 (BoDV1) and variegated squirrel bornavirus 1 (VSBV-1), both of which can cause fatal encephalitis in humans (4, 5). Carboviruses and culterviruses, on the other hand, have only been found in reptiles and fish, respectively, and no zoonotic spillover has been reported (2, 3, 6, 7).

Bornavirids possess enveloped virions that enclose a linear, negative-sense, monopartite RNA molecule of approximately 9,000 nucleotides (nt) (3, 8). 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 six viral proteins: nucleoprotein (N), accessory protein (X), phosphoprotein (P), matrix protein (M), glycoprotein (G), and the large protein (L) which contains an RNA-directed RNA polymerase (3, 8). Recently, the largest bornavirid genome, comprising 11,090 nt, was identified for the cartilovirus little skate bornavirus (LSBV). The LSBV genome encodes two additional ORFs, potentially encoding viral proteins 1 and 2 (vp1 and −2), and represents a unique genomic architecture 3′-N-vp1-vp2-X/P-G-M-L-5′ (2). Studies of BoDV-1 have demonstrated that its RNA is transcribed and its genome replicated in the host cell nucleus (9). In bornavirids, the generation of a diverse array of mRNAs is regulated by differential use of transcription start and termination sites (10), as well as alternative splicing of polycistronic primary transcripts to increase and control their transcriptional complexity (11, 12).

Wuhan sharpbelly bornavirus (WhSBV, NC_055169; species *Cultervirus hemicultri*), a member of the genus *Cultervirus*, was initially identified by RNA sequencing of the gut, liver, and gill tissues from a freshwater sharpbelly [*Hemiculter leucisculus* (Basilewsky, 1855), family Xenocyprididae, order Cypriniformes] from China (3, 6). Our recent study using *in silico* data mining of publicly available sequencing data sets led to the identification of novel fish culterviruses, including the first complete genome of the Murray-Darling carp bornavirus (MDCBV, BK063521) from a goldfish [*Carassius auratus* (Linnaeus, 1758), family Cyprinidae, order Cypriniformes] tissue pool data set and a WhSBV variant (BK063520) detected in grass carp [*Ctenopharyngodon idella* (Valenciennes, 1844), family Xenocyprididae, order Cypriniformes] CIK (kidney) and L8824 (liver) cell lines (2). The two viruses belong to the same bornavirid species (*Cultervirus hemicultri*) and share 78% nt overall identity. However, current knowledge of these culterviruses remains limited to *in silico* genetic discoveries and bioinformatic predictions, with no virus isolates available for experimental research.

To deepen our understanding of the biology and abundance of culterviruses, we screened 48 fish-derived cell lines for the presence of known or novel culterviruses using metagenomic RNA sequencing. This revealed the presence of WhSBV in two grass carp swim bladder cell lines, with genomes nearly identical to the aforementioned WhSBV variant BK063520 from the liver (L8824) and kidney (CIK) cell lines. This study represents the first investigation into the biology of WhSBV and lays the groundwork for future research of culterviruses.

RESULTS

Metagenomic screening of different fish cell lines reveals the presence of WhSBV

During metagenomic sequencing of 48 fish cell lines, the cypriniform cell lines GCSB1441 and GCSB1542 were found to contain sequences matching WhSBV (Fig. S1). Viral genomes were assembled from both cell lines, sharing 99.9% nt identity with each other and 99.9% and 87.8% nt identity with WhSBV variants CIK (BK063520) and DSYS4497 (NC_055169), respectively. Both currently published WhSBV reference genomes (NC_055169 and BK063520) were identified through metagenomic sequencing and *de novo* assembly. The NC_055169 genome was 8,989 nt in length, while

the BK063520 variant was 8,985 nt long. Here, specific sequencing extended the WhSBV genome to 9,008 nt (Fig. S2). These results did not alter the internal genome sequence but revealed additional nucleotides beyond those captured in the metagenomic assembly. Compared to the current reference genome sequence in GenBank (RefSeq NC_055169), the complete WhSBV genome sequence presented here is 7 and 11 nt longer at the 3' and 5' ends, respectively.

Based on the complete WhSBV genome sequences, gene-specific reverse transcription quantitative polymerase chain reaction (RT-qPCR) assays targeting the N, P, G, M, and L gene regions were established. All five assays were evaluated using total RNA from the persistently WhSBV-infected GCSB1441 and GCSB1542 and uninfected GK cells as the control. The Cq value of the individual WhSBV gene-specific RT-qPCRs was comparable in GCSB1441 and GCSB1542, while no amplification was observed in GK cells (see Fig. S3). Among the targets, the G gene-specific RT-qPCR assay demonstrated the highest sensitivity and consistently produced the lowest Cq values. It was therefore selected for further screening and for subsequent analyses. Using serial dilutions, a Cq cut-off value of 37 was defined for the G gene-specific RT-qPCR, corresponding to the highest dilution that consistently produced detectable amplification.

Subsequent testing of the 48 cell lines used for metagenomic analysis with the G gene assay identified WhSBV RNA in GCSB1441 and GCSB1542 cell lines, while all other cell lines tested negative (see Table S1). Furthermore, qPCR assays targeting the WhSBV N, P, G, M, and L gene regions without prior reverse transcription were negative for GCSB1441 and GCSB1542 cells, indicating that the detected WhSBV sequences likely represent genuine viral RNA rather than endogenous viral elements integrated into the host DNA genome (see Fig. S3).

Cell lysis is necessary for efficient release of WhSBV viral particles

Five inoculum preparation methods were compared to assess their ability to induce WhSBV replication in CCB cells (Table S2). In the inoculated cultures, WhSBV and cellular ACTB RNA levels were monitored using RT-qPCR over a time period of 240 hours post-infection (hpi) (Fig. 1A). Three cycles of freeze/thawing without subsequent centrifugation appeared to yield the highest amounts of infectious virus as higher viral RNA loads were detectable in the inoculated cells already at early time points, as compared to methods employing sonication with or without freeze/thaw cycles or hypertonic treatment (Fig. 1B). Cells infected with the untreated supernatant barely showed any increase in viral RNA over time, suggesting that the measured viral RNA might have been the residual inoculum rather than a result of an active viral replication (Fig. 1A). Based on its performance, consistency, and practical ease, freeze/thaw lysis was selected for use in subsequent infection experiments.

WhSBV establishes persistency only in cypriniform fish cells

In order to test the infectivity of WhSBV and its potential *in vitro* host range, we used cell lysates from persistently WhSBV-infected GCSB1441 cells to inoculate various cell lines from bony fish, reptiles, birds, and mammals (see Table S2). Following inoculation, the cells were propagated for several passages, and both cell pellets and supernatants were tested for presence of viral RNA using specific RT-qPCR.

WhSBV RNA was detected in the supernatant and cell pellet of all fish cell lines tested during the first one to two passages. In cells from the order Cypriniformes, including cyprinid (CaPi, CCB, and GTS-9), leuciscid (EPC), and danionid (ZF4) species, viral RNA was retained over all six consecutive passages. The amount of viral RNA detected started to increase compared to freshly inoculated cells (T1) and eventually reached a plateau at around Cq 22 in both supernatants and cell pellets. This stabilization occurred between passages 2 and 4 for the CaPi, EPC, CCB, and GTS-9 cell lines. For zebrafish cell line ZF4, the amount of viral RNA initially decreased after inoculation and passaging until passages 2–3, but then recovered, reaching Cq values of about 27 in the supernatant and cell pellets (Fig. 2A)

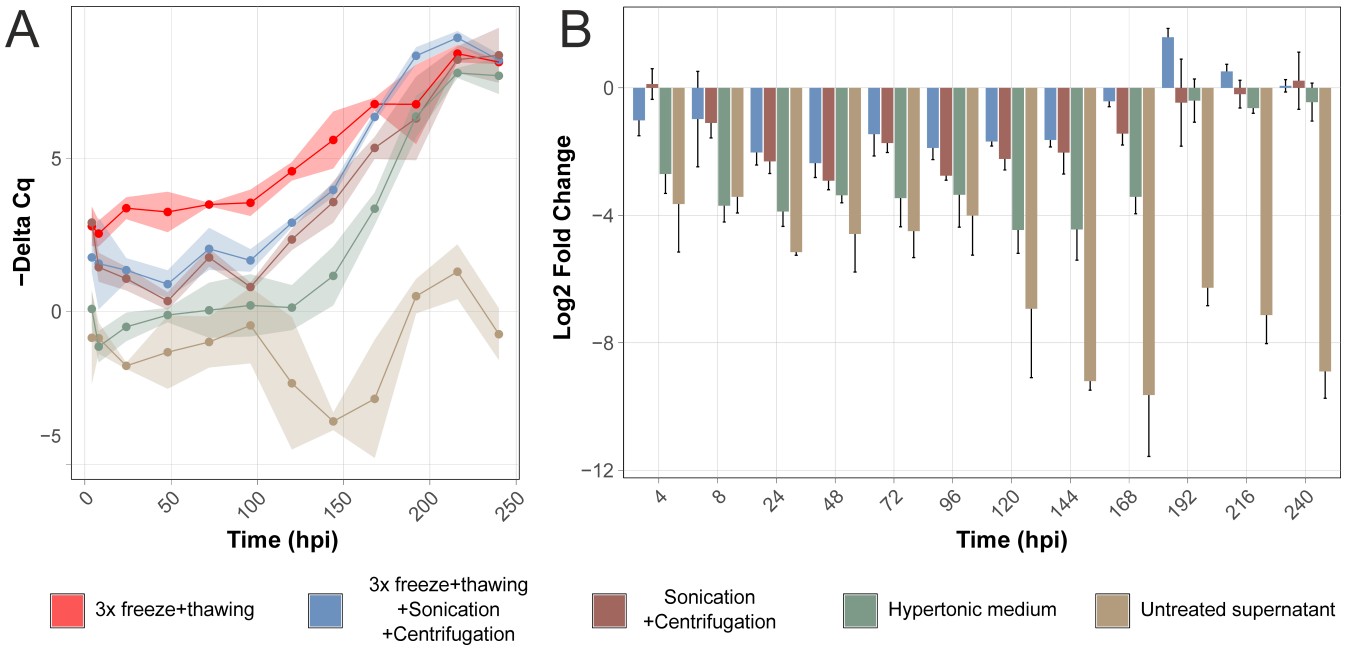

FIG 1 Evaluation of inoculum preparation for WhSBV propagation in CCB cells. (A) CCB cells were experimentally infected with WhSBV preparations, which had been produced from persistently WhSBV-infected GCSB1441 cells using five different methods (each represented by color). Cell pellets were harvested at 0, 4, 8, 24, 48, 72, 96, 120, 144, 168, 192, 216, and 240 hours post infection (hpi), and RNA levels of the viral glycoprotein (G) and cellular beta actin (ACTB) genes were detected by reverse transcription quantitative PCR (RT-qPCR). Values are presented as negative "Delta Cq." (B) Relative change of viral RNA levels in CCB cells over time in comparison to "3 x freeze/thawing" method. Values are presented as $\log_2$ fold changes calculated using the "Delta Delta Cq" method. Each point/bar represents the mean of three technical replicates (± standard deviation).

In contrast, the non-cypriniform fish, centrarchid (RT/F), and sparid (SAF-1) showed no evidence of WhSBV replication after inoculation (Fig. 2B). Immediately after inoculation (T1), viral RNA levels in the supernatant were similar to those of the cypriniform cell lines (Fig. 2A), but dropped below the detection limit in cell pellets and supernatant after passage 2 (Fig. 2B). Although viral RNA was detected in SAF-1 cell pellets at passage 4, the levels remained below the defined Cq cut-off and were not indicative of active replication. The cichlid (TiB) and salmonid (RTG-2/f and CHSE-214) cell lines exhibited reduced cell growth over the course of the experiment, which was reflected by a decline in the measured ACTB RNA levels, a trend also observed in corresponding uninfected controls. Viral RNA remained detectable in cell pellets until the end of the experiment up to passage 6 for (TiB) and passage 4 for (RTG-2/f and CHSE-214), but without a detectable increase in levels. In several instances, Cq values fluctuated near or fell below the established detection cut-off, indicating the absence of sustained viral replication (Fig. 2B).

In the reptilian cell lines SKL-R, VH-2, and CDSK, WhSBV RNA levels in the culture supernatant decreased after inoculation but remained detectable up to passage 6, with Cq values stabilizing around 28 and 35; however, viral RNA levels fell below the defined Cq cut-off at passages 5 and 6 in SKL-R (Fig. 3). In the SKL-R cell pellets, viral RNA was still detectable up to passage 6, although consistently at low levels, yet above the cut-off. In contrast, CDSK and VH-2 pellets showed an increase in viral RNA levels during passages 5 and 6, reaching Cq values of 28 and 30, respectively. A decline in ACTB RNA levels was also observed in the reptilian cell lines, mirroring the trend in cichlid and salmonid cell lines. Similar decreases in ACTB expression were also noted in the corresponding uninfected controls.

In the mammalian Vero and C6 cell lines and the avian QM7 and DF-1 cell lines, WhSBV RNA levels in the cell pellets or supernatants dropped below or near the detection limit already after passages 1–3 (Fig. 3).

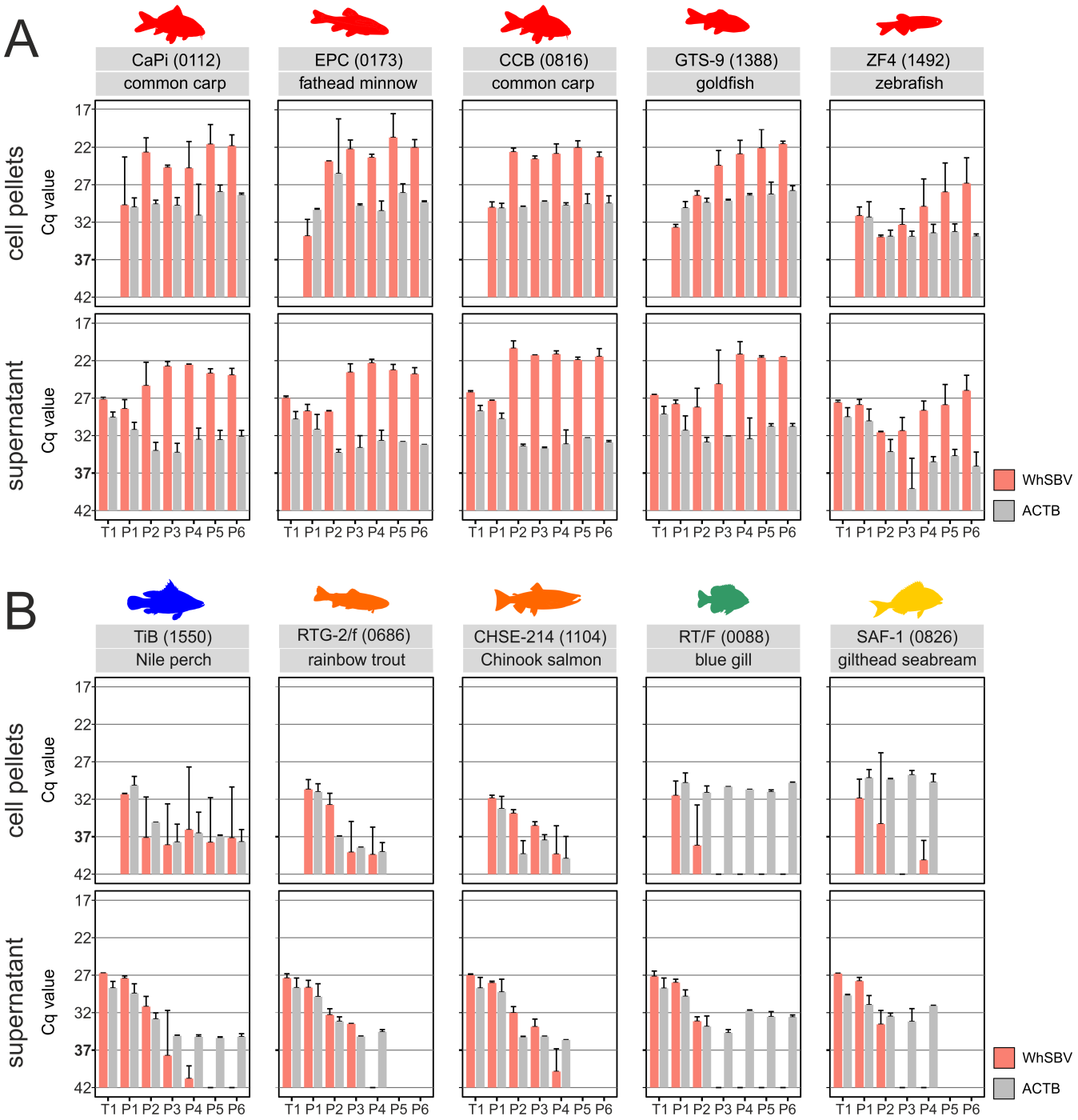

FIG 2 WhSBV RNA levels over successive passages in different fish cell lines. Different fish cell lines were inoculated with WhSBV. The cells were then passaged for four to six times by splitting. (A) Cell lines from the taxonomic order of Cypriniformes (red); (B) cell lines from the taxonomic orders of Cichliformes (blue), Salmoniformes (orange), Centrarchiformes (green), and Spariformes (yellow). RNA levels of viral glycoprotein (WhSBV; red bars) and cellular beta actin (ACTB; gray 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. A Cq cut-off value of 37 was set as the limit of detection for WhSBV.

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

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

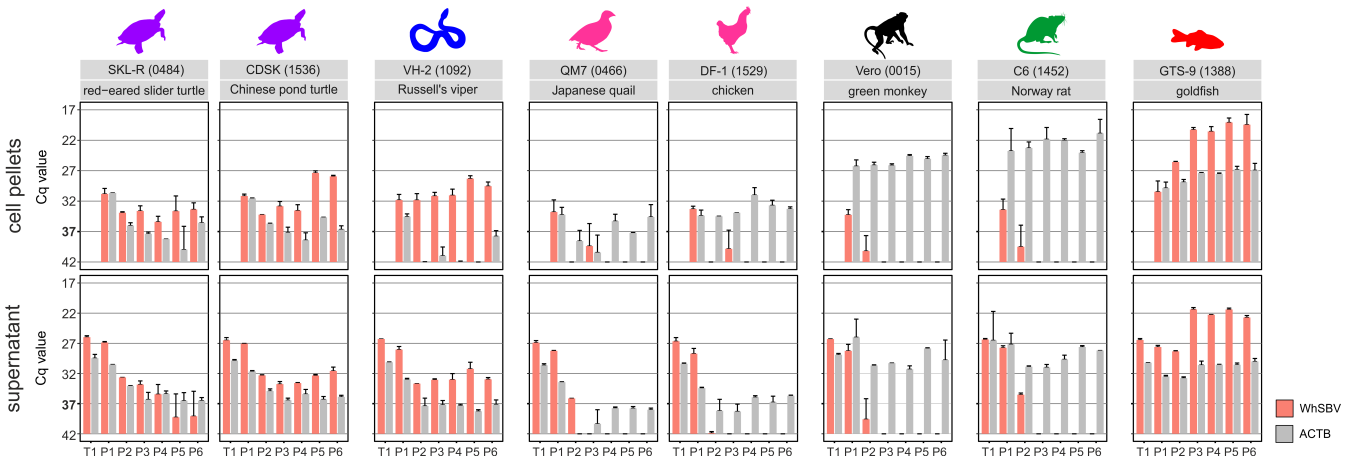

FIG 3 WhSBV RNA levels over successive passages in different non-fish cell lines. Different cell lines from the taxonomic orders of Testudines (purple), Squamata (blue), Galliformes (pink), Primates (black), and Rodentia (green) were inoculated with WhSBV and subsequently passaged for six times by splitting 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; gray 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. A Cq cut-off value of 37 was set as the limit of detection for WhSBV.

Using transmission electron microscopy, putative virus particle-like structures were observed in WhSBV-inoculated ZF4 cell lines, whereas similar structures were absent in non-inoculated controls. These structures appeared spherical to pleomorphic without visible peplomers and a diameter of approximately 150–180 nm (Fig. 4A). Potential membrane-associated budding was rarely observed at the plasma membrane of the persistently WhSBV-infected GCSB1441 cells (Fig. 4B)

WhSBV replicates efficiently only in cypriniform cell lines

In order to gain deeper insights into viral propagation in different fish cell lines, a kinetic infection experiment was performed over 10 days using both cypriniform (EPC and CCB) and non-cypriniform fish cells (RTG-2/f and TiB). The level of viral RNA load was assessed

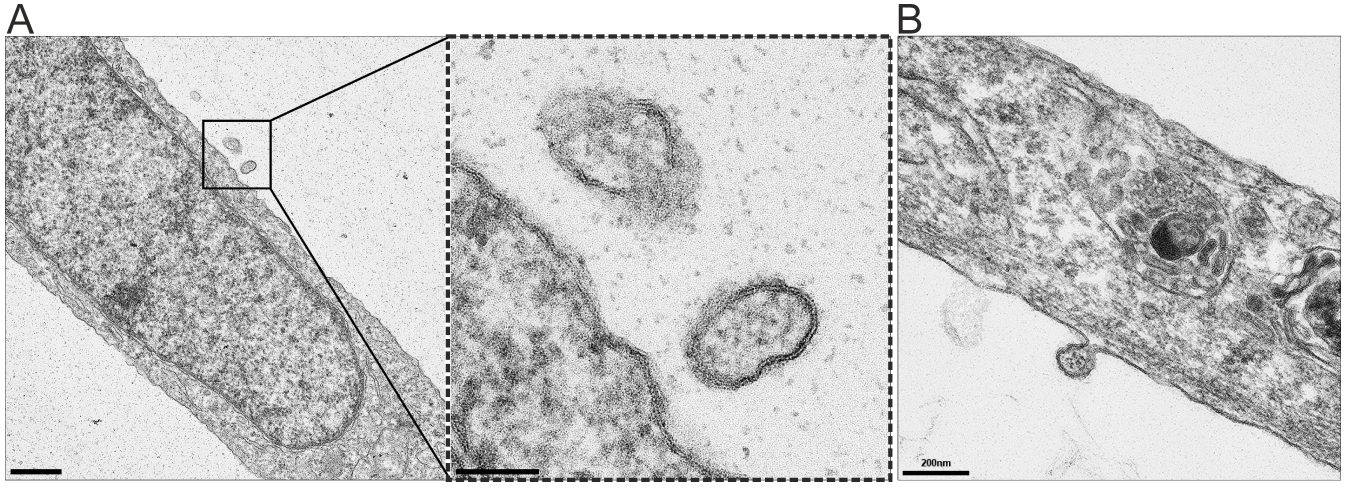

FIG 4 Detection of putative WhSBV-like particles by transmission electron microscopy. (A) Putative virion-like structures observed in the experimentally infected ZF4 cell lines. These structures appeared spherical to pleomorphic, with diameters ranging from approximately 150–180 nm. (B) Rare membrane-associated budding events were observed at the plasma membrane of persistently WhSBV-infected GCSB1441 cells, which may represent viral budding.

by RT-qPCR, and the presence of viral RNA at the cellular level was assessed by RNA *in situ* hybridization (RNA ISH) at different time points.

In the cypriniform EPC and CCB cell lines, viral RNA levels increased exponentially until 12 hpi, reaching Cq values of about 18.6 in EPC and 19.5 in CCB. In contrast, the viral RNA levels in salmonid RTG-2/f and cichlid TiB cell lines showed stagnation or a linear increase (Fig. 5A). The presence of viral (-)RNA was detected by specific RNA ISH in both cypriniform (EPC and CCB) and non-cypriniform fish cell lines (RTG-2/f and TiB) from 4 hpi onward. In detail, RNA ISH confirmed the RT-qPCR results from the kinetic experiment as the number of infected cells, indicated by the presence of viral (-)RNA, increased in a time-dependent manner for the cypriniform cell lines EPC and CCB. At early time points (4–72 hpi), infected single cells appeared, followed by foci of positive cells, later confluent and finally diffuse labeling of the cell pellet (Fig. 5B). Percentage quantification of individual cell pellets for EPC showed widespread infection, with 95% of the fields analyzed showing positive detection as early as 4 hpi. Infection remained widespread (63-100% of infected test fields) throughout the course of the experiment until 240 hpi (see Fig. S4). In CCB cells, only 29% of the test fields were infected at 4 hpi and longer cultivation correlated with a higher proportion of positive test fields, eventually reaching 100% at 96, 120, and 240 hpi (see Fig. S4). Similar observations were made using a custom-designed probe against WhSBV (+)RNA on CCB cell lines (Fig. S5). In terms of the labeling signal, infected cells harvested at early time points were more likely to show fine-granular cytoplasmic labeling, and larger, globular signals were consistently found at 48 hpi and later (Fig. 5B). In clear contrast, the infection pattern in the non-cypriniform fish cell lines was essentially confined to single cells and showed no evidence of increasing infection abundance over time (Fig. 5B). Only the intensity of the chromogen signal within the cell increased in TiB from 72 hpi, but in RTG-2/f; this phenomenon was limited to 72 and 96 hpi. RTG-2/f and TiB cells ultimately yielded 28% and 16% infected test fields at 240 hpi, respectively (see Fig. S4).

For mammalian Vero and avian DF-1 cell lines, only a single time point at 240 hpi was collected, alongside CCB which was used as a positive control. RNA ISH analysis of mammalian and avian DF-1 cell lines showed no detectable WhSBV RNA at either the control 0 hpi or at 240 hpi (data not shown).

Using routine hematoxylin-eosin staining, neither viral inclusion bodies nor other morphological evidence of a cytopathogenic effect could be identified in EPC, CCB, RTG-2/f, TiB, Vero, and DF-1 cell lines (see Fig. S6).

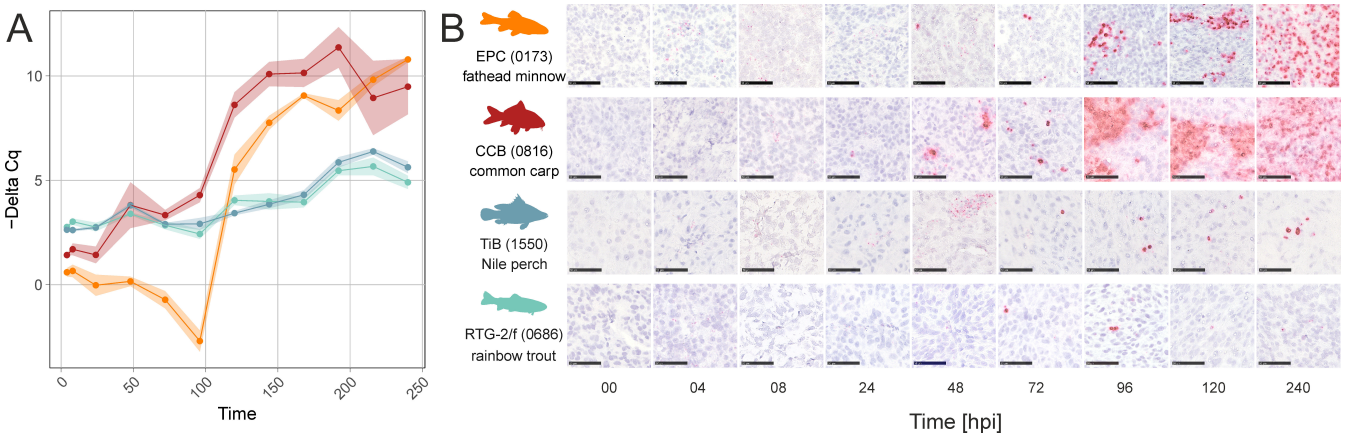

FIG 5 Detection of viral RNA by RT-qPCR and RNA *in situ* hybridization. (A) RNA levels of the viral glycoprotein (WhSBV) and cellular beta actin (ACTB) genes were detected by reverse transcription quantitative PCR (RT-qPCR) and are presented as negative Delta cycle of quantification (-Delta Cq) values for RNA extracted from cell pellets of EPC (orange), CCB (red), RTG-2/f (greenish-blue) and TiB (blue-gray) at different hours post inoculation (hpi). Results are presented as arithmetic means (± standard deviation) of three replicates. A Cq cut-off value of 37 was set as the limit of detection for WhSBV. (B) Detection of viral RNA by RNA *in situ* hybridization 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.

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

To investigate cellular transcriptional changes associated with WhSBV infection, we performed RNA sequencing (RNAseq) of RNA transcripts at early and late time points following viral inoculation in CCB cells (Fig. 6A). Quantitative RT-qPCR and sequencing confirmed the presence of WhSBV RNA and the onset of infection at 4 hpi, with viral RNA levels progressively increasing at all subsequent time points (Fig. 6B). This trend was consistent with the RNA ISH data, which showed a gradual increase in the percentage of infected CCB cells over time, reaching 100% infection at 96, 120, and 240 hpi (Fig. 6B)

The number of differentially expressed genes (DEGs) remained minimal at early and late time points but became more pronounced at 240 hpi, indicating a delayed transcriptional response to infection (Fig. 6C). Principal component analysis (PCA) of early time points revealed clear transcriptional differences between WhSBV-infected and control cells at 4 and 8 hpi, which diminished at 24 and 48 hpi, with both groups becoming more similar (Fig. 6C). In contrast, the PCA of late time points experiment showed no significant transcriptional changes at 72 and 96 hpi. However, minimal transcriptional alterations were observed at 120 hpi, followed by distinct clustering of infected and control samples at 240 hpi, suggesting a change in the host transcriptional response at later stages of infection (Fig. 6D).

Differential gene expression analysis of early time points identified three DEGs at 4 or 8 hpi. Specifically, RASAL2 (Ras GTPase-activating activator like 2 [LOC109087896]) and IRF1B (interferon regulatory factor 1b) were upregulated at 4 hpi and 8 hpi (Fig. 6E), whereas EDEM3 (ER degradation enhancing alpha-mannosidase-like 3 [LOC109095466]) was upregulated exclusively at 8 hpi. Conversely, ABCC1 (multidrug resistance-associated 1-like [LOC109090229]) was downregulated at 4 hpi. The expression of these genes returned to baseline levels by 24 and 48 hpi. At later stages of infection (72 and 96 hpi), no DEGs were detected between infected and control cells. At 120 hpi, eight DEGs were upregulated, including IRF7 (interferon regulatory factor 7) and MxB (MX dynamin like GTPase 2) (Fig. 6E). At 240 hpi, 113 DEGs were upregulated, including IRF7, MxB, and several other key antiviral genes involved in early innate immune responses in fish, such as IRF3, ISG15 (interferon-stimulated gene 15), and RSDA2 (viperin, radical S-adenosyl methionine domai-containing 2), along with other genes involved in interferon pathways, RIG-I (antiviral innate immune response receptor RIG-I [LOC109081892]) and GVINP1 (interferon-induced very large GTPase 1-like [LOC109058642]) (Fig. 6E; Table S3 for complete list of DEGs).

Detection of spliced WhSBV gene transcripts

Mapping of sequence reads from the persistently WhSBV-infected GCSB1441 and GCSB1542 cells to the complete WhSBV genome (Fig. 7A) revealed the presence of three potential introns (In1-3) located at the beginning of the L ORF. Conventional RT-PCR and Sanger sequencing confirmed the presence of In1-3 within the L ORF and also revealed an additional, unpredicted splice site, In4, within the L ORF (Fig. 7B). Based on the above findings, we predicted several potential transcripts for the M and L genes, which share the same start site S3 and terminate at either T3 or T4. The M gene could produce a short transcript encoding the M ORF, which can either remain unspliced (M t1; size: 741 nt) or undergo splicing at In1 to produce a shorter transcript (M t2; 565 nt), both theoretically terminating at T3. Although the unspliced transcript M t1 was detected by RT-PCR, it appears to be in the minority of transcripts based on RNAseq results. In addition, four transcripts were predicted, termed M-L t1-4, all sharing the same start site S3 and termination site T4, but differing in their splicing patterns. The longest unspliced transcript, M-L t1, contains the full-length M and L ORFs (5,680 nt), whereas the other transcripts (M-L t2-4) are progressively truncated by splicing at different introns, resulting in transcripts of size 5,218 nt, 5,150 nt, and 4,945 nt, respectively (Fig. 7C). These transcripts may encode either truncated versions of the L ORF or only the M ORF.

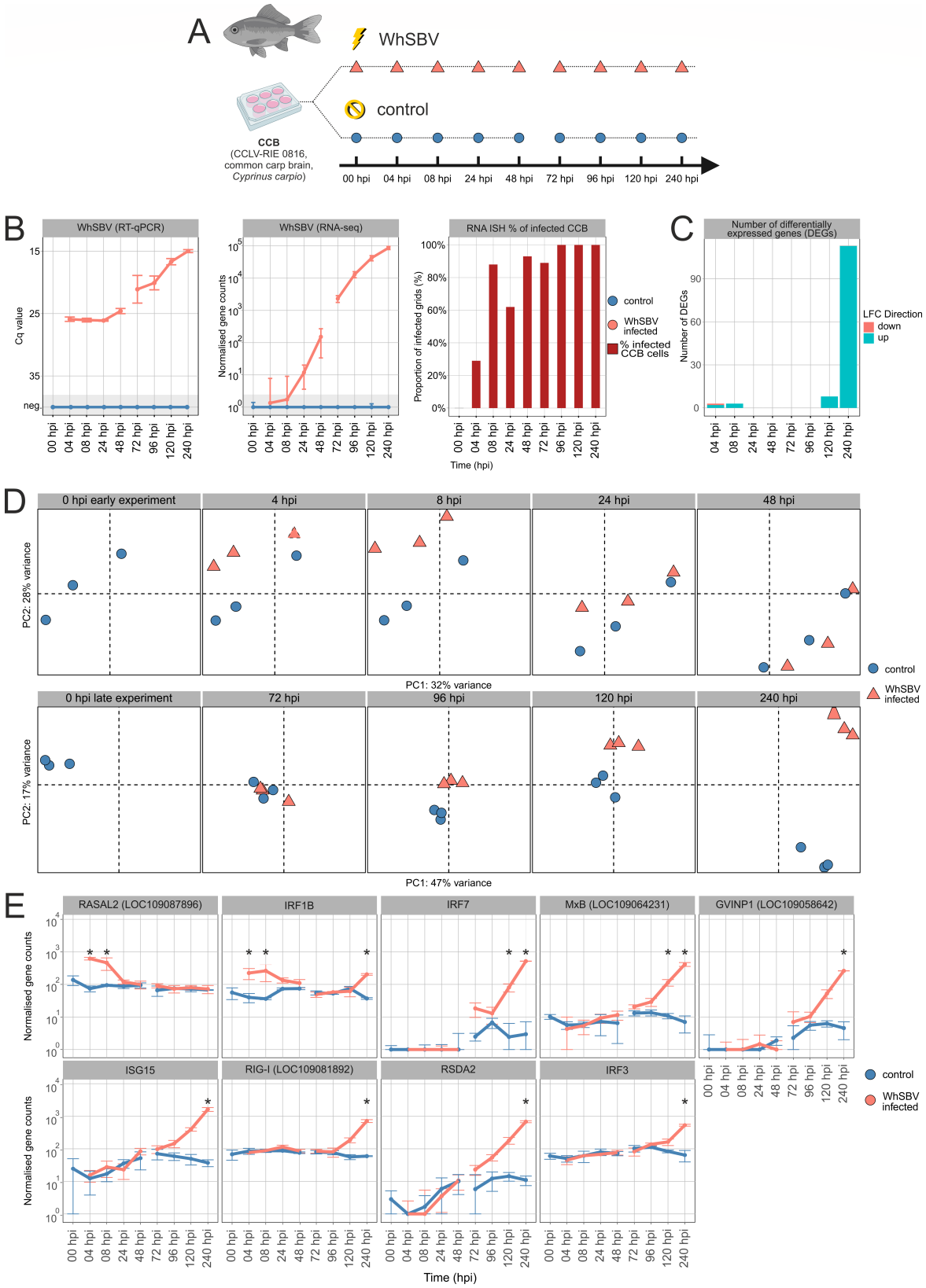

FIG 6 Cellular response to early and late WhSBV infection in the cypriniform brain cells CCB. (A) CCB cells were experimentally WhSBV-infected (orange triangles) and mock-infected (blue dots) in two consecutive experiments, covering early (0, 4, 8, 24, and 48 hours post infection (hpi)) and late (0, 72, 96, 120, and 240 hpi) infection stages. (B) Viral RNA was detected by reverse transcription quantitative PCR (RT-qPCR; left panel) and RNA sequencing (RNAseq; middle panel).

(Continued on next page)

Fig 6 (Continued)

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 normalized gene counts. Arithmetic means (± standard deviation, SD) of three replicate wells are shown for Cq values and normalized gene counts. WhSBV-specific RNA *in situ* hybridization (RNA ISH) data from a separate experiment are shown for comparison (right panel; see Fig. S4). (C) Differentially expressed genes (DEGs) with an absolute log2-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 RAS GTPase-activating protein 2-like (LOC109087896) (RSAL2), interferon regulatory factor 1 (IRFIB), interferon regulatory factor 7 (IRF7), interferon regulatory factor 3 (IRF3), interferon-induced GTP-binding protein MxB-like (MxB), interferon-induced very large GTPase 1-like (LOC109058642) (GVINP1), interferon-stimulated gene 15 (ISG15), antiviral innate immune response receptor RIG-I (LOC109081892) (RIG 1), and viperin, radical S-adenosyl methionine domain-containing 2 (RSDA2). Asterisks indicate time points with significantly differentially gene expression as compared to uninfected controls. The complete list of DEGs can be found in Table S3. Results are presented as the arithmetic/geometric mean (± standard deviation) of three replicate wells.

Detection of multicistronic WhSBV gene transcripts

The observed sequence coverage across the genome was uneven, with abrupt increases and decreases observed in certain potential intergenic regions (Fig. 7A). These coverage variations corresponded in part to transcription start and termination sites identified by motif prediction (see Table S4). In detail, three transcriptional start sites (S1-3) were located upstream of the N, X, and M ORFs. Start sites S1 and S3 were marked by an increase in read coverage, indicating transcription initiation. Four transcription termination sites were detected by motif prediction: T1, T2, and T4 were predicted downstream of the N, G, and L ORFs, respectively, while T3 was found within the L ORF. The termination sites T2 and T3 were characterized by a marked decrease in sequence coverage, and termination sites T1-3 coincided with reads transitioning to poly(A)-tails. S2 and S3 were located immediately adjacent to T1 and T2, respectively. Sequencing of 5′ ends of cDNA identified the likely transcription start site S1 at nucleotide positions 33 to 35 in the viral genome, marked by the sequence "GAA," matching the predicted sequence of S2 and S3. Regions of the genome with comparable levels of sequence coverage between adjacent start and termination sites were interpreted as belonging to the same viral RNA transcript or mRNA. We observed no coverage drop at termination signal T1, and the majority of transcripts seem to skip T1, indicating that N, X/P and G genes can be expressed from polycistronic mRNA (Fig. 7A). However, a small number of reads transitioning to poly(A)-tails at T1 indicate the presence of monocistronic N mRNA. The mRNAs containing the M and L genes shared S3 as the common start site, but their transcript levels differed, with the M gene being comparably highly expressed to the L gene (Fig. 7A). We did not detect a possible short upstream ORF (uORF) in the X-P-G transcript as observed for BoDV-1 (13).

A complementary Northern blot analysis of RNA from the persistently WhSBV-infected GCSB1441 and the experimentally WhSBV-infected EPC cells supported some findings from the sequence data analysis (Fig. 8A through C). All probes (N, X, P, G, M, or L ORF specific) detected the full genome of WhSBV from the GCSB1441 and EPC cells at the expected size of 9,008 nt. Probes X, P, and G probably labeled the same RNA species at approximately 2,100 nt, matching the predicted length for the polycistronic X-P-G mRNA. The M and L probes both labeled an RNA between 6 and 8 kb, likely representing the M-L t1 transcript. A faint band at 500 nt was detected using the M probe, but not with the L probe, which may represent the transcripts encoding the M ORF alone. The N probe identified an RNA species that was consistently higher than expected and was also recognized by the X, P, and G probes. This observation suggests the presence of an RNA spanning N, X, P, and G, while the monocistronic N mRNA was not detected at the expected size.

High-resolution mass spectrometry was performed on the persistently WhSBV-infected GCSB1441 cell suspension to provide an overview of the viral proteins. The analysis revealed peptides that matched all the predicted proteins of WhSBV, except for the L ORF

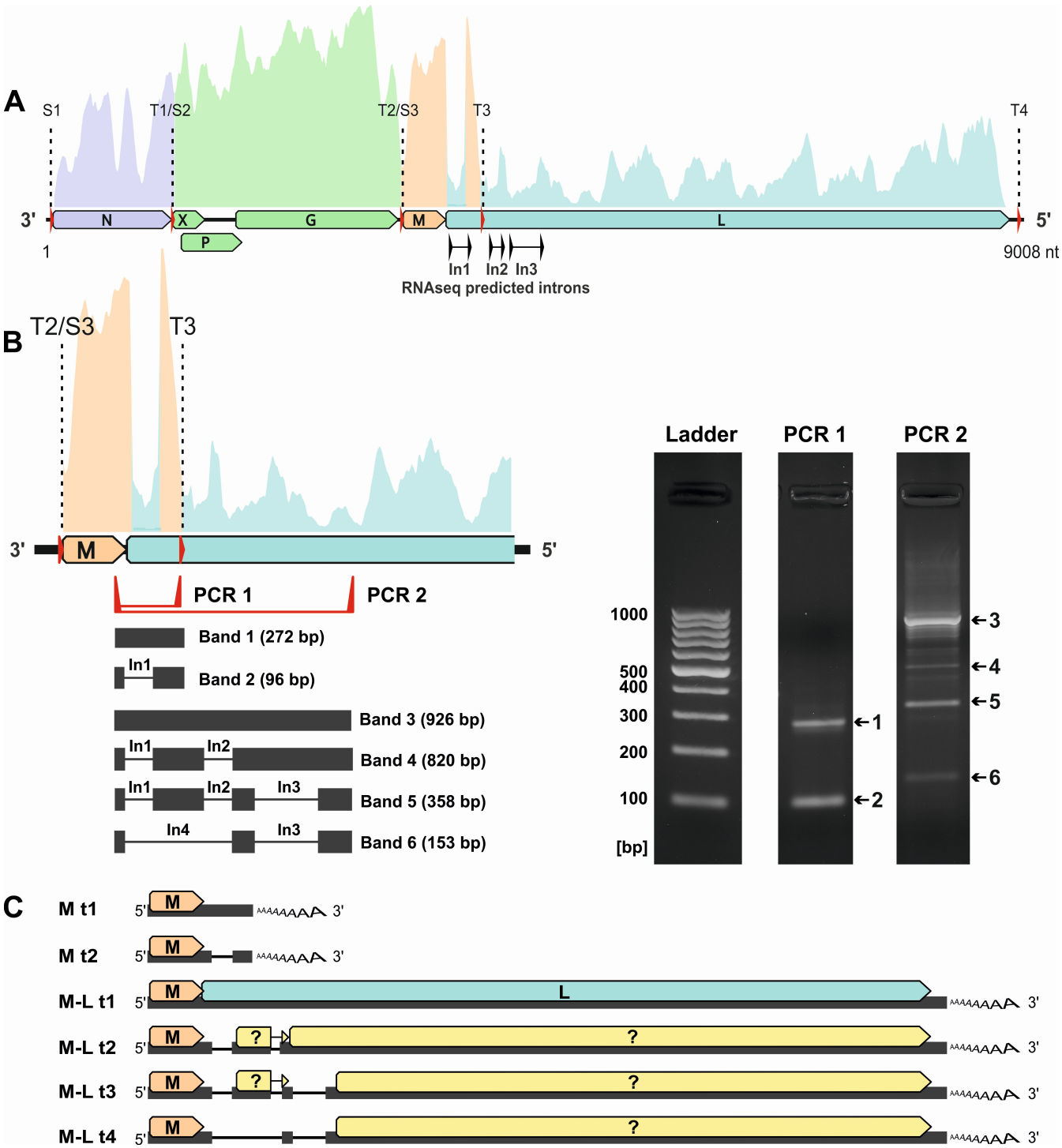

FIG 7 Detection of spliced WhSBV RNA in persistently WhSBV-infected GCSB1542 cells. (A) The trimmed raw reads from the metagenomic RNA data sets from persistently WhSBV-infected GCSB1441 and GCSB1542 cells were mapped to the WhSBV genome (PV171101) using "bbmap" (version 39.33). In the schematic WhSBV genome, open reading frames (ORFs) are shown as arrows and transcription start sites (S1–S3) and termination sites (T1-4) are indicated by red arrows and dashed lines. ORFs that are likely to be located on the same RNA transcript are shown in the same color. Potential introns (In1-3) were detected using STAR (version 2.7.11b) and are indicated by black arrows. (B) RNA was extracted from the persistently WhSBV-infected GCSB1542 cell line, and cDNA was synthesized using oligo-dT primers. A PCR was performed using this cDNA, and primers target 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.

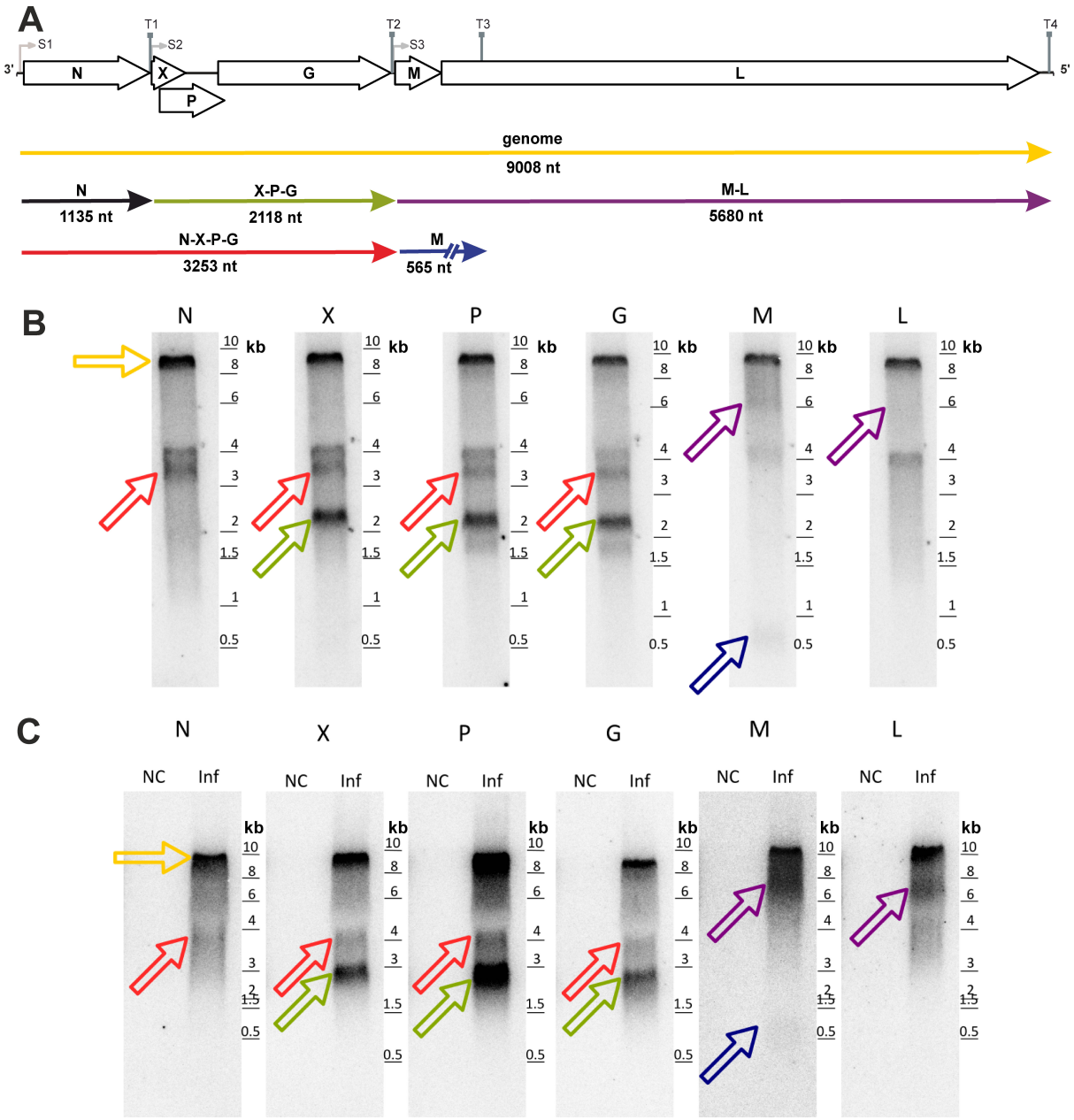

FIG 8 Detection of multicistronic WhSBV RNA in persistently WhSBV-infected and WhSBV-inoculated cells. WhSBV RNA transcripts were detected by Northern blot using ^{32}P-CTP-labeled probes specific for WhSBV genes (arrows) encoding for the nucleoprotein (N), accessory protein (X), phosphoprotein (P), matrix protein (M), glycoprotein (G), and the large protein (L). For reference, a schematic representation of the WhSBV genome, along with open reading frames (arrows) and possible RNA molecules (colored lines) is shown in (A). Total RNA from the persistently WhSBV-infected GCSB1441 cell line (B) and WhSBV-inoculated (Inf) EPC cells compared to non-WhSBV-inoculated (NC) EPC cells (C) was used for the Northern blot. Probes that detected RNA species of the same size were interpreted as corresponding to the same RNA molecule. The colored arrows in (B) and (C) correspond to the respective WhSBV RNA molecules in the schematic. A ladder is provided in kilobase pairs (kb).

where no matching peptides were detected. No fused peptides were detected either, suggesting no additional splicing (Fig. 9).

DISCUSSION

This study represents the first *in vitro* molecular characterization of WhSBV, a member of the family *Bornaviridae*, genus *Cultervirus* (3, 6), and thus the first biological

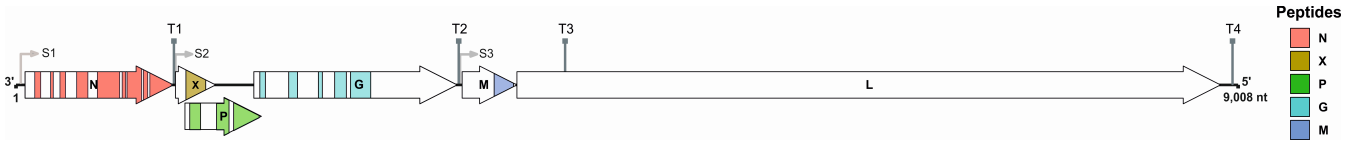

FIG 9 Detection of WhSBV viral peptides in persistently WhSBV-infected GCSB1441 cells. High-resolution mass spectrometry was performed on protein extracts from the persistently WhSBV-infected GCSB1441 cells. Mass spectrometry raw files were processed with MaxQuant (version 2.4.13.0) (14) using a custom virus database (six entries) and the *C. idella* database (C_idella_female_genemodels.v1; 32,811 entries) with default settings. The figure shows a schematic representation of the complete viral genome, with open reading frames (ORF) indicated by arrows and transcription start (S1–S3) and termination sites (T1-4). Areas corresponding to detected peptides are color-coded for each associated ORF. No peptides matching the L protein or fused peptides were detected.

characterization of a bornavirid not belonging to the well-studied genus *Orthobornavirus*. Initially, identified by RNA sequencing of sharpbelly tissues (gut, liver, and gill) (2, 6), WhSBV was later detected by data mining in grass carp cell lines CIK and L8824 (2, 6, 15). Here, we further identified WhSBV in two grass carp swim bladder-derived cell lines by metagenomic RNA sequencing.

The identification of WhSBV in established cell lines allowed us to study the virus *in vitro*, offering insights into WhSBV persistence, potential *in vitro* host range, and transcriptional strategies. Although the lack of WhSBV-specific antibodies and the absence of a cytopathic effect in cell culture limited the ability to conduct classic virus titration, we used RT-qPCR and RNA ISH instead. Other studies have shown a good correlation between viral titers and Cq values, enabling a semi-quantitative assessment of viral load to be made in the absence of virus titration (16–18). Our RT-qPCR assay targeting the G gene was the most reliable and sensitive method of detecting WhSBV RNA. This is consistent with the findings of Northern blot analysis and genome read coverage, which showed that the X-P-G coding RNA was expressed at higher levels than other viral RNAs.

Our comparative analysis across cypriniform cell lines revealed that WhSBV can establish a persistent infection without inducing cytopathic effects and maintains stable viral RNA levels over multiple passages (Fig. 2A). Similar behavior has been observed for BoDV-1, VSBV-1, and several avian orthobornaviruses, where some inoculated cell lines became persistently infected without cytopathic effect and could be sub-cultivated without loss of infectivity (19–26). As some orthobornaviruses show a relatively broad host spectrum *in vitro,* we tested WhSBV inoculation of non-cypriniform, reptile, avian, and mammalian cell lines.

Avian and mammalian cells were clearly not susceptible to WhSBV infection. This may be due to the absence of essential host factors in avian and mammalian cells required for WhSBV replication, the activation of an effective antiviral response, or the virus's potential intolerance to the elevated incubation temperature of 37°C. In some non-cypriniform fish (RTG-2/f, CHSE-214, and TiB) and reptilian cell lines, susceptibility to WhSBV infection was difficult to assess because ACTB levels declined over several passages, with eventual loss of cultures. This phenomenon was also observed in uninfected control cultures, suggesting the deterioration was likely due to suboptimal culture conditions, rather than WhSBV infection. Nevertheless, in these cell lines, viral RNA levels over the successive passages fluctuated near or below the detection Cq cut-off value. While these low-level signals may reflect residual inoculum, RNA ISH confirmed the presence of viral RNA within some cells for RTG-2/f and TiB, indicating that these cells were indeed infected by WhSBV. However, infection remained confined to individual cells, and there was no progressive increase in the percentage of infected cells, unlike in cypriniform cells where infection spread over time. As RNA ISH was not possible for reptilian cells, there is currently no conclusive evidence whether these are susceptible to WhSBV infection, and further investigations will be necessary.

The difference in WhSBV infectivity between cypriniform and non-cypriniform fish cell lines may be attributed, in part, to inherent cytological and physiological distinctions between these cells. The fish cell lines used in this study originate from various

tissues and primarily represent epithelial or fibroblast lineages. These cells may differ substantially in the membrane composition, receptor expression profiles, and basal innate immune activity, all of which can influence their permissiveness to viral entry, replication, and persistence (27). Furthermore, variation in metabolic activity, temperature tolerance, and stress responses can further modulate viral propagation *in vitro* (28–31). The restricted replication of WhSBV in non-cypriniform fish cells may also indicate an evolutionary history of adaptation to cypriniform hosts, which is supported by the absence of confirmed detections in other taxonomic groups to date (2, 6, 15).

For the production of virus preparations, active disruption of the infected cells by, e.g., freeze/thawing was necessary, while comparably little infectious virus appeared to be present in the supernatant of infected culture. This is in agreement with orthobornaviruses, such as BoDV-1 (32, 33). We observed virus particle-like structures and budding in WhSBV-inoculated ZF4 cells, which were absent in the respective controls. Although these structures cannot be definitively attributed to WhSBV, their morphology is consistent with that of viral particles. With 150–180 nm in diameter, their detected size exceeds the typical size range reported for BoDV-1 (70–130 nm) (34, 35) and avian bornaviruses (60–104 nm) (36–38). While the observed structures may represent viral particles and budding, it is currently not possible to confirm their identity as WhSBV without specific staining tools, such as immunogold labeling.

In order to investigate the cellular response to WhSBV infection and to gain insights into the possible mechanisms underlying its ability to establish a persistent infection without inducing cytopathic effects, we analyzed the host gene expression at different time points of infection. The innate antiviral immune system, including the interferon (IFN) pathway, is well conserved between mammals and teleost fish (39–41). In teleosts, as in higher vertebrates, viral nucleic acids are sensed by pattern recognition receptors (PRRs), notably RIG-I-like receptors such as RIG-I and MDA5 (42–44). Upon sensing, these receptors initiate signaling cascades that culminate in the production of type I IFNs and pro-inflammatory cytokines, which then upregulate ISGs to restrict viral replication and enhance viral clearance (44–48). Transcriptomic analysis of the experimentally WhSBV-infected CCB cells indicated a limited initial activation of the innate immune response shortly after inoculation at 4 and 8 hpi by upregulation of the genes IRF1B and RASAL2 (Fig. 6E). Interferon regulatory factors, such as IRF1B, are key regulators of the innate antiviral response in vertebrates, controlling the expression of type I IFNs and ISGs (49, 50). In zebrafish, IRF1B alone was shown to induce the expression of IFNs and ISGs, resulting in effective protection against viral infection (51). While the role of RASAL2 in fish remains unclear, in mammals, it has been characterized as a tumor and metastasis suppressor (52). Despite this initial immune activation, no significant changes in host cell gene expression were observed between 24 and 96 hpi (Fig. 6C), even as viral RNA levels and the proportion of infected cells increased steadily during this period (Fig. 6B). These findings suggest that WhSBV may evade or suppress early innate immune signaling through yet-unknown mechanisms, thereby enabling efficient viral replication and the establishment of persistent infection. In comparison, orthobornaviruses have been shown to evade IFN induction by possessing monophosphorylated genome ends (53, 54). BoDV-1 further suppresses the interferon response via P-protein-mediated inhibition of TANK-binding kinase 1 (55) and N-protein interference with IRF7 activation (56).

Only at 120 and 240 hpi did we observe the upregulation of RIG-I and IRFs other than IRF1B, which may have contributed to a delayed innate immune response at later stages. This response was marked by significant upregulation of many ISGs, such as MxB, GVINP1, ISG15, and RSDA2, all of which play antiviral roles in teleost fish (57). This suggests that the host's innate immune system did not mount a response until the infection had spread to nearly all cells. However, this late response appeared insufficient to clear the infection as the cells remained persistently infected, suggesting either the activation of a secondary immune response aimed to control viral load or that the increasing intracellular viral burden eventually triggered the innate immune system

of the host. Despite the upregulation of ISGs, no significant induction of apoptosis- or necrosis-related genes was observed throughout infection, suggesting that WhSBV establishes infection without inducing programmed cell death. This may contribute to continued viral replication and long-term or even persistent infection. In contrast, common carp hepatocytes exposed to ammonia stress undergo apoptosis via the P53-BAC/BCL-2 pathway, demonstrating the species' ability to induce cell death under certain conditions (58). The absence of such responses in WhSBV-infected cells raised the possibility that the virus actively suppresses apoptosis. Our study provides a temporal overview of host gene expression during a 10-day infection course in the cypriniform CCB cells. To fully elucidate the molecular strategies employed by WhSBV to evade and/or modulate the host immune response, future investigations should focus on identifying the specific viral genes, proteins, or genomic structures involved in immune suppression and persistence. The establishment of persistence without cytopathic effects is a common feature of orthobornaviruses (3) and has also been observed in other mononegaviruses, such as arenavirids (family *Arenaviridae*), which suppress host antiviral responses, particularly by inhibiting type I interferon production, to maintain chronic infections (59, 60).

One strategy to evade the host's immune response might be the strict regulation of WhSBVs transcription that involves expression gradients, polycistronic mRNAs, and splicing, as seen in some orthobornaviruses (11, 12, 61, 62). Some regulatory features, such as the splicing of M-L transcripts, and the position and sequence of transcription start and termination sites (11, 33), seem to be highly conserved among culterviruses and therefore suggest conserved mechanisms. This transcription strategy is in contrast to the predominantly monocistronic transcription seen in other mononegaviruses (63), such as paramyxovirids (family *Paramyxoviridae*), where the viral genome is transcribed into monocistronic unspliced mRNAs, each encoding a single protein (64). However, there are some exceptions, such as Ebola virus (family *Filoviridae*), that uses RNA editing of its glycoprotein gene in order to generate multiple variants, thereby optimizing replication and pathogenicity without increasing the genome size (65).

Northern blot and sequence coverage analysis indicated that the N proteins are expressed not only from a monocistronic mRNA but also co-expressed with X/P and G from a polycistronic mRNA. However, it remains to be clarified whether the detected RNA of approximately 3,000–4,000 nt in length is actually mRNA or a sub-genomic RNA. For BoDV-1, the non-coding region between the N and X genes shows structural similarities to that of WhSBV, where T1 is located downstream of the transcription start site S2 in BoDV-1 (66). The consensus motif of the BoDV-1 termination sites, $U(A)_{6\text{-}7}$ (67), serves as a core signal sequence but is insufficient for efficient transcription termination on its own, as the surrounding upstream and downstream nucleotides significantly influence T1 utilization. Consequently, when transcription is initiated at S1 and terminated at T1, these regulatory elements contribute to the production of a monocistronic (1,200 nt) mRNA encoding the N protein (66). However, studies have demonstrated that the viral RNA-directed RNA polymerase frequently bypasses the T1 transcription termination site, resulting in the production of a longer (1,900 nt) polycistronic mRNA containing the N, X, and P ORFs (66). The presence of such transcripts in BoDV-1-infected cell cultures and rat brains suggests that they are likely authentic mRNAs, rather than by-products of aberrant replication, and their production may be related to viral fitness (66). Conversely, other studies have suggested that these long mRNAs may represent aborted replication products or an extended form of "leader RNA," similar to those described in certain other negative-strand RNA viruses (10, 68).

In BoDV-1, the 5′ untranslated region of the X/P mRNA contains a short uORF that modulates ribosomal scanning and translation initiation, contributing to controlled protein expression (13). A comparable feature was not detected for WhSBV, suggesting that the translation of the multicistronic mRNAs is regulated by leaky scanning rather than ribosomal termination/reinitiation, as seen in other mononegaviruses (69). Future studies are needed to investigate the function of these multicistronic mRNAs, particularly

those co-expressing N with X/P and G, and to determine whether they encode proteins or contribute to viral fitness.

The relatively low expression level of the L gene observed in WhSBV may be explained by a combination of transcriptional and posttranscriptional mechanisms. One potential factor is the presence of a transcription termination signal (T3) located within the L ORF. In BoDV-1, the T3 signal mediates premature transcription termination of the L ORF, resulting in lower transcript abundance and consequently lower protein expression (10). A similar regulatory structure in WhSBV could result in early termination of L gene transcripts, as supported by our transcriptional profiling, which shows high coverage starting at S3 and ending at T3. Additionally, PCR and Sanger sequencing revealed the presence of introns within the L ORF, suggesting a splicing event that may disrupt the integrity of the L ORF. This could serve to downregulate L expression while allowing the expression of the overlapping M ORF. This phenomenon has also been reported in BoDV-1, where alternative splicing within the L gene results in expression of the overlapping G ORF (8, 11, 61). Although L transcripts were detected by Northern blot, proteomic analysis did not identify the L protein, further supporting its comparable low expression levels. Collectively, these findings suggest that the L gene in WhSBV may be tightly controlled to fine-tune L protein levels. Future studies using more sensitive proteomic techniques and functional assays are warranted to clarify whether L protein expression is truly low or simply undetectable under current conditions. Additionally, further investigation is needed to elucidate the functional relevance of the observed introns. Based on the current data, possible hypotheses include the following: facilitating increased expression of the M gene by truncating the L ORF, regulating L expression at low levels, or generating an alternative transcript encoding a yet-unidentified protein. A similar mechanism has been observed for vesicular stomatitis virus (VSV; family *Rhabdoviridae*), which controls L gene expression by polycistronic transcription, co-expressing it with other viral genes (70). Ultimately, the transcriptional regulation of WhSBV, including both polycistronic and spliced transcripts, appears to play a role in balancing viral protein levels and minimizing host cell stress, while optimizing viral replication and assembly.

Conclusion

This study provides a molecular characterization of WhSBV, highlighting its evolutionary adaptation to cypriniform hosts. The virus replicates efficiently in cypriniform cell cultures, establishing persistent infections without inducing any visible cell damage and suppressing the host's early innate immune responses. Its limited or absent replication in non-cypriniform cultures highlights the need for *in vivo* validation to determine its full host range. Additionally, this study sheds light on the transcriptional mechanisms of WhSBV, revealing its regulatory strategies for modulating gene expression. Further research is essential to unravel the molecular mechanisms underlying WhSBV persistence and to assess its ecological impact, particularly in aquaculture, where persistent viral infections could significantly impact fish health and disease management. Our findings also emphasize the importance of transitioning from *in silico* screening to *in vitro* validation—a critical step in bioinformatics-driven virology research. Although metagenomic approaches facilitate virus discovery, a key limitation is often the lack of direct validation from real virus samples.

MATERIALS AND METHODS

Nucleic acid extraction and metagenomic sequencing

A total of 48 fish cell lines were provided by the Collection of Cell Lines in Veterinary Medicine (CCLV) at the Friedrich-Loeffler-Institut, Riems, Germany (see Table S1 for a complete list of cell lines). Fresh-frozen cell lines were lysed in 1 mL TRIzol LS reagent (Life Technologies) and shaken vigorously for 10 min at room temperature. After the addition

of 200 µL chloroform (Carl Roth) and centrifuging at 13,000 × *g*, 10 min, 4°C, 200 µL of the upper aqueous phase was used for RNA extraction. Total RNA was extracted using the Agencourt RNAdvance Tissue Kit (Beckman Coulter) with a KingFisher Flex Purification System (Thermo Fisher Scientific) according to the manufacturer's instructions. DNase I digestion (Qiagen) was carried out prior to the final elution to remove genomic DNA residues. The quantity of total RNA was determined using the NanoDrop Lite Spectrophotometer (Thermo Fisher Scientific). Poly(A) RNA was isolated using the Dynabeads mRNA DIRECT Purification Kit (Invitrogen) following the manufacturer's instructions. Fragmentation and strand-specific construction of whole-transcriptome libraries was performed using the Collibri Stranded RNA Library Prep Kit for Illumina Systems (Invitrogen). The quality and integrity of the RNA and the final library were verified using the Agilent TapeStation 4150 (Agilent Technologies) with appropriate chips and reagents. The final libraries were quantified using the Qubit dsDNA HS Assay Kit (Invitrogen) in combination with the Qubit 2.0 Fluorometer (Invitrogen). Equimolar amounts of the library were then sequenced using Novaseq 6000 with SP flow cell in the single-end 101 bp mode (CeGaT) according to the manufacturer's instructions.

Metagenomic analysis

The raw reads obtained from the 48 fish cell lines were trimmed using TrimGalore (v0.6.10 [71]) and cutadapt (v4.0 [72]). The metagenomic pipeline SqueezeMeta (v1.6.4 [73]) was then used to individually assemble the trimmed reads from each data set and then sequentially merge them into a single transcriptome ("seqmerge" mode). Poly(A/T) sequences at the sequence termini of the assembled and merged transcripts were then removed using cutadapt (v4.0 [72]). The polished transcriptome was then used as input for taxonomic assignment and abundance estimation using SqueezeMeta (v1.6.4 [73], database version Sep-3-2023). The results were further analyzed using the R (v4.3.1 [74]) package SQMtools (v1.6.3 [73]).

Validation of WhSBV genome termini

Persistently WhSBV-infected grass carp swim bladder (GCSB) cell lines GCSB1441 (CCLV no. 1441) and GCSB1542 (1542; see Table S1) were grown in non-vented T25 tissue culture flasks (Corning) in a suitable growth medium supplemented with 10% fetal bovine serum (FBS) to 95%–100% confluence. Total RNA was extracted using the TRIzol-chloroform method, followed by purification using the RNeasy Mini Kit (Qiagen) and DNase I digestion prior to elution. The quality and quantity of the RNA were assessed using the Agilent TapeStation 4150 in combination with appropriate chips and reagents, and the NanoDrop Lite spectrophotometer.

The 5′-Rapid Amplification of cDNA Ends (RACE) kit, version 2.0 (Invitrogen), was used to determine the 3' end of the WhSBV genome. Reverse transcription (RT) was performed on the extracted RNA using a gene-specific primer WhSBV-945-R (0.4 µM) (see Table S5 for a complete list of primers). The resulting cDNA was then digested with RNase cocktail (Invitrogen) at 37°C for 20 min and subsequently purified using AMPure XP beads (Beckman Coulter). A tailing reaction was then performed using dCTP and dATP (New England Biolabs). A second round of amplification was performed using gene-specific primers WhSBV-777-R (0.2 µM) and the AAP primer (0.2 µM) for purified C-tailed cDNA or the adapter primer (AP) (0.2 µM) for purified A-tailed cDNA (see Table S5). The hemi-nested PCR was performed with gene-specific primers WhSBV-545-R (0.2 µM) and the AUAP primer (0.2 µM) (see Table S5). To identify the 5' end, total RNA samples were treated with the *E. coli* Poly(A) Polymerase Kit (New England Biolabs) according to the manufacturer's instructions in order to add poly(A) tails. The poly(A) tailed RNA was then purified using AMPure XP beads. This purified poly(A) tailed RNA was used as input to the 3′-RACE system for rapid amplification of cDNA ends (Invitrogen), using the AP primer (see Table S5) for RT. The cDNA was then digested with an RNase cocktail and purified using AMPure XP beads. For the second round of amplification, six different gene-specific primers (WhSBV_7388F, WhSBV_7528F, WhSBV_7911F, WhSBV_8215F, WhSBV_8737F,

and WhSBV_8901F; 0.2 µm) were combined with the AUAP primer (see Table S5), for individual reactions. The resulting PCR amplicons were separated by 2% agarose gel electrophoresis. The remaining product was purified with AMPure XP beads and sequenced using the BigDye Terminator v1.1 Cycle Sequencing Kit on a 3500 Genetic Analyzer (Applied Biosystems). The final WhSBV genome sequences were uploaded to GenBank under accession number PV171101.

RT-qPCR and qPCR assays for specific detection of WhSBV

For the following assays, persistently WhSBV-infected GCSB1441 and GCSB1542 cells and the WhSBV-uninfected GK (0119) cell line (see Table S1) were grown to 95%–100% confluence in non-vented T25 tissue culture flasks using appropriate growth medium supplemented with 10% FBS. Cell pellets were harvested using Alsever's trypsin-versus-versus-solution (ATV; FLI, Riems) after discarding the supernatant, and RNA/DNA was extracted using the NucleoMag VET kit (Macherey-Nagel) according to the manufacturer's instructions.

WhSBV-specific RT-qPCR assays were established by designing primers and probe sets targeting the viral N, P, G, M, and L genes (see Table S5 for a complete list of primers and probes). WhSBV-specific RNA (and potentially DNA) was detected using the AgPath-ID one-step RT-PCR reagents (Thermo Fisher Scientific). Briefly, 5 µL of extracted RNA was reverse-transcribed and amplified in a reaction mix of 25 µL total volume containing the specific primer/probe mix (final primer and probe concentration: 0.4 and 0.2 µM). For the internal control, primers ACT-1030-F (final concentration: 0.2 µM) and ACT-1135-R (0.2 µM) along with probe ACT-1081-HEX (0.2 µM) were included in each reaction to target the cellular actin beta (ACTB) gene (75). The reaction was performed on a Bio-Rad CFX96 qPCR cycler (Bio-Rad) using the following cycling conditions: 45°C for 10 min, 95°C for 10 min, 42 cycles of 95°C for 15 s, 57°C for 20 s and 72°C for 30 s. The efficiency and limit of detection of the WhSBV G gene RT-qPCR assay were tested using serial dilutions. A Cq cut-off value of 37 was set as the limit of detection for all performed assays.

To exclude the presence of WhSBV sequences as endogenous viral elements (EVE) integrated into the cellular genome of GCSB1441, GCSB1542, and GK, we specifically tested extracted DNA from these cells using a non-reverse transcription qPCR system. WhSBV cDNA from the 3′-RACE experiment, as previously described, was used as a positive control. The Maxima Probe/ROX qPCR Master Mix (Thermo Fisher Scientific) was used according to the manufacturer's protocol, targeting the viral genes N, P, G, M, and L (final primer and probe concentrations: 0.4 and 0.2 µM) and the cellular ACTB gene (final primer and probe concentrations: 0.2 and 0.2 µM). The reaction was performed on a Bio-Rad CFX96 qPCR cycler using the following cycling conditions: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s.

Optimization of WhSBV inoculum preparation

Persistently WhSBV-infected GCSB1441 cells were maintained in non-vented T75 culture flasks (75 cm^2; Corning) at 26°C until reaching 95%–100% confluency in growth medium supplemented with 10% FBS. Five different methods were evaluated to prepare the WhSBV inoculum in order to determine the most efficient and reliable protocol for subsequent experiments: (i) 5 mL of the supernatant (equivalent to 1 mL per 15 cm^2 of infected cell layer) was retained, and the cells were subjected to three freeze/thaw cycles freezing at −80°C and thawing at room temperature for 30 minutes each. (ii) After three freeze/thaw cycles as described above, the resulting cell suspension was collected in a Falcon tube (Sarstedt), sonicated (Branson Sonifier 450), and centrifuged at 1,200 × *g* for 20 minutes at 4°C, and only the supernatant was collected. (iii) Cell lines were suspended in 5 mL supernatant, which was subsequently sonicated and centrifuged as described above, and only the supernatant was collected. (iv) The culture medium was replaced with 5 mL hypertonic medium containing 150 mM NaCl. After 2 h of incubation at 26°C, the cells were scraped using a cell scraper, and the cell suspension was centrifuged as

previously described, and only the supernatant was collected. (v) 5 mL of the untreated supernatant was collected. All inocula were stored at −80°C until further use.

In order to test the infectivity of the different inocula, CCB cells (see Table S2) were seeded in 12-well plates (3.8 cm^2 per well; Corning) and grown to 95%–100% confluency in the appropriate medium supplemented with 10% FBS at 26°C. Inoculation was performed at an infection ratio of 1:0.8, corresponding to the ratio between the area of the cell layer from which the inoculum originated and the area of the cell layer inoculated. Specifically, 300 µL of the medium was replaced by 300 µL of WhSBV inoculum. Infections were conducted in three independent replicate wells for each time point. Cell pellets were harvested at 0, 4, 8, 24, 48, 72, 96, 120, 144, 168, 192, 216, and 240 hpi. At each time point, supernatants were discarded, and cell pellets were collected by trypsinization using ATV. RNA was extracted using the NucleoMag VET kit as described previously. The extracted nucleic acids were used for RT-qPCR analysis, as described above, using specific primer/probe mixes for WhSBV G and ACTB. A Delta-Delta Cq (ΔΔCq) analysis was conducted using R (version 4.3.1; [74) to compare different inoculation methods, with ACTB serving as the reference gene and method (i) as the reference method. To ensure consistency, a single batch of inoculum prepared by method (i) was used throughout most assays.

Experimental infection of different cell cultures with WhSBV

Bony fish cells RT/F (CCLV no. 0088), CaPi (0112), EPC (0173), RTG-2/f (0686), CCB (0816), SAF-1 (0826), CHSE-214 (1104), GTS-9 (1388), ZF4 (1492), and TiB (1550), the reptilian cells SKL-R (0484), VH-2 (1092), and CDSK (1536), the avian cells QM7 (0466) and DF-1 (1529), and the mammalian cells Vero (0015) and C6 (1452) were obtained from CCLV (see Table S2 for a complete list of cell lines). All fish and reptile cell lines were grown in non-vented T25 flasks (25 cm^2; Corning) to 95%–100% confluence in an appropriate 10 mL growth medium supplemented with 10% FBS. Incubation temperatures were set at 20°C for the fish cells CHSE-214 and RTG-2/f, 26°C for all other fish cells and reptilian cell SKL-R, and 28°C for the reptile cells VH-2 and CDSK. Mammalian and avian cells were grown in vented T25 flasks at 37°C in a humidified atmosphere containing 5% CO_2.

To initiate the infection experiment, cells were grown to confluence in T25 flasks, as described above. Prior to inoculation (T0), 1 mL of the supernatant was collected from each cell culture. The cells were then inoculated with 1 mL of the WhSBV inoculum, prepared using the optimized protocol involving three freeze/thaw cycles, as previously described. The inoculum was applied at a ratio of 1:1.7, as described above. Cultures were then incubated for 1 h under the above conditions. After this incubation, a further 1 mL of the supernatant was collected (T1). Cells were then incubated and split at regular intervals of 2–3 days at a 2-fold splitting factor for a total of 4–6 passages (P1-6). The splitting process involved collecting the supernatant before each passage, washing and detachment of the cells with ATV, and collecting half of the detached cells for analysis. The remaining cells were used for further culture by adding fresh growth medium supplemented with 10% FBS. The collected supernatants and cell pellets were stored at −80°C for subsequent RNA extraction and RT-qPCR analysis. Negative controls were treated in the same way as infected samples, except that the WhSBV inoculum was replaced by 1 mL RNase-free water (Qiagen). The infection experiment was performed in two independent runs for each cell line. As the infection of reptile, avian, and mammalian cell lines was performed separately from the comparison of different fish cell lines, the cypriniform cell line GTS-9 was used as a positive control to validate the infectivity of the inoculum.

Nucleic acid extraction from supernatant and cell pellet samples was performed using the NucleoMag VET kit according to the manufacturer's instructions. RT-qPCR was then performed using WhSBV G gene and ACTB specific RT-qPCR primer/probe mixes, as described earlier.

Electron microscopy

WhSBV-infected and uninfected ZF4 cells (see Table S2) from the previous infection experiment and persistently WhSBV-infected GCSB1441 cells were transferred to non-vented T75 flasks for electron microscopy analysis. Cells were fixed in 2.5% glutaraldehyde buffered in 0.1 M sodium cacodylate pH 7.2 (SERVA Electrophoresis), and 1% aqueous OsO_4 was used for post-fixation and 2.5% uranyl acetate in ethanol for *en bloc* staining (SERVA Electrophoresis). After a graded dehydration in ethanol, the samples were cleared in propylene oxide and infiltrated with Glycid Ether 100 (SERVA Electrophoresis). For polymerization, samples were incubated at 60°C for 3 days. Ultrathin sections were transferred to formvar-coated nickel grids (slot grids; Plano). All grids were counterstained with uranyl acetate and lead citrate before examination on a Talos F200i transmission electron microscope (FEI) at an accelerating voltage of 80 kV.

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

The cypriniform cell lines (EPC and CCB), the non-cypriniform fish cells (RTG-2/f and TiB), the mammalian Vero line, and the avian DF-1 cells (see Table S2) were maintained in non-vented T75 flasks in a suitable medium supplemented with 10% FBS until they reached 95%–100% confluence. Cultures were maintained under the conditions described in the previous infection experiment. For infection, 3 mL of the medium was replaced with an equal volume of the inoculum prepared via three freeze/thaw cycles, to achieve an infection ratio of 1:1.7, as previously described. For cypriniform and non-cypriniform fish, cells were harvested from two independent experiments at early and late time points: 0, 4, 8, 24, and 48 hpi for the early time points and 0, 72, 96, 120, and 240 hpi for the late time points. In a separate experiment, the Vero and DF-1 cell lines were harvested exclusively at 0 and 240 hpi, alongside the CCB cell line, which served as a positive cypriniform control. All cells were harvested using a cell scraper. The harvested cells were collected in 50 mL Falcon tubes, centrifuged at 3,500 rpm for 15 minutes, and the supernatants discarded. The cell pellets were then fixed in 10% neutral buffered formalin (Carl Roth) for at least 24 h. After paraffin embedding, 4-µm-thick sections were processed for subsequent RNA *in situ* hybridization (RNA ISH) using RNAScope 2-5 HD Reagent Kit Red and custom-designed probes targeting WhSBV (-)RNA (genomic RNA; WhSBV-L-O2, catalog number 1567161-C1, Advanced Cell Diagnostics, Hayward, USA) on all samples according to the manufacturer's instructions. In addition, custom-designed probes targeting WhSBV (+)RNA (mRNA and antigenomic RNA; WhSBV-G-O1, catalog number 1567171-C1) were used on CCB samples. As technical controls, probes against the genes for peptidylprolyl isomerase B (PPIB) and dihydrodipicolinate reductase (DapB) were included in each run. In addition, routine staining using hematoxylin-eosin was performed on all samples taken at 0 and 240 hpi for histopathologic evaluation. All slides were scanned using a Hamamatsu S60 scanner and evaluated using NDPview.2 plus software (version 2.8.24, Hamamatsu Photonics, K.K. Japan) by a board-certified pathologist (AB, DiplECVP). The abundance of infected cells was recorded as a percentage in a 10 × 10 grid (size 100 × 100 µm per grid cell). Thus, a positive signal within a grid cell resulted in one percentage point.

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

Gene expression analysis of early and late WhSBV infection

The CCB cells (see Table S2) were used to study the effect of WhSBV on cellular gene expression. Cells were seeded in 6-well plates (9.6 cm^2 per well; Corning) and cultured for 24 hours to 95%–100% confluence in 10% FBS supplemented medium. They were maintained under the conditions described in the previous kinetics experiment. Infection with WhSBV was performed at a 1:1.3 ratio as described above by replacing 500 µL of the culture medium in each well with 500 µL of WhSBV inoculum prepared via three freeze/thaw cycles. Negative controls were inoculated with 500 µL of the same medium used for persistently WhSBV-infected GCSB1441 cultivation, replacing 500 µL of the cell culture medium. Cell pellets were collected in two independent experiments to assess early and late infection stages. In the early infection experiment, mock-infected and WhSBV-infected cell pellets were collected at 4, 8, 24, and 48 hpi. For the late-stage experiment, samples were collected at 72, 96, 120, and 240 hpi. In addition, pre-infection cell pellets were collected at 7 and 24 hours before inoculation, the latter representing the 0 hpi time point. Infection was performed in three independent replicate wells for each time point. Cell pellets were collected by trypsinization after supernatants were discarded. Total RNA was extracted using Trizol, chloroform, and the Agencourt RNAdvance Tissue Kit, as previously described. RT-qPCR was performed on all samples using G-gene and ACTB-specific RT-qPCR primer/probe mixes to validate samples prior to sequencing. Transcriptomic libraries were prepared as previously described, but including ERCC spike-in mix (Invitrogen) as an internal control. Subsequently, the libraries were pooled and sent for sequencing on an Illumina NovaSeq 6000 system (Novogene GmbH) running in the paired-end 150 bp mode. Raw reads were subsequently trimmed using TrimGalore (v0.6.10; [71]) and then analyzed using the nf-core "RNAseq" pipeline (version 3.14.0; [76, 77]). In detail, the common carp reference genome GCF_018340385.1 was used for splice-aware mapping with STAR and subsequent quantification with Salmon. The quantification data were then further analyzed in R (version 4.3.1; [74]) and differentially expressed genes (DEGs) were deduced using the DESeq2 package (v1.40; [78]). Log2-fold changes were shrunk using the function "lfcShrink" with the "apeglm" shrinkage estimator (79). The cut-off values for DEGs were set to absolute shrinked log2-fold change > 1 and adjusted *P*-value < 0.05. To explore global gene expression patterns, PCA was performed using R (version 4.3.1; [74]). Transcriptomic data have been deposited in ArrayExpress under accession E-MTAB-14894.

Prediction of regulatory and splice sites in the WhSBV genome

The trimmed raw reads from the metagenomic RNA data sets from persistently WhSBV-infected GCSB1441 and GCSB1542 cells (see Table S1) were mapped back to the RACE corrected WhSBV genome (PV171101) using the "bbmap" function from BBTools (version 39.33 [80]). Duplicates were removed from mapped reads using the "MarkDuplicates" function from the Picard toolkit (version 2.20.4 [81]), and reads were separated by mapping orientation using Samtools (version 1.22.1 [82]). The genome coverage for forward and reverse deduplicated reads was obtained using Bedtools (version 2.31.1 [83]) and was plotted using R (version 4.3.1 [74]). Poly(A)-addition sites were deduced using the same data set and reference with ContextMap (version 2.7.9 [84]) and bowtie (version 1.3.1 [85]). Potential introns were detected using STAR (version 2.7.11b [86]) running in the "2-pass" mode, enabling sensitive novel junction discovery. The function "FIMO" from the MEME Suite (v5.5.2 [87]) was used to identify the conserved motif "AKUUAAYAAAAACAU**GAA**" (2) of potential transcription termination and start sites.

Validation of predicted splice sites in the WhSBV genome

To experimentally validate each predicted splice site, primers were designed to amplify the specific genomic regions by RT-PCR. Purified RNA was extracted by using the RNeasy Mini kit as described above. cDNA from poly(A) RNA was synthesized using the Protoscript II First Strand cDNA Synthesis Kit and oligo-dT primers (New England

Biolabs). The cDNA products were then digested with an RNase cocktail and purified using AMPure XP beads. Conventional Taq-PCR was performed using the Accuprime *Taq* DNA Polymerase System (Invitrogen) with primers specific for the predicted splice sites within the viral M and L genes (see Table S5). RNase-free water was used instead of cDNA as a negative control. The amplified DNA products were run on a 2% agarose gel, and bands were excised. The DNA was extracted using a QIAquick Gel Extraction Kit (Qiagen) and then sequenced using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) on a 3500 Genetic Analyzer (Applied Biosystems).

Specific detection of viral RNA using Northern blot

RNA was extracted from the persistently WhSBV-infected GCSB1441, WhSBV-infected, and uninfected EPC cells (see Tables S1 and S2) from the previous infection experiment using the TRIzol-chloroform method and the RNeasy Mini Kit with DNase I (Qiagen), as previously described. RNA quality and quantity were assessed using the TapeStation System 4150 (Agilent Technologies) and NanoDrop Lite spectrophotometer, as described previously. Subsequently, 2–4 µg of total RNA was glyoxylated using Glyoxal (Sigma-Aldrich) at 56°C for 45 min before separation on a 0.9% denaturing agarose gel (4.7% formaldehyde, Carl Roth) in a phosphate buffer system. RNA was transferred overnight to Hybond-N membranes (Cytiva Amersham) using SSC transfer buffer. RNA was cross-linked to the membrane, and viral RNA (mRNA and genomic RNA) was detected using DNA probes. Probes were generated by RT-PCR using specific forward and reverse primers against each gene of WhSBV (see Table S5) using the Platinum Script III One-step RT-PCR Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The amplified cDNA products were then run on a 2% agarose gel. After amplification, the products were purified using AMPure XP beads according to the manufacturer's protocol. The quantity of purified products was measured using a NanoDrop Lite spectrophotometer and labeled with ^{32}P-CTP (Hartmann-Analytic) using a Nick translation kit (GE Healthcare). Hybridization was performed overnight at 60°C–65°C, and the signal was analyzed after washing using the CR35 phosphoimager system (Dürr Medical).

Specific detection of WhSBV proteins

Approximately 3 × 10^5 persistently WhSBV-infected GCSB1441 cells were harvested, resuspended in 1 × LDS Buffer (Thermo Fisher Scientific) containing 100 mM DTT (Sigma-Aldrich), and heated for 10 min at 70°C shaking at 1,400 rpm in a Thermomixer (Eppendorf). Reconstitution buffer (50 mM HEPES/KOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, pH 8.0, 0.5% [wt/vol] Sodium deoxycholate, 1% [vol/vol] Triton X-100) was added to a final volume of 100 µL. Samples were reduced in 10 mM DTT for 30 min at 60°C, followed by alkylation in 25 mM iodoacetamide (Sigma-Aldrich) for 25 min at RT in the dark with orbital shaking at 800 rpm in a Thermomixer (Eppendorf). The reaction was quenched by incubation with 10 mM DTT f.c. for 5 min at RT. For protein binding, 125 µg hydrophilic and 125 µg hydrophobic Sera-Mag Carboxylate-Modified Magnetic Particles (GE Healthcare) together with 150 µL absolute ethanol (Carl Roth) were added to each sample and incubated for 15 min at RT at 1000 rpm in a Thermomixer (Eppendorf). Beads were washed three times with 80% LC-/MS-grade ethanol (Supelco). Residual ethanol was removed prior to adding 1 µg LC-grade trypsin (Serva) in 100 µL of 50 mM ammonium bicarbonate buffer pH 8. After incubation for 4 h at 37°C and 1,000 rpm in a Thermomixer (Eppendorf), peptides were collected and desalted on StageTips (88). The amount of 400 ng digested peptides was analyzed with a nanoElute2 HPLC system (Bruker) coupled to a TimsTOF HT mass spectrometer (Bruker). Peptides were separated on an Ultimate CSI 25 × 75 C18 UHPLC column (Ionopticks) by running an optimized 100 min gradient of 2 to 95% MS-grade acetonitrile with 0.1% formic acid at 250 nl/min at 50°C. The mass spectrometer was operated with the manufacturer-provided DDA_PASEF_1.1 sec_cycletime method. Mass spectrometry raw files were processed with MaxQuant (version 2.4.13.0) (14) using a custom virus database (six entries) and

the *C. idella* database (C_idella_female_genemodels.v1; 32,811 entries) with standard settings.

ACKNOWLEDGMENTS

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, as well as Mathias Lenk from the Collection of Cell Lines in Veterinary Medicine (CCLV), and the contributions of Robin Brandt and Kristin Vorpahl.

AUTHOR AFFILIATIONS

[1]Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Greifswald, Germany
[2]Department of Experimental Animal Facilities and Biorisk Management, Friedrich-Loeffler-Institut, Greifswald, Germany
[3]Institute of Infectology, Friedrich-Loeffler-Institut, Greifswald, Germany
[4]Institute of Molecular Virology and Cell Biology, Friedrich-Loeffler-Institut, Greifswald, Germany

AUTHOR ORCIDs

Mirette I. Y. Eshak http://orcid.org/0009-0004-3254-8422
Angele Breithaupt https://orcid.org/0000-0002-6373-5923
Birke A. Tews http://orcid.org/0000-0002-8647-0497
Christine Luttermann http://orcid.org/0000-0002-6977-9900
Kati Franzke http://orcid.org/0000-0003-0827-5407
Sören Woelke http://orcid.org/0009-0006-4196-9709
Martin Beer http://orcid.org/0000-0002-0598-5254
Dennis Rubbenstroth http://orcid.org/0000-0002-8209-6274
Florian Pfaff http://orcid.org/0000-0003-0178-6183

FUNDING

Funder	Grant(s)	Author(s)
BMEL | Friedrich-Loeffler-Institut (FLI)	FLI-IVD-XX-2021-83	Florian Pfaff

AUTHOR CONTRIBUTIONS

Mirette I. Y. Eshak, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review and editing | Florian Pfaff, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing.

DATA AVAILABILITY

All sequencing data associated with this study have been deposited in public repositories and are freely accessible. The WhSBV 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 GCSB1441 are available in the Sequence Read Archive (SRA) under SRR32424572, and RNAseq data have been deposited at ArrayExpress under E-MTAB-14894.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental figures (JVI01322-25-s0001.docx). Figures S1 to S6.
Table S1 (JVI01322-25-s0002.xlsx). Summary of cell lines used for metagenomic RNA sequencing.
Table S2 (JVI01322-25-s0003.xlsx). Complete list of cell lines used for infection experiments in this study.
Table S3 (JVI01322-25-s0004.xlsx). Summary of differentially expressed genes.
Table S4 (JVI01322-25-s0005.xlsx). Transcription start and termination motifs of WhSBV.
Table S5 (JVI01322-25-s0006.xlsx). Primers and probes used in this study.

REFERENCES

1. Kuhn JH, Dürrwald R, Bào Y, Briese T, Carbone K, Clawson AN, deRisi JL, Garten W, Jahrling PB, Kolodziejek J, Rubbenstroth D, Schwemmle M, Stenglein M, Tomonaga K, Weissenböck H, Nowotny N. 2015. Taxonomic reorganization of the family *Bornaviridae*. Arch Virol 160:621–632. https://doi.org/10.1007/s00705-014-2276-z
2. Eshak MIY, Rubbenstroth D, Beer M, Pfaff F. 2023. Diving deep into fish bornaviruses: uncovering hidden diversity and transcriptional strategies through comprehensive data mining. Virus Evol 9:vead062. https://doi.org/10.1093/ve/vead062
3. Rubbenstroth D, Briese T, Dürrwald R, Horie (堀江真行) M, Hyndman TH, Kuhn JH, Nowotny N, Payne S, Stenglein MD, Tomonaga (朝長啓造) K, ICTV Report Consortium. 2021. ICTV Virus Taxonomy Profile: Bornaviridae. J Gen Virol 102. https://doi.org/10.1099/jgv.0.001613
4. Niller HH, Angstwurm K, Rubbenstroth D, Schlottau K, Ebinger A, Giese S, Wunderlich S, Banas B, Forth LF, Hoffmann D, et al. 2020. Zoonotic spillover infections with Borna disease virus 1 leading to fatal human encephalitis, 1999-2019: an epidemiological investigation. Lancet Infect Dis 20:467–477. https://doi.org/10.1016/S1473-3099(19)30546-8
5. Hoffmann B, Tappe D, Höper D, Herden C, Boldt A, Mawrin C, Niederstraßer O, Müller T, Jenckel M, van der Grinten E, Lutter C, Abendroth B, Teifke JP, Cadar D, Schmidt-Chanasit J, Ulrich RG, Beer M. 2015. A variegated squirrel bornavirus associated with fatal human encephalitis. N Engl J Med 373:154–162. https://doi.org/10.1056/NEJMoa1415627
6. Shi M, Lin X-D, Chen X, Tian J-H, Chen L-J, Li K, Wang W, Eden J-S, Shen J-J, Liu L, Holmes EC, Zhang Y-Z. 2018. The evolutionary history of vertebrate RNA viruses. Nature 556:197–202. https://doi.org/10.1038/s41586-018-0012-7
7. Hyndman TH, Shilton CM, Stenglein MD, Wellehan JFX. 2018. Divergent bornaviruses from Australian carpet pythons with neurological disease date the origin of extant *Bornaviridae* prior to the end-cretaceous extinction. PLoS Pathog 14:e1006881. https://doi.org/10.1371/journal.ppat.1006881
8. Briese T, Schneemann A, Lewis AJ, Park YS, Kim S, Ludwig H, Lipkin WI. 1994. Genomic organization of Borna disease virus. Proc Natl Acad Sci USA 91:4362–4366. https://doi.org/10.1073/pnas.91.10.4362
9. Briese T, de la Torre JC, Lewis A, Ludwig H, Lipkin WI. 1992. Borna disease virus, a negative-strand RNA virus, transcribes in the nucleus of infected cells. Proc Natl Acad Sci USA 89:11486–11489. https://doi.org/10.1073/pnas.89.23.11486
10. Schneemann A, Schneider PA, Kim S, Lipkin WI. 1994. Identification of signal sequences that control transcription of borna disease virus, a nonsegmented, negative-strand RNA virus. J Virol 68:6514–6522. https://doi.org/10.1128/JVI.68.10.6514-6522.1994
11. Schneider PA, Schneemann A, Lipkin WI. 1994. RNA splicing in Borna disease virus, a nonsegmented, negative-strand RNA virus. J Virol 68:5007–5012. https://doi.org/10.1128/JVI.68.8.5007-5012.1994
12. Tomonaga K, Kobayashi T, Lee BJ, Watanabe M, Kamitani W, Ikuta K. 2000. Identification of alternative splicing and negative splicing activity of a nonsegmented negative-strand RNA virus, Borna disease virus. Proc Natl Acad Sci USA 97:12788–12793. https://doi.org/10.1073/pnas.97.23.12788
13. Watanabe Y, Ohtaki N, Hayashi Y, Ikuta K, Tomonaga K. 2009. Autogenous translational regulation of the Borna disease virus negative control factor X from polycistronic mRNA using host RNA helicases. PLoS Pathog 5:e1000654. https://doi.org/10.1371/journal.ppat.1000654
14. Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. https://doi.org/10.1038/nbt.1511
15. Costa VA, Mifsud JCO, Gilligan D, Williamson JE, Holmes EC, Geoghegan JL. 2021. Metagenomic sequencing reveals a lack of virus exchange between native and invasive freshwater fish across the Murray-Darling Basin, Australia. Virus Evol 7:veab034. https://doi.org/10.1093/ve/veab034
16. Brandolini M, Taddei F, Marino MM, Grumiro L, Scalcione A, Turba ME, Gentilini F, Fantini M, Zannoli S, Dirani G, Sambri V. 2021. Correlating qRT-PCR, dPCR and viral titration for the identification and quantification of SARS-CoV-2: a new approach for infection management. Viruses 13:1022. https://doi.org/10.3390/v13061022
17. Hitchman RB, Siaterli EA, Nixon CP, King LA. 2007. Quantitative real-time PCR for rapid and accurate titration of recombinant baculovirus particles. Biotechnol Bioeng 96:810–814. https://doi.org/10.1002/bit.21177
18. Falsey AR, Formica MA, Treanor JJ, Walsh EE. 2003. Comparison of quantitative reverse transcription-PCR to viral culture for assessment of respiratory syncytial virus shedding. J Clin Microbiol 41:4160–4165. https://doi.org/10.1128/JCM.41.9.4160-4165.2003
19. Herzog S, Rott R. 1980. Replication of Borna disease virus in cell cultures. Med Microbiol Immunol 168:153–158. https://doi.org/10.1007/BF02122849
20. Rinder M, Ackermann A, Kempf H, Kaspers B, Korbel R, Staeheli P. 2009. Broad tissue and cell tropism of avian bornavirus in parrots with proventricular dilatation disease. J Virol 83:5401–5407. https://doi.org/10.1128/JVI.00133-09
21. Pham PH, Leacy A, Deng L, Nagy É, Susta L. 2020. Isolation of Ontario aquatic bird bornavirus 1 and characterization of its replication in immortalized avian cell lines. Virol J 17:16. https://doi.org/10.1186/s12985-020-1286-6
22. 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*). Vet Microbiol 161:36–42. https://doi.org/10.1016/j.vetmic.2012.07.004
23. Rubbenstroth D, Schmidt V, Rinder M, Legler M, Corman VM, Staeheli P. 2014. Discovery of a new avian bornavirus genotype in estrildid finches (Estrildidae) in Germany. Vet Microbiol 168:318–323. https://doi.org/10.1016/j.vetmic.2013.11.032
24. Rubbenstroth D, Rinder M, Stein M, Höper D, Kaspers B, Brosinski K, Horie M, Schmidt V, Legler M, Korbel R, Staeheli P. 2013. Avian bornaviruses are widely distributed in canary birds (*Serinus canaria f. domestica*). Vet Microbiol 165:287–295. https://doi.org/10.1016/j.vetmic.2013.03.024
25. Schlottau K, Nobach D, Herden C, Finke S, Beer M, Hoffmann D. 2020. First isolation, *in-vivo* and genomic characterization of zoonotic variegated squirrel Bornavirus 1 (VSBV-1) isolates. Emerg Microbes Infect 9:2474–2484. https://doi.org/10.1080/22221751.2020.1847604
26. Nowotny N, Kolodziejek J, Jehle CO, Suchy A, Staeheli P, Schwemmle M. 2000. Isolation and characterization of a new subtype of Borna disease virus. J Virol 74:5655–5658. https://doi.org/10.1128/jvi.74.12.5655-5658.2000
27. Davidovich M, Dishon A, Ilouze M, Kotler M. 2007. Susceptibility of cyprinid cultured cells to cyprinid herpesvirus 3. Arch Virol 152:1541–1546. https://doi.org/10.1007/s00705-007-0975-4
28. Marsella A, Buratin A, Pascoli F, Abbadi M, Toson M, Cuenca A, Toffan A, Vendramin N. 2025. Temperature impact on replication and virulence of

European infectious hematopoietic necrosis viruses. Aquaculture 609:742786. https://doi.org/10.1016/j.aquaculture.2025.742786
29. Li M, Yang J, Ye C, Bian P, Yang X, Zhang H, Luo C, Xue Z, Lei Y, Lian J. 2021. Integrated metabolomics and transcriptomics analyses reveal metabolic landscape in neuronal cells during JEV infection. Virol Sin 36:1554–1565. https://doi.org/10.1007/s12250-021-00445-0
30. Barton BA, Iwama GK. 1991. Physiological changes in fish from stress in aquaculture with emphasis on the response and effects of corticosteroids. Annu Rev Fish Dis 1:3–26. https://doi.org/10.1016/0959-8030(91)90019-G
31. Ramnani B, Powell S, Shetty AG, Manivannan P, Hibbard BR, Leaman DW, Malathi K. 2023. Viral hemorrhagic septicemia virus activates integrated stress response pathway and induces stress granules to regulate virus replication. Viruses 15:466. https://doi.org/10.3390/v15020466
32. Matsumoto Y, Hayashi Y, Omori H, Honda T, Daito T, Horie M, Ikuta K, Fujino K, Nakamura S, Schneider U, Chase G, Yoshimori T, Schwemmle M, Tomonaga K. 2012. Bornavirus closely associates and segregates with host chromosomes to ensure persistent intranuclear infection. Cell Host Microbe 11:492–503. https://doi.org/10.1016/j.chom.2012.04.009
33. Tomonaga K, Kobayashi T, Ikuta K. 2002. Molecular and cellular biology of Borna disease virus infection. Microbes Infect 4:491–500. https://doi.org/10.1016/s1286-4579(02)01564-2
34. Zimmermann W, Breter H, Rudolph M, Ludwig H. 1994. Borna disease virus: immunoelectron microscopic characterization of cell-free virus and further information about the genome. J Virol 68:6755–6758. https://doi.org/10.1128/JVI.68.10.6755-6758.1994
35. Kohno T, Goto T, Takasaki T, Morita C, Nakaya T, Ikuta K, Kurane I, Sano K, Nakai M. 1999. Fine structure and morphogenesis of Borna disease virus. J Virol 73:760–766. https://doi.org/10.1128/JVI.73.1.760-766.1999
36. Wünschmann A, Honkavuori K, Briese T, Lipkin WI, Shivers J, Armien AG. 2011. Antigen tissue distribution of Avian bornavirus (ABV) in psittacine birds with natural spontaneous proventricular dilatation disease and ABV genotype 1 infection. J Vet Diagn Invest 23:716–726. https://doi.org/10.1177/1040638711408279
37. Hoppes S, Gray PL, Payne S, Shivaprasad HL, Tizard I. 2010. The isolation, pathogenesis, diagnosis, transmission, and control of avian bornavirus and proventricular dilatation disease. Vet Clin North Am Exot Anim Pract 13:495–508. https://doi.org/10.1016/j.cvex.2010.05.014
38. Gancz AY, Kistler AL, Greninger AL, Farnoushi Y, Mechani S, Perl S, Berkowitz A, Perez N, Clubb S, DeRisi JL, Ganem D, Lublin A. 2009. Experimental induction of proventricular dilatation disease in cockatiels (*Nymphicus hollandicus*) inoculated with brain homogenates containing avian bornavirus 4. Virol J 6:100. https://doi.org/10.1186/1743-422X-6-100
39. Altmann SM, Mellon MT, Distel DL, Kim CH. 2003. Molecular and functional analysis of an interferon gene from the zebrafish, *Danio rerio*. J Virol 77:1992–2002. https://doi.org/10.1128/jvi.77.3.1992-2002.2003
40. Schultz U, Kaspers B, Staeheli P. 2004. The interferon system of non-mammalian vertebrates. Dev Comp Immunol 28:499–508. https://doi.org/10.1016/j.dci.2003.09.009
41. Zou J, Tafalla C, Truckle J, Secombes CJ. 2007. Identification of a second group of type I IFNs in fish sheds light on IFN evolution in vertebrates. J Immunol 179:3859–3871. https://doi.org/10.4049/jimmunol.179.6.3859
42. Takeuchi O, Akira S. 2010. Pattern recognition receptors and inflammation. Cell 140:805–820. https://doi.org/10.1016/j.cell.2010.01.022
43. Rao Y, Su J. 2015. Insights into the antiviral immunity against grass carp (*Ctenopharyngodon idella*) reovirus (GCRV) in grass carp. J Immunol Res 2015:670437. https://doi.org/10.1155/2015/670437
44. Chang M, Collet B, Nie P, Lester K, Campbell S, Secombes CJ, Zou J. 2011. Expression and functional characterization of the RIG-I-like receptors MDA5 and LGP2 in Rainbow trout (*Oncorhynchus mykiss*). J Virol 85:8403–8412. https://doi.org/10.1128/JVI.00445-10
45. Zou J, Bird S, Secombes C. 2010. Antiviral sensing in teleost fish. Curr Pharm Des 16:4185–4193. https://doi.org/10.2174/138161210794519093
46. Biacchesi S, LeBerre M, Lamoureux A, Louise Y, Lauret E, Boudinot P, Brémont M. 2009. Mitochondrial antiviral signaling protein plays a major role in induction of the fish innate immune response against RNA and DNA viruses. J Virol 83:7815–7827. https://doi.org/10.1128/JVI.00404-09
47. Verrier ER, Langevin C, Benmansour A, Boudinot P. 2011. Early antiviral response and virus-induced genes in fish. Dev Comp Immunol 35:1204–1214. https://doi.org/10.1016/j.dci.2011.03.012
48. Chang M-X, Zou J, Nie P, Huang B, Yu Z, Collet B, Secombes CJ. 2013. Intracellular interferons in fish: a unique means to combat viral infection. PLoS Pathog 9:e1003736. https://doi.org/10.1371/journal.ppat.1003736
49. Nehyba J, Hrdlicková R, Bose HR. 2009. Dynamic evolution of immune system regulators: the history of the interferon regulatory factor family. Mol Biol Evol 26:2539–2550. https://doi.org/10.1093/molbev/msp167
50. Negishi H, Taniguchi T, Yanai H. 2018. The interferon (IFN) class of cytokines and the IFN regulatory factor (IRF) transcription factor family. Cold Spring Harb Perspect Biol 10:a028423. https://doi.org/10.1101/cshperspect.a028423
51. Feng H, Zhang Y-B, Zhang Q-M, Li Z, Zhang Q-Y, Gui J-F. 2015. Zebrafish IRF1 regulates IFN antiviral response through binding to IFNϕ1 and IFNϕ3 promoters downstream of MyD88 signaling. J Immunol 194:1225–1238. https://doi.org/10.4049/jimmunol.1402415
52. McLaughlin SK, Olsen SN, Dake B, De Raedt T, Lim E, Bronson RT, Beroukhim R, Polyak K, Brown M, Kuperwasser C, Cichowski K. 2013. The RasGAP gene, RASAL2, is a tumor and metastasis suppressor. Cancer Cell 24:365–378. https://doi.org/10.1016/j.ccr.2013.08.004
53. Habjan M, Andersson I, Klingström J, Schümann M, Martin A, Zimmermann P, Wagner V, Pichlmair A, Schneider U, Mühlberger E, Mirazimi A, Weber F. 2008. Processing of genome 5′ termini as a strategy of negative-strand RNA viruses to avoid RIG-I-dependent interferon induction. PLoS One 3:e2032. https://doi.org/10.1371/journal.pone.0002032
54. Reuter A, Ackermann A, Kothlow S, Rinder M, Kaspers B, Staeheli P. 2010. Avian bornaviruses escape recognition by the innate immune system. Viruses 2:927–938. https://doi.org/10.3390/v2040927
55. Unterstab G, Ludwig S, Anton A, Planz O, Dauber B, Krappmann D, Heins G, Ehrhardt C, Wolff T. 2005. Viral targeting of the interferon-β-inducing Traf family member-associated NF-κB activator (TANK)-binding kinase-1. Proc Natl Acad Sci USA 102:13640–13645. https://doi.org/10.1073/pnas.0502883102
56. Song W, Kao W, Zhai A, Qian J, Li Y, Zhang Q, Zhao H, Hu Y, Li H, Zhang F. 2013. Borna disease virus nucleoprotein inhibits type I interferon induction through the interferon regulatory factor 7 pathway. Biochem Biophys Res Commun 438:619–623. https://doi.org/10.1016/j.bbrc.2013.08.006
57. Ortega-Villaizan M del M, Chico V, Perez L. 2022. Fish innate immune response to viral infection-an overview of five major antiviral genes. Viruses 14:1546. https://doi.org/10.3390/v14071546
58. Xue S, Lin J, Han Y, Han Y. 2021. Ammonia stress–induced apoptosis by p53–BAX/BCL-2 signal pathway in hepatopancreas of common carp (*Cyprinus carpio*). Aquacult Int 29:1895–1907. https://doi.org/10.1007/s10499-021-00724-3
59. Meyer B, Ly H. 2016. Inhibition of innate immune responses is key to pathogenesis by arenaviruses. J Virol 90:3810–3818. https://doi.org/10.1128/JVI.03049-15
60. Kane M, Zang TM, Rihn SJ, Zhang F, Kueck T, Alim M, Schoggins J, Rice CM, Wilson SJ, Bieniasz PD. 2016. Identification of interferon-stimulated genes with antiretroviral activity. Cell Host Microbe 20:392–405. https://doi.org/10.1016/j.chom.2016.08.005
61. Cubitt B, Oldstone C, Valcarcel J, Carlos de la Torre J. 1994. RNA splicing contributes to the generation of mature mRNAs of Borna disease virus, a non-segmented negative strand RNA virus. Virus Res 34:69–79. https://doi.org/10.1016/0168-1702(94)90120-1
62. Cubitt B, Ly C, de la Torre JC. 2001. Identification and characterization of a new intron in Borna disease virus. J Gen Virol 82:641–646. https://doi.org/10.1099/0022-1317-82-3-641
63. Amarasinghe GK, Ayllón MA, Bào Y, Basler CF, Bavari S, Blasdell KR, Briese T, Brown PA, Bukreyev A, Balkema-Buschmann A, et al. 2019. Taxonomy of the order Mononegavirales: update 2019. Arch Virol 164:1967–1980. https://doi.org/10.1007/s00705-019-04247-4
64. Wignall-Fleming EB, Hughes DJ, Vattipally S, Modha S, Goodbourn S, Davison AJ, Randall RE. 2019. Analysis of paramyxovirus transcription and replication by high-throughput sequencing. J Virol 93:e00571-19. https://doi.org/10.1128/JVI.00571-19
65. Yu D-S, Weng T-H, Wu X-X, Wang FXC, Lu X-Y, Wu H-B, Wu N-P, Li L-J, Yao H-P. 2017. The lifecycle of the Ebola virus in host cells. Oncotarget 8:55750–55759. https://doi.org/10.18632/oncotarget.18498
66. Poenisch M, Wille S, Staeheli P, Schneider U. 2008. Polymerase read-through at the first transcription termination site contributes to regulation of borna disease virus gene expression. J Virol 82:9537–9545. https://doi.org/10.1128/JVI.00639-08

67. Schneemann A, Schneider PA, Lamb RA, Lipkin WI. 1995. The remarkable coding strategy of borna disease virus: a new member of the nonsegmented negative strand RNA viruses. Virology (Auckl) 210:1–8. https://doi.org/10.1006/viro.1995.1311
68. Banerjee AK, Barik S, De BP. 1991. Gene expression of nonsegmented negative strand RNA viruses. Pharmacol Ther 51:47–70. https://doi.org/10.1016/0163-7258(91)90041-j
69. Firth AE, Brierley I. 2012. Non-canonical translation in RNA viruses. J Gen Virol 93:1385–1409. https://doi.org/10.1099/vir.0.042499-0
70. Barr JN, Whelan SPJ, Wertz GW. 2002. Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. Biochim Biophys Acta 1577:337–353. https://doi.org/10.1016/s0167-4781(02)00462-1
71. Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G. 2023. FelixKrueger/TrimGalore: v0.6.10 - add default decompression path. Zenodo.
72. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10. https://doi.org/10.14806/ej.17.1.200
73. Tamames J, Puente-Sánchez F. 2018. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. Front Microbiol 9:3349. https://doi.org/10.3389/fmicb.2018.03349
74. R Core Team. 2024. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.
75. Wernike K, Hoffmann B, Kalthoff D, König P, Beer M. 2011. Development and validation of a triplex real-time PCR assay for the rapid detection and differentiation of wild-type and glycoprotein E-deleted vaccine strains of Bovine herpesvirus type 1. J Virol Methods 174:77–84. https://doi.org/10.1016/j.jviromet.2011.03.028
76. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. 2020. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol 38:276–278. https://doi.org/10.1038/s41587-020-0439-x
77. Patel H, Ewels P, Manning J, Garcia MU, Peltzer A, Hammarén R, Botvinnik O, Talbot A, Syme R, Kelly G, et al. 2024. nf-core/rnaseq: nf-core/rnaseq v3.16.0 - Further Fire Ferret. Zenodo.
78. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8
79. Zhu A, Ibrahim JG, Love MI. 2019. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics 35:2084–2092. https://doi.org/10.1093/bioinformatics/bty895
80. Bushnell B, Rood J, Singer E. 2017. BBMerge - accurate paired shotgun read merging via overlap. PLoS One 12:e0185056. https://doi.org/10.1371/journal.pone.0185056
81. Picard toolkit. 2019. https://broadinstitute.github.io/picard/.
82. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008. https://doi.org/10.1093/gigascience/giab008
83. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033
84. Bonfert T, Kirner E, Csaba G, Zimmer R, Friedel CC. 2015. ContextMap 2: fast and accurate context-based RNA-seq mapping. BMC Bioinformatics 16:122. https://doi.org/10.1186/s12859-015-0557-5
85. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25
86. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. https://doi.org/10.1093/bioinformatics/bts635
87. Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36.
88. Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2:1896–1906. https://doi.org/10.1038/nprot.2007.261