Development of a rapid and cost-effective multiplex PCR-Nanopore sequencing assay for accurate diagnosis of four tilapia pathogens

Jérôme Delamare-Deboutteville[1*#], Watcharachai Meemetta[2#], Khaettareeya Pimsannil[2], Han Ming Gan[3], Laura Khor Li Imm[1], Chadag Vishnumurthy Mohan[1], Ha Thanh Dong[4], Saengchan Senapin[2,5*]

[1]WorldFish, Penang, Malaysia
[2]Fish Health Platform, Center of Excellence for Shrimp Molecular Biology and Biotechnology (Centex Shrimp), Faculty of Science, Mahidol University, Bangkok, Thailand
[3]Patriot Biotech Sdn Bhd, Bandar Sunway, 47500, Selangor, Malaysia
[4]School of Environment, Resources and Development, Asian Institute of Technology, Pathum Thani, 12120, Thailand
[5]National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani, Thailand

*Corresponding authors:
J. Delamare-Deboutteville, j.delamare@cgiar.org
S. Senapin, saengchan@biotec.or.th
 # these authors contributed equally to this work.

Abstract

Tilapia aquaculture faces significant threats posed by four prominent pathogens: tilapia lake virus (TiLV), infectious spleen and kidney necrosis virus (ISKNV), *Francisella orientalis*, and *Streptococcus agalactiae*. Currently, employed molecular diagnostic methods for these pathogens rely on multiple singleplex PCR reactions, which are both time-consuming and expensive. In this study, we present a pioneering approach utilizing a novel multiplex PCR (mPCR) assay, coupled with rapid Nanopore sequencing, enabling for the one-tube simultaneous detection and one-reaction Nanopore sequencing-based identification of all four pathogens. Our one-tube multiplex assay exhibits a detection limit of 1,000 copies per reaction for TiLV, ISKNV, and *S. agalactiae*, while for *F. orientalis*, the detection limit is 10,000 copies per reaction. This capability allows for the detection of single infections as well as co-infections in clinical samples within a single day. Moreover, the consensus sequences generated from the amplicons of each sample demonstrate 100% sequence identity with publicly available data, providing strong support for the improving accuracy of Nanopore sequencing. The integration of multiplex PCR and Nanopore sequencing provides a promising and cost-effective platform for rapid and precise diagnostics of major tilapia pathogens, making it a valuable tool for enhancing health management practices within the aquaculture industry, ultimately contributing to improved disease control and prevention.

Introduction

Tilapia (*Oreochromis* spp.) is one of the most widely farmed freshwater fish species globally due to its high adaptability, fast growth, and excellent meat quality. Global production is estimated at 6,100,719 tonnes in 2020 [1]. By providing sustenance, employment opportunities, and domestic and export revenues, this valuable species supports large populations worldwide [2]. In the past decade, tilapia production has almost doubled [3], largely attributed to its ease of cultivation, market demand, and stable pricing [2].

However, the production of tilapia has been threatened by several viral and bacterial diseases that can
cause significant economic losses [4]. Among the most important tilapia pathogens are tilapia lake virus
(TiLV), infectious spleen and kidney necrosis virus (ISKNV), *Francisella orientalis*, and *Streptococcus*
agalactiae (SAG) also called group B *streptococcus* (GBS) [5–9]. Effective diagnosis and monitoring of these
pathogens are crucial for disease management and control [10]. Traditional diagnostic techniques such as
virus and bacterial isolation, immunofluorescence assay (IFA), and enzyme-linked immunosorbent assay
(ELISA) are time-consuming, labor-intensive, and require specialized equipment and expertise [10]. In
contrast, molecular methods such as polymerase chain reaction (PCR) and quantitative PCR (qPCR) have
gained widespread use in pathogen detection due to their high sensitivity and specificity, rapid turnaround
time, and ability to detect multiple pathogens in a single reaction [10–17].

In current real-world situations, it is uncommon to perform subsequent sequencing of the positive PCR or
qPCR products to validate the pathogen identity due to the time and resources involved in conventional
Sanger sequencing. Sequence-based verification is particularly valuable in multiplex PCR since non-
specific amplification is more likely to occur in a multiplex PCR reaction due to the presence of multiple
primer combinations. Fortunately, this is now possible with Oxford nanopore technology (ONT) that allows
on-site amplicon sequencing [18]. Proper integration of multiplex PCR and ONT-based genotyping may
therefore present a promising tool for the cost-effective, efficient, and precise detection and characterization
of pathogens in tilapia aquaculture.

Here, we present a novel multiplex PCR assay capable of detecting TiLV, ISKNV, *F. orientalis*, and *S.*
agalactiae simultaneously in tilapia samples. By leveraging the high-throughput sequencing capabilities of
the nanopore platform, we were able to obtain sequence-level confirmation of the targeted pathogens and
gather valuable genetic information about the amplicons. Our assay provides a rapid, cost effective, and
feasible solution for the identification and characterization of multiple pathogens in a single reaction, with
potential applications in disease diagnosis, targeted surveillance, and control.

Methods

Ethics declarations

70 The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page,
71 have been adhered to. No ethical approval was required as no animals were used in this study. Virus
72 sequences were generated from archived samples.

Clinical samples and nucleic acid extraction

We utilized archival clinical samples of fry, fingerling, juvenile and adult Nile tilapia, red tilapia, and Asian
sea bass from challenge experiments or from various disease outbreaks between 2015 and 2020. Our
investigation included samples that were either confirmed to be caused by a single pathogen using PCR
diagnosis or suspected to have resulted from co-infections. All sample details are comprehensively listed
in Table 1. To extract the nucleic acid from the samples, some were processed using the PathoGen-spin
DNA/RNA extraction kit from iNtRON Biotechnology, while others were archival RNA samples extracted
using Trizol reagent from Invitrogen, and DNA samples extracted using the conventional phenol/chloroform
ethanol precipitation method as described by [19]. If done close to the farm, the diagnostic workflow from the
point of sample collection to the final data analysis can take less than 12 hours, the process of which
includes nucleic acid extraction, multiplex PCR, library preparation, Nanopore sequencing, and data
analysis, which is illustrated in Figure 1.
86

Primers used in this study

We obtained the primers for the four selected target pathogens from Bio Basic (Canada), and the primer sequences are summarized in Supplemental Table 1. Primers for TiLV, ISKNV, *F. orientalis*, and *S. agalactiae* were reported in previous studies [16,20–24]. Expected amplicon sizes were 137, 190, 203, and 351 bp, respectively. Specificity of each primer pair was assessed *in silico* using the Primer-BLAST program (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

Plasmid positive controls and the analytical sensitivity assay

All plasmid positive controls except for *S. agalactiae* (constructed in this study) were obtained from our previous studies (Supplemental Table 2). Positive plasmid control containing *S. agalactiae groEL* partial gene was obtained by cloning a 351 bp-*groEL* amplified fragment purified before being ligated into pGEM-T easy vector (Promega). A recombinant plasmid was subjected to DNA sequencing (Macrogen). Copy number of each plasmid was calculated based on its size in base pair (bp) and amount in nanogram (ng) using a web tool at http://cels.uri.edu/gsc/cndna.html. For the multiplex PCR sensitivity assay, a combination of four recombinant plasmids was mixed. This mixture was then subjected to 10-fold serial dilution, resulting in a range of 1 to 10^6 copies/µl. Subsequently, 4 µl of the diluted series was used in the multiplex PCR reaction. To simulate a clinical sample from a fish, each PCR reaction was spiked with 50 ng of RNA extracted from a healthy tilapia.

Multiplex PCR condition optimization and the detection of clinical samples

Previous studies claimed that qPCR reagents usually contain PCR additives and enhancers that can improve amplification efficiency [30]. To optimize our multiplex PCR, we used KAPA SYBR FAST One-Step qRT-PCR master mix (Roche), which is known to contain such additives. Each reaction was prepared in a 25 µl volume, comprising 1X master mix, 4 µl of template, and 80-240 nM of each primer pair (Supplemental Table 3). We determined the optimal annealing temperature (Ta) using gradient PCR with Ta ranging between 55-65 °C (Supplemental Figure 1). To identify the best combination of PCR components, we added ammonium sulfate, BSA, dNTPs, and MgCl2 in different proportions. The final cycling conditions comprised a reverse transcription step at 42 °C for 5 min, followed by inactivation at 95 °C for 3 min, and 40 cycles of 95°C for 10 s and 60°C for 30 sec, with a final extension step at 60 °C for 5 min (Supplemental Table 3). The thermocycle machine used was Biometra TOne (Analytik Jena). We then analyzed 10 µl of each product by electrophoresis on a 3.5% agarose gel stained with ethidium bromide. The newly optimized multiplex PCR assay was subsequently used to detect the presence of TiLV, ISKNV, *F. orientalis*, and *S. agalactiae* in archival clinical samples (Table 1). Most of these samples were previously tested for a single pathogen using PCR or qPCR assays.

Quantification of pathogens by qPCR assays

A total of 15 clinical samples, whose mPCR products underwent Nanopore sequencing, were analyzed to determine the presence and quantity of each pathogen. The established protocols [16,21,22] were followed to detect TiLV, ISKNV, and SAG, as described in Supplemental Table 4. For FnO qPCR, the same primers used in mPCR were utilized in a SYBR Green-based qPCR assay conducted in this study.

Nanopore sequencing

The present study employed amplicons generated from both single and multiplex PCR reactions (Table 1, samples highlighted in grey) as templates for library preparation using the ligation sequencing kit (SQK-LSK109) and the native barcoding expansion 1-12 kit (EXP-NBD104) according to the standard protocols of Oxford Nanopore Technologies (ONT) adapted for the Flongle flow cell. A total of three sequencing runs (r1-r3) were conducted, with 250 ng of PCR product per sample and a unique native barcode (BC) assigned to each sample. Run 1 (r1) involved the sequencing of amplicons obtained from individual PCR reactions as well as a combination of these amplicons (Table 1). Run 2 (r2) focused on the sequencing of multiplex PCR (mPCR) products derived from clinical samples. Run 3 (r3) employed mPCR products obtained from the remaining clinical samples. Additionally, different concentrations of single PCR products were included in this run. The library of pooled barcoded samples was subjected to a Short Fragment Buffer (SFB) wash prior to the final elution step of the protocol. The DNA concentration was quantified at every step using the Qubit assay. Subsequently, the prepared library was loaded onto a Flongle flow cell (FLO-FLG106), following the Nanopore standard protocol, and each Flongle flow cell was fitted to a Flongle adapter (FLGIntSP) for MinION.

Reads filtering and read abundance calculation

Basecalled reads in fastq format were primer-trimmed with cutadapt v4.3 [31] and reads that have been trimmed with length ranging from 75-400 bp will be retained for the subsequent analysis, leaving out overly short or long reads without intact primer sequence on both ends. The filtered reads were aligned to the 4 pathogen gene segments using Minimap2 v2.17 [32]. Reads that aligned uniquely (only single hit reported) with more than 80% coverage to the target region were used for abundance calculation and variant calling (consensus generation). Calculation of raw, trimmed and aligned read statistics used seqkit v2.20 [33]. Reads failing to align at this stringent level were extracted and re-aligned using the default Minimap2 setting, followed by a less stringent blastN alignment. The host genome was included in the reference sequence to assess the fraction of reads mapping to the host.

Generation of consensus and variant sequences

Only reads with unique alignments were used as the template for Minimap2 and medaka variant calling based on the ARTIC pipeline [34,35]. The reads were mapped to the reference gene segments of four pathogens, followed by variant calling using the medaka variant model r941_min_g507. Variant filtering was then performed based on mapping quality and read depth. The number of ambiguous bases in the consensus sequences was calculated using QUAST v5 [36], and only sequences without any ambiguous base were selected for subsequent variant sequence generation. Variants were generated using cd-hit v.4.8.1 [37], whereby sequences with 100% nucleotide identity and exact sequence length were clustered into the same genotype. The variant sequences were compared against the NCBI Blast nt database (accessed on 25 April 2023) to identify the top 5 BLAST hits for each sequence.

Code availability

The Linux scripts used to generate initial FastQ files, assembled amplicons (public and from this study) are publicly available in the Zenodo.org dataset https://doi.org/10.5281/zenodo.7866295.

Results

Primer testing by single PCR

To validate the selected primer pairs for each pathogen, single PCR reactions were performed using each primer pair with its respective target template. The amplification of the target regions was confirmed by the presence of amplicons of the expected sizes, as illustrated in Supplemental Figure 2. To simulate the expected products from multiplex PCR (mPCR), an equal amount of each individual PCR product was loaded into lane C of the Supplemental Figure 2. Gel electrophoresis analysis revealed distinct and separable bands, demonstrating the potential for further utilization in the multiplex PCR analysis. Note that the products generated from this step were subsequently employed in Nanopore sequence run 1 (r1).

Multiplex PCR detection of four pathogens among clinical samples

We have successfully optimized a multiplex PCR assay for simultaneous detection of four important tilapia pathogens, TiLV, ISKNV, *F. orientalis*, and *S. agalactiae*, in a single reaction. Detection sensitivity of the new assay in the presence of spiked host RNA was shown to be 10^3 copies/reaction for TiLV, ISKNV, and *S. agalactiae*. In the presence of 10^4 copies of each template, the assay could detect *F. orientalis* as well as three other pathogens (Supplemental Figure 3). The assay could detect each pathogen efficiently in tilapia clinical samples, as confirmed by the presence of distinct PCR bands (Figure 2). The assay was able to detect TiLV in heavily infected samples (3-7) both on the gel and by nanopore but failed to detect it on the gel for samples (8-12) with low levels of TiLV infection. Similarly, the highest number of reads for ISKNV were found in sample #25 (twice more than in sample #27) and yet it had a much feinter band on the gel. Good correlation between read numbers and bands intensity for *F. orientalis* in four samples (#30-33) but only a weak band for *S. agalactiae* in one sample (#39) compared to number of amplicons (Figure 2, 3, Supplemental Table 5). In some samples, we detected dual infections with more than one pathogen, as evidenced by the presence of multiple bands, including a possible co-infection with *F. orientalis* and TiLV in samples 3-7. However, differentiation of *F. orientalis* and ISKNV was challenging as their respective bands have similar sizes (Figure 2).

Significant gains in readability achieved through iterative filtering processes

A total of 246,756 reads were successfully demultiplexed from three separate Flongle runs (r1-r3), each containing different sample types (Table 1, Supplemental Table 5). Following primer and length filtering, a substantial reduction in read count was observed, resulting in an average of 64% of reads being removed (range: 46%-87%) (Supplemental Table 5). An additional 10% of reads were removed after alignment filtering. After these filtering steps, a total of 75,203 reads remained, providing an average of 2,500 filtered reads per sample for subsequent read abundance and consensus sequence generation (Supplemental Table 5).

Successful recovery of pathogen variant from infected samples

Given the small sampling size and possibly highly conserved nature of some of the pathogen gene segments, one variant was generated per pathogen gene segment except for the TiLV gene segment whereby two variants were recovered (Table 2). All recovered variants exhibited 100% nucleotide identity to at least one publicly available sequence in the NCBI database. In addition, samples that were derived from the single PCR product inputs and their combined mixture were all clustered together in their respective variants. In the case of the TiLV amplicons, we observed two unique variants (Var 0 and Var 1)

for TiLV gene segment 9, one of which displayed 100% nucleotide identity to the gene region of a relatively
divergent TiLV strain from Vietnam (Table 2). This variant was detected in samples r2_NT_F1 to r2_NT_F5,
which were collected from the same sampling site at the same time (Table 2).

Effect of template input on data distribution and pathogen detection specificity on the Nanopore
platform

In the first run (r1), using well-defined pathogen nucleic acids as templates for the detection assay, confident
alignments were observed for each sample to their expected pathogen gene segments, except for sample
r1_TILV_1 which consisted purely of TiLV amplicons and did not sequence well enough to enable
subsequent variant calling (Figure 3 and Table 2). For sample 4PAT, which consisted of a combination of
products from 4 pathogens, the read distribution was broadly similar, as was the case for sample 4PAT_DIL,
which used a ten-fold diluted version of the template (Figure 3). In the second run (r2), actual clinical
samples with PCR-verified infections were used, and over 50% and up to 99% of reads were primarily
mapped to the primary suspected pathogen gene fragments, with the remaining reads mapping to the other
three pathogens. Interestingly, in the third run (r3), a significant percentage (>5%) of reads aligned to other
pathogens were observed among samples spiked with different amounts of single amplified products
(Figure 3). For example, among the TiLV samples (TILV05, 15, 30, 50), the lowest amount of PCR product
template (TILV05) resulted in low number of TiLV reads and a higher percentage of reads belonging to non-
TiLV pathogens. A similar trend was observed for pure SAG samples, with the most diluted SAG sample
(SAG05) having the highest percentage of non-SAG reads. However, in contrast to the pure synthetic TiLV
samples, the percentage of non-specific reads decreased to less than 2% in the SAG50 sample with the
highest read count, while it remained around 10% in the TILV50 sample.

Reads that failed stringent alignment to the pathogen gene panel were host-derived or partial
pathogen sequences with lower nucleotide identity

Upon investigation of reads that failed to align, a significant proportion were found to map not only to the
four pathogen gene segments, but also to the host genome, when a more lenient alignment configuration
was employed (Figure 4). Samples from run 2 (r2), which were predominantly derived from fish tissues,
exhibited a higher percentage of reads mapping to the host genome compared to other samples. As
expected, given the use of input PCR products as the template for all samples in run 1 (r1), no reads were
found to align to the host genome when they were aligned with Minimap2 (default setting). On the contrary,
a small portion of reads belonging to the host genome were found among different concentrations of single
PCR products from run 3 (r3) (Figure 4 and Supplemental Table 5) that also consists of samples derived
from tilapia organs (Table 1).

Discussion

The intensification of tilapia farming has led to a surge in emerging and re-emerging infectious diseases,
resulting in substantial losses in the aquaculture industry [6,9,38]. These diseases continue to evolve and
spread under intense selective pressures, particularly in regions with high-density tilapia farming. To
address this issue, we built upon our previous study [18] and improved the detection and genotyping of
multiple pathogens by switching to multiplex PCR coupled with Nanopore sequencing. Our updated
approach enabled the simultaneous and sequence-based detection and verification of four major tilapia
pathogens: ISKNV, TiLV, *S. agalactiae*, and *F. orientalis*.

The requirement for simultaneous diagnosis of multiple pathogens is particularly relevant to high mortalities observed in larger grow-out tilapia. These mortalities are likely exacerbated by co-infections, as indicated by [14] previous research findings, which revealed the active co-infection of ISKNV-positive fish with Streptococcus agalactiae and other bacterial pathogens.

The discrepancies between gel-based interpretation of multiplex PCR products and qPCR or direct amplicon sequencing approaches are not surprising, given the subjective nature of gel visualization and the labor-intensive process of running multiple samples. The accuracy of pathogen detection can be significantly impacted by faint or similar-sized bands, making it challenging to standardize results. To streamline the multiplex PCR approach and increase throughput, the gel visualization step can be eliminated in favor of direct Nanopore sequencing. With high-degree sample-based multiplexing, it is now possible to perform high throughput sequence-based detection of multiple tilapia pathogens using Nanopore sequencing. The utilization of this application is paramount for the aquaculture industry, given that disease outbreaks are frequently linked with multiple infections. Furthermore, as aquaculture farms are often situated in remote locations far from diagnostic reference laboratories, the ability to mobilize diagnostic testing with rapid result, high accuracy and less complicated equipment at the pond side is preferred.

Although our approach is scalable, there are limitations due to the old version of MinKNOW that was used at the time of the study. Short amplicons are sequenced less efficiently and have lower read accuracy. We addressed this issue by using a stringent read filtering and alignment approach, but this led to a reduced read recovery in the final abundance calculation. Some of the synthetic samples analyzed contained reads from other pathogens, which was unexpected and suggests cross-ligation of native barcodes during library preparation. The availability of unbound adapters is higher among samples with low DNA input, leading to higher levels of crosstalk and the presence of reads from other samples. New improvements in the current Nanopore technology (as of 13 May 2023) should mitigate these issues: (1) The use of "short read mode" in the latest MinKNOW software, (2) switching to PCR-barcoding kit to only enrich for amplicon with Nanopore partial adapter, and (3) the use of new Q20+ LSK114 kit with improved read accuracy.

By enforcing a stringent alignment setting, the reads that align to the pathogen gene segment are directly suitable for subsequent reference-based variant calling, producing highly accurate variants that are useful for biological interpretation. This is exemplified by the consistent recovery of a TiLV segment 9 variant previously only found in a Vietnamese TiLV strain in all Thai tilapia samples from the same sampling batch. Considering the close geographical proximity of Thailand and Vietnam, this may suggest a possible spillover event, highlighting the need for more thorough sampling and sequencing to elucidate the degree of TiLV diversity in both regions.

For field diagnostics, some of the equipment used in the present study can be replaced with portable items that are small enough to fit in a bag for use in remote settings. An example of the field application of this diagnostic workflow is the Lab-in-a-backpack concept developed by WorldFish [39–44]. It includes all the necessary sampling, extraction, and sequencing kits, pipettes, consumables, as well as the use of the miniPCR thermocycler and blueGel electrophoretic system from miniPCR bio™, a minicentrifuge, a magnetic rack, a fluorometer, a Nanopore MinION sequencer, and flow cells for sequencing. Additionally, a laptop with minimum requirements to run the MinKNOW software is required. Similar PCR-sequencing approaches using portable equipment have been employed for the rapid identification of a diverse range of biological specimens, including plants, insects, reptiles [45], and viruses such as COVID-19, Ebola, and Zika [46,47].

Improvements in the remote and low-resource applications of the current methodology will broaden its use for capacity building and implementations in low-income countries, where high mortalities due to unknown

causes are prevalent in tilapia farming. Accurate and rapid genomic detection of fish pathogens through our amplicon-based approach will allow health professionals to promptly provide advice and take necessary actions for producers. Moreover, it can facilitate informed decisions regarding additional investigations, such as whole-genome sequencing (WGS), on isolates stored in biobanks. The sequence data obtained from WGS offer crucial epidemiological insights, which can be utilized to develop customized multivalent autogenous vaccines using local pathogens [48]. These vaccines can then be administered to broodstock and seeds prior to distribution among grow-out farmers for restocking purposes.

In summary, our study presents an attractive approach for detecting and verifying multiple tilapia pathogens, which can also be applied to pathogens in other livestock, fish, or crustaceans, as long as the genetic information of the pathogen is publicly available. Although there are limitations to the current version of the pipeline, we are optimistic about the potential for future improvements in Nanopore technology, which will further enhance the accuracy and scalability of our approach. We believe that our workflow can serve as a valuable tool for disease surveillance and health management in aquaculture and beyond. We hope that this novel, cost-effective, multiplex PCR-Nanopore sequencing platform will stimulate further research and development in this field and contribute to the advancement of diagnostic technology in aquaculture.

Data Availability

Demultiplexed FastQ files for all 30 samples can be found under BioProject PRJNA957495 with the corresponding BioSample accession numbers from SAMN34371653-SAMN34371700. The intermediary files generated during the bioinformatic analyses are publicly available in the Zenodo.org dataset https://doi.org/10.5281/zenodo.7866295.

References

1. FAO. FishStatJ: universal software for fishery statistical time series: aquaculture production 1950–2020. FAO, Rome. (2022).
2. Wang, M. & Lu, M. Tilapia polyculture: a global review. *Aquaculture Research* **47**, 2363–2374 (2016).
3. FAO. *The State of World Fisheries and Aquaculture 2020: Sustainability in action*. (FAO, Rome, 2020). doi:10.4060/ca9229en.
4. Debnath, S. C. *et al.* Tilapia aquaculture, emerging diseases, and the roles of the skin microbiomes in health and disease. *Aquacult Int* (2023) doi:10.1007/s10499-023-01117-4.
5. Alathari, S. *et al.* A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia. *Viruses* **15**, 965 (2023).
6. Haenen, O. L. M. *et al.* Bacterial diseases of tilapia, their zoonotic potential and risk of antimicrobial resistance. *Reviews in Aquaculture* **15**, 154–185 (2023).
7. Kawasaki, M. *et al.* Microevolution of Streptococcus agalactiae ST-261 from Australia Indicates Dissemination via Imported Tilapia and Ongoing Adaptation to Marine Hosts or Environment. *Applied and Environmental Microbiology* **84**, (2018).
8. Mabrok, M. *et al.* Rapid visualization in the specific detection of Flavobacterium columnare, a causative agent of freshwater columnaris using a novel recombinase polymerase amplification (RPA) combined with lateral flow dipstick (LFD) assay. *Aquaculture* **531**, 735780 (2021).
9. Machimbirike, V. I. *et al.* Viral infections in tilapines: More than just tilapia lake virus. *Aquaculture* **503**, 508–518 (2019).

10. Dong, H. T. *et al.* From the basics to emerging diagnostic technologies: What is on the horizon for tilapia
disease diagnostics? *Reviews in Aquaculture* **15**, 186–212 (2023).

11. Kralik, P. & Ricchi, M. A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and
Everything. *Frontiers in Microbiology* **8**, (2017).

12. Liamnimitr, P., Thammatorn, W., U-thoomporn, S., Tattiyapong, P. & Surachetpong, W. Non-lethal sampling
for Tilapia Lake Virus detection by RT-qPCR and cell culture. *Aquaculture* **486**, 75–80 (2018).

13. López-Porras, A. *et al.* Application of multiplex quantitative polymerase chain reaction methods to detect
common bacterial fish pathogens in Nile tilapia, Oreochromis niloticus, hatcheries in Costa Rica. *Journal of the*
World Aquaculture Society **50**, 645–658 (2019).

14. Ramírez-Paredes, J. G. *et al.* First detection of infectious spleen and kidney necrosis virus (ISKNV) associated
with massive mortalities in farmed tilapia in Africa. *Transboundary and Emerging Diseases* **68**, 1550–1563
(2021).

15. Soto, E., Bowles, K., Fernandez, D. & Hawke, J. P. Development of a real-time PCR assay for identification and
quantification of the fish pathogen Francisella noatunensis subsp. orientalis. *Diseases of Aquatic Organisms*
89, 199–207 (2010).

16. Taengphu, S. *et al.* Concentration and quantification of Tilapia tilapinevirus from water using a simple iron
flocculation coupled with probe-based RT-qPCR. *PeerJ* **10**, e13157 (2022).

17. Waiyamitra, P. *et al.* A TaqMan RT-qPCR assay for tilapia lake virus (TiLV) detection in tilapia. *Aquaculture* **497**,
184–188 (2018).

18. Delamare-Deboutteville, J. *et al.* Rapid genotyping of tilapia lake virus (TiLV) using Nanopore sequencing.
Journal of Fish Diseases **44**, 1491–1502 (2021).

19. Meemetta, W., Domingos, J. A., Dong, H. T. & Senapin, S. Development of a SYBR Green quantitative PCR
assay for detection of Lates calcarifer herpesvirus (LCHV) in farmed barramundi. *Journal of Virological*
Methods **285**, 113920 (2020).

20. Dong, H. T. *et al.* Duplex PCR assay and in situ hybridization for detection of Francisella spp. and Francisella
noatunensis subsp. orientalis in red tilapia. *Dis Aquat Organ* **120**, 39–47 (2016).

21. Kawato, Y. *et al.* Development of New Real-time PCR Assays for Detecting *Megalocytivirus* Across Multiple
Genotypes. *Fish Pathology* **56**, 177–186 (2021).

22. Leigh, W. J. *et al.* Evaluation of PCR primers targeting the groEL gene for the specific detection of
Streptococcus agalactiae in the context of aquaculture. *J Appl Microbiol* **125**, 666–674 (2018).

23. Paria, A. *et al.* Evaluation of candidate reference genes for quantitative expression studies in Asian seabass
(Lates calcarifer) during ontogenesis and in tissues of healthy and infected fishes. *Indian journal of*
experimental biology **54**, 597–605 (2016).

24. Yang, C. G. *et al.* Evaluation of reference genes for quantitative real-time RT-PCR analysis of gene expression in
Nile tilapia (Oreochromis niloticus). *Gene* **527**, 183–192 (2013).

25. Nguyen, V. V., Dong, H. T., Senapin, S., Pirarat, N. & Rodkhum, C. *Francisella noatunensis* subsp. *orientalis* , an
emerging bacterial pathogen affecting cultured red tilapia (*Oreochromis* sp.) in Thailand. *Aquaculture*
Research **47**, 3697–3702 (2016).

26. Dong, H. T., Ataguba, G. A., Khunrae, P., Rattanarojpong, T. & Senapin, S. Evidence of TiLV infection in tilapia
hatcheries from 2012 to 2017 reveals probable global spread of the disease. *Aquaculture* **479**, 579–583
(2017).

27. Nguyen, V. V. *et al.* Transmission of Francisella noatuensis subsp. orientalis from subclinically infected hybrid
red tilapia broodstock (Oreochromis sp.) to their offspring. *Microbial Pathogenesis* **136**, 103670 (2019).

28. Thawornwattana, Y. *et al.* Tilapia lake virus (TiLV): Genomic epidemiology and its early origin. *Transboundary*
and Emerging Diseases **68**, 435–444 (2021).

29. Kerddee, P. *et al.* Molecular evidence for homologous strains of infectious spleen and kidney necrosis virus
(ISKNV) genotype I infecting inland freshwater cultured Asian sea bass (Lates calcarifer) in Thailand. *Arch Virol*
166, 3061–3074 (2021).
30. Karunanathie, H., Kee, P. S., Ng, S. F., Kennedy, M. A. & Chua, E. W. PCR enhancers: Types, mechanisms, and
applications in long-range PCR. *Biochimie* **197**, 130–143 (2022).
31. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**,
10–12 (2011).
32. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
33. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation.
PLOS ONE **11**, e0163962 (2016).
34. Core Pipeline - artic pipeline. https://artic.readthedocs.io/en/latest/minion/.
35. sars-cov-2-ont-artic-variant-calling/COVID-19-ARTIC-ONT. https://workflowhub.eu/workflows/111.
36. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies.
Bioinformatics **29**, 1072–1075 (2013).
37. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological
sequences. *Bioinformatics* **26**, 680–682 (2010).
38. Shinn, A. P. *et al.* A global review of problematic and pathogenic parasites of farmed tilapia. *Reviews in*
Aquaculture **15**, 92–153 (2023).
39. Huso, D., Wilkinson, S., Barnes, A. & Delamare-Deboutteville, J. Lab-in-a-backpack: Rapid Genomic Detection
to revolutionize control of disease outbreaks in fish farming. *FISH CRP* https://fish.cgiar.org/impact/stories-of-
change/lab-backpack-rapid-genomic-detection-revolutionize-control-disease (2020).
40. Cagua, F., Wilkinson, S., Barnes, A. & Delamare-Deboutteville, J. Lab in a Backpack: Rapid and accurate
diagnosis for the control and prevention of diseases in aquatic animals. https://labinabackpack.com/ (2021).
41. Barnes, A. *et al.* Diagnosis in a fish farmer's backpack. in *Food & Nutrition Security – The Biosecurity, Health,*
Trade Nexus. https://www.crawfordfund.org/events/other-events/2021-conference/speakers-chairs/dr-
andrew-barnes/ (Crawford Fund, 2021).
42. Chadag, V., Delamare-Deboutteville, J., Beveridge, M. & Marwaha, N. Reducing disease risks in fish through
better detection, management and prevention. *WorldFish. https://worldfishcenter.org/publication/reducing-*
disease-risks-fish-through-better-detection-management-and-prevention (2021).
43. The University of Queensland. Diagnosis in a fish farmer's backpack. YouTube.
https://www.youtube.com/watch?v=pUqmgQ-AuXM. (2021).
44. WorldFish. A New Concept for Rapid Genomic Detection of Fish Disease. YouTube.
https://www.youtube.com/watch?v=iFxbodO6Fos. (2020).
45. Pomerantz, A. *et al.* Rapid in situ identification of biological specimens via DNA amplicon sequencing using
miniaturized laboratory equipment. *Nat Protoc* **17**, 1415–1443 (2022).
46. González-González, E. *et al.* Validation of use of the miniPCR thermocycler for Ebola and Zika virus detection.
PLOS ONE **14**, e0215642 (2019).
47. González-González, E., Santiago, G. T., Lara-Mayorga, I. M., Martínez-Chapa, S. O. & Alvarez, M. M. Portable
and accurate diagnostics for COVID-19: Combined use of the miniPCR thermocycler and a well-plate reader for
SARS-CoV-2 virus detection. *PLOS ONE* **15**, e0237418 (2020).
48. Barnes, A. C. *et al.* Autogenous vaccination in aquaculture: A locally enabled solution towards reduction of the
global antimicrobial resistance problem. *Reviews in Aquaculture* **14**, 907–918 (2022).
49. Dong, H. T. *et al.* Infectious spleen and kidney necrosis disease (ISKND) outbreaks in farmed barramundi (Lates
calcarifer) in Vietnam. *Fish Shellfish Immunol* **68**, 65–73 (2017).
50. Kawato, Y. *et al.* Isolation and characterisation of an ISKNV-genotype megalocytivirus from imported angelfish
Pterophyllum scalare. *Diseases of Aquatic Organisms* **140**, 129–141 (2020).

Acknowledgments

We would like to extend our utmost gratitude and profound recognition to the late Dr. Pattanapon Kayansamruaj for his valuable contributions during the nascent stages of conceiving this research endeavor. This work was undertaken as part of the CGIAR Inspire Challenge project on Rapid Genomic Detection of Aquaculture Pathogens, the CGIAR Initiatives on *Aquatic Foods*, and *Protecting Human Health Through a One Health Approach*, and the Norwegian Agency for Development Cooperation (Norad) project *Increased Sustainability in the Aquaculture Sector in sub-Saharan Africa, through Improved Aquatic Animal Health Management*. We would like to thank all funders who support this research through their contributions to the CGIAR Trust Fund: https://www.cgiar.org/funders/, and Norad. Funders were not involved in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

The supplemental tables and figures have been submitted in separate files.

468 **Table 1.** Sources of clinical tilapia samples and PCR detection results

Code	Sample_name	Life stage	Country	Year Isolation	Previous Diagnosis sPCR	mPCR	Refs/comments
1	Healthy NT 1	juvenile (NT)	Thailand	2020	-	-	[16]
2	Healthy NT 2	juvenile (NT)	Thailand	2020	-	-	[16]
3	*r2_NT F1*	fingerling (NT)	Thailand	2020	TiLV (++)	TiLV*	[16]
4	*r2_NT F2*	fingerling (NT)	Thailand	2020	TiLV (++)	TiLV*	[16]
5	*r2_NT F3*	fingerling (NT)	Thailand	2020	TiLV (++)	TiLV*	[16]
6	*r2_NT F4*	fingerling (NT)	Thailand	2020	TiLV (++)	TiLV*	[16]
7	*r2_NT F5*	fingerling (NT)	Thailand	2020	TiLV (++)	TiLV*	[16]
8	NT fingerling pond C1	fingerling (NT)	Thailand	2020	TiLV (+)	-	[16]
9	NT fingerling pond C2	fingerling (NT)	Thailand		TiLV (+)	-	[16]
10	NT female brood	adult (NT)	Thailand		TiLV (+)	-	[16]
11	NT male brood	adult (NT)	Thailand		TiLV (+)	-	[16]
12	NT juvenile 4	juvenile (NT)	Thailand		TiLV (+)	-	[16]
13	RT F1	fingerling (RT)	Thailand	2015	FnO (+)	-	[25]
14	RT F2	fingerling (RT)	Thailand		FnO (+)	-	[25]
15	RT F3	fingerling (RT)	Thailand		FnO (+)	-	[25]
16	RT F4	fingerling (RT)	Thailand		FnO (+)	-	[25]
17	NT 1.2	fingerling (NT)	Thailand	2015	TiLV (+)	-	[26]
18	NT 1.3	fingerling (NT)	Thailand		TiLV (+)	-	[26]
19	NT 2.1	fingerling (NT)	Thailand		TiLV (+)	-	[26]
20	NT 2.3	fingerling (NT)	Thailand		TiLV (+)	-	[26]
21	NT 3.1	fingerling (NT)	Thailand		TiLV (+)	-	[26]
22	NT 3.3	fingerling (NT)	Thailand		TiLV (+)	-	[26]
23	NT Gh 21	juvenile (NT)	West Africa	2019	ISKNV (++)	ISKNV*	-
24	NT Gh 22	juvenile (NT)	West Africa		ISKNV (++)	ISKNV*	-
25	*r2_NT Gh 23*	juvenile (NT)	West Africa		ISKNV (++)	ISKNV*	-
26	*r2_NT Gh 24*	adult (NT)	West Africa		ISKNV (++)	ISKNV*	-
27	*r2_NT Gh 25*	juvenile (NT)	West Africa		ISKNV (++)	ISKNV*	-
28	RT EX 01	fry (RT)	Thailand	2019	FnO (++)	FnO#	[27]
29	RT EX02	fry (RT)	Thailand		FnO (++)	FnO#	[27]
30	*r2_RT EX03*	fry (RT)	Thailand		FnO (++)	FnO#	[27]
31	*r2_RT FM1a*	adult (RT)	Thailand		FnO (++)	FnO#	[27]
32	*r2_RT M1*	adult (RT)	Thailand		FnO (++)	FnO#	[27]
33	*r2_RT FM1b*	adult (RT)	Thailand		FnO (++)	FnO#	[27]
34	NT PV F8	adult (NT)	Thailand	2020	Not done	*	-
35	NT PV F9	adult (NT)	Thailand		Not done	#	-
36	NT PV F10	adult (NT)	Thailand		Not done	#	-
37	*r3_NT Group B*	fry (NT)	Thailand	2019	SAG (++)	SAG	-
38	*r3_NT Zone A*	fry (NT)	Thailand		SAG (++)	SAG*	-
39	*r3_NT Zone B*	fry (NT)	Thailand		SAG (++)	SAG	-
40	*r1_TILV_1*	fingerling (RT)	Thailand	2018-07	TiLV	TiLV	[28]
41	*r1_ISKNV_1*	fry (AS)	Thailand	2018-11	ISKNV	ISKNV	[29]
42	*r1_FNO_1*	fingerling (RT)	Thailand	2015	FnO	FnO	[25]
43	*r1_SAG_1*	fry (NT)	Thailand	2019-07	SAG	SAG	
44	*r3_TILV50*	fingerling (RT)	Thailand	2019-07	TiLV	TiLV	TiLV PCR amplicon of 100 ng
45	*r3_TILV30*	fingerling (RT)	Thailand	2019-07	TiLV	TiLV	TiLV PCR amplicon of 75 ng
46	*r3_TILV15*	fingerling (RT)	Thailand	2019-07	TiLV	TiLV	TiLV PCR amplicon of 50 ng
47	*r3_TILV05*	fingerling (RT)	Thailand	2019-07	TiLV	TiLV	TiLV PCR amplicon of 25 ng
48	*r3_SAG50*	fry (NT)	Thailand	2019-07	SAG	SAG	SAG PCR amplicon of 100 ng
49	*r3_SAG30*	fry (NT)	Thailand	2019-07	SAG	SAG	SAG PCR amplicon of 75 ng
50	*r3_SAG15*	fry (NT)	Thailand	2019-07	SAG	SAG	SAG PCR amplicon of 50 ng
51	*r3_SAG05*	fry (NT)	Thailand	2019-07	SAG	SAG	SAG PCR amplicon of 25 ng
52	*r1_4PAT_1*	N/A	Thailand	N/A	4PAT	4PAT	PCR amplicons 4 PAT (combined1)
53	*r1_4PAT_DIL_1*	N/A	Thailand	N/A	4PAT	4PAT	PCR amplicons 4 PAT (combined2)
54	*r3_4PAT_DIL_3*	N/A	Thailand	N/A	4PAT	4PAT	PCR amplicons 4 PAT (combined3)

Notes and abbreviations: sPCR, singleplex PCR; mPCR, multiplex PCR; NT, Nile tilapia; RT, red tilapia; AS, Asian sea bass; -,
negative test; ++, high pathogen load; +, low pathogen load; *, suspected dual infections with other pathogen(s); #, probably non-
specific products; TiLV, tilapia lake virus; ISKNV, infectious spleen and kidney necrosis virus; FnO, *Francisella noatunensis* subsp.
orientalis; SAG, *Streptococcus agalactiae*; 4PAT: mixture of PCR amplicons 4 pathogens (TiLV, ISKNV, FnO, SAG). Samples
sequenced are italicized and highlighted in grey, and the designation "r1" to "r3" indicates the run number in Nanopore.

Table 2. List of variants for the pathogen gene region, along with the inferred variants and their top 5 NCBI
BLAST hits. Pat: Pathogen; Var, Variants; %ID, % nucleotide identity. The samples were grouped based
on the sequencing run from which they were generated.

Pat	Samples	Var	Top 5 BLAST hits	%ID	Accession
FnO	r1_4PAT_1,r1_4PAT_DIL_1,r1_FNO_1,r1_SAG_1 r2_NT_F1,r2_NT_F2,r2_NT_F3,r2_NT_F4,r2_NT_F5,r2_NT_Gh23,r2_NT_Gh24,r2_NT_Gh25,r2_RT_EX03,r2_RT_FM1A,r2_RT_FM1B,r2_RT_M1 r3_4PAT_DIL_3	0	*Francisella noatunensis* subsp. *orientalis* strain LPM2-AR2019	100	MN385384.1
			Francisella noatunensis subsp. *orientalis* strain FO371	100	CP022953.1
			Francisella noatunensis subsp. *orientalis* strain FNO364	100	CP022952.1
			Francisella noatunensis subsp. *orientalis* strain FNO222	100	CP022951.1
			Francisella noatunensis subsp. *orientalis* strain FNO215	100	CP022950.1
ISKNV	r1_4PAT_1,r1_4PAT_DIL_1,r1_ISKNV_1 r2_NT_F2,r2_NT_Gh23,r2_NT_Gh24,r2_NT_Gh25	0	ISKNV isolate ISKNV/10	100	MT178422.1
			ISKNV isolate ISKNV/48	100	MT178418.1
			Angelfish iridovirus AFIV-16	100	MK689685.1
			ISKNV isolate M6	100	MK084827.1
			ISKNV isolate SB04	100	KY440040.1
SAG	r1_SAG_1,r1_4PAT_1,r1_4PAT_DIL_1,r1_SAG_1 r3_4PAT_DIL_3,r3_NT_GroupB,r3_NT_ZoneA,r3_NT_ZoneB,r3_SAG05,r3_SAG15,r3_SAG30,r3_SAG50	0	*Streptococcus agalactiae* strain 01173	100	CP053027.1
			Streptococcus agalactiae strain Sag153	100	CP036376.1
			Streptococcus agalactiae strain ZQ0910	100	CP049938.1
			Streptococcus agalactiae strain NJ1606	100	CP026084.1
			Streptococcus agalactiae strain BSE009	100	CP020387.1
TiLV	r1_4PAT_1,r1_4PAT_DIL_1 r3_4PAT_DIL_3,r3_TILV15,r3_TILV30,r3_TILV50	0	TiLV isolate WVL19054 segment 9	100	MN193531.1
			TiLV isolate WVL19031-01A segment 9	100	MN193521.1
			TiLV isolate EC-2012 segment 9	100	MK392380.1
			TiLV isolate Til-4-2011 segment 9	100	KU751822.1
			TiLV AD-2016 Contig 20	100	KU552140.1
	r2_NT_F1,r2_NT_F2,r2_NT_F3,r2_NT_F4,r2_NT_F5	1	TiLV isolate RIA2-VN-2019 segment 9	100	ON376590.1
			TiLV isolate WVL19054 segment 9	98.91	MN193531.1
			TiLV isolate WVL19031-01A segment 9	98.91	MN193521.1
			TiLV strain EC-2012 segment 9	98.91	MK392380.1
			TiLV isolate Til-4-2011 segment 9	98.91	KU751822.1

477 Abbreviations: FnO, *Francisella noatunensis* subsp. *orientalis*; ISKNV, infectious spleen and kidney necrosis virus;
478 SAG, *Streptococcus agalactiae;* TiLV, tilapia lake virus

Figure legends

Figure 1. Simplified workflow illustration of diseased fish sampling on farm to Nanopore sequencing results.
The entire process can take less than 12 hours.

Figure 2. Multiplex PCR detection of TiLV, ISKNV, *F. orientalis* and *S. agalactiae* from clinical diseased
tilapia. Sample codes are listed in Table 1. The expected amplicon size for each pathogen is indicated on

the right. -ve, no template control, +ve, mix positive control plasmids; M, 100 bp DNA marker (New England
BioLabs). Samples marked *r2* to *r3* were subjected to Nanopore sequencing run number 2 and 3,
respectively.

Figure 3. Percentage of reads mapping to each pathogen after primer trimming (left axis) and the total
filtered read count for each sample (right axis). Table at the top: single pathogen tested by qPCR assay,
with Cq value indicated; ud, undetectable.

Figure 4. Distribution of reads assigned to the host genome and four pathogens at various stringency
levels, using five different mapping strategies. (1) 4pat_Pass, Reads that were directly assigned/mapped
to one of the four pathogens using the reported stringent Minimap2 alignment strategy; (2) 4pat_M, Raw
reads that failed to map to the pathogen genes using the stringent Minimap2 alignment strategy but
subsequently aligned to the pathogen genes after using the default Minimap2 alignment setting; (3)
Tilapia_M, Raw reads that failed to map to the pathogen genes using the stringent Minimap2 alignment
approach but were subsequently aligned to the host genome after using the default Minimap2 alignment
setting; (4) Tilapia_BN, Raw reads that could only be aligned to the host genome using BlastN after failing
to be aligned with the default Minimap2 alignment strategy; (5) 4pat_BN, Raw reads that could only be
aligned to the pathogen genes using BlastN after failing to be aligned with the default Minimap2 alignment
strategy.

Supplemental figure legends

Supplemental Figure 1 Sensitivity assay of the developed multiplex PCR was performed using indicated
serial dilution of mix control plasmids as template. Two representative replicates out of four are shown. M,
DNA marker (New England Biolabs), -ve, no template control, 3.5% agarose gel electrophoresis using 1X
TAE buffer was used.

Supplemental Figure 2. Agarose gel electrophoresis (3.5% gel) of amplified products from single PCR
amplification on respective templates of TiLV (T), ISKNV (I), *F. orientalis* (F), and *S. agalactiae* (S).
Expected bands sizes are indicated on the right side of the gel. C: combined PCR products loaded in the
same well. M, DNA marker (New England Biolabs).

Supplemental Figure 3. Temperature gradient PCR results of multiplex PCR reactions using two different
master mix: top: Luna Universal One-Step qRT-PCR reagent, bottom: KAPA SYBR FAST One-Step qRT-
PCR reagent. Annealing temperatures ranged from 55-65 °C.

Supplemental table legends

Supplemental Table 1. Primers used in this experiment.

Supplemental Table 2. Sources of plasmid template used in sensitivity assay.

Supplemental Table 3. Reaction mixture and cycling conditions for multiplex PCR diagnosis.

Supplemental Table 4. Summary of qPCR testing for the detection and quantification of pathogens in
tilapia tissue samples

Supplemental Table 5. Sequencing and alignment statistics

1. Fish sampling for disease diagnostics

30 min

2. Nucleic acid extraction

30 min

3. Multiplex PCR

2 h

4. Gel electrophoresis & DNA clean up

30 min

5. Nanopore multiplex Library preparation

90 min

6. MinION sequencing

30 min

7. Mapping to reference

15 min

8. Disease control in aquaculture.

Cq values of the qPCR assay

TiLV	13	11	12	10	12	ud	ud	ud	34	ud	35	35	33	32	32
ISKNV	ud	ud	ud	ud	38	30	33	14	ud	ud	ud	ud	ud	ud	ud
FnO	ud	ud	ud	ud	35	ud	ud	ud	23	14	18	13	ud	ud	ud
SAG	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	24	25	33