Multiplex PCR detection and Nanopore-based rapid genotyping of tilapia (*Oreochromis* spp.) pathogens.

Jérôme Delamare-Deboutteville[1*#], Watcharachai Meemetta[2#], Khaettareeya Pimsannil[2], Han Ming Gan[3], Pattanapon Kayansamruaj[4], Laura Khor Li Imm[1], Ha Thanh Dong[5], Chadag Vishnumurthy Mohan[1], Saengchan Senapin[2,6*]

[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]Department of Aquaculture, Faculty of Fisheries, Kasetsart University, Bangkok, Thailand
[5]School of Environment, Resources and Development, Asian Institute of Technology, Pathum Thani, 12120, Thailand
[6]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, senapin2010@gmail.com

 # First authors with equal contributions

Acknowledgments

This research was undertaken as part of the CGIAR Initiatives on Aquatic Foods led by WorldFish, and the One CGIAR Initiative "Protecting Human Health Through a One Health Approach", supported by contributors to the CGIAR Trust Fund: https://www.cgiar.org/funders/. Funding provided support in the form of salary for authors [J.D.D; L.K.L.I, C.V.M], travels, laboratory consumables and analytical costs.

Data Availability Statement

Demultiplexed FastQ files for all 30 samples can be found under BioProject PRJNA957495 with the corresponding BioSample accession numbers from SAMN34371653-SAMN34371700. The linux scripts used to generate initial FastQ files, assembled amplicons (public and from this study) along with the intermediate files generated from this study are publicly available in the Zenodo.org dataset https://doi.org/10.5281/zenodo.7866295.

Conflict of interest statement

The authors declare no conflict of interest.

Ethics approval statement

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

ABSTRACT

Tilapia aquaculture is threatened by multiple pathogens, including Tilapia Lake virus (TiLV), Infectious spleen and kidney necrosis virus (ISKNV), *Francisella orientalis*, and *Streptococcus agalactiae*. Current diagnostic methods rely on multiple singleplex PCR reactions, which are time-consuming and expensive. Here, we introduce a novel multiplex PCR (mPCR) assay coupled with rapid Nanopore sequencing that allows for the simultaneous detection and sequence-based identification of all four pathogens. Our multiplex assay exhibits high sensitivity, detecting as few as 1,000 copies/reaction for ISKNV and *S. agalactiae*, and 10,000 copies/reaction for TiLV and *F. orientalis*, making it capable of detecting single and dual infections in clinical samples within a day. The consensus sequences generated from the amplicons for each sample exhibit 100% sequence identity to publicly available data, providing strong support for the improving accuracy of Nanopore sequencing. Combining multiplex PCR and Nanopore sequencing provides a promising and cost-effective platform for rapid and accurate diagnostics of major pathogens of tilapia, making it a valuable tool for the aquaculture industry.

Keywords: Tilapia, TiLV, ISKNV, *Streptococcus agalactiae*, *Francisella orientalis*, multiplex PCR, diagnostic by sequencing, Oxford Nanopore Technologies, bioinformatics analysis

INTRODUCTION

Tilapia (*Oreochromis* spp.) is one of the most widely farmed freshwater fish species globally due to its adaptability, fast growth, and excellent meat quality. Global production is estimated at 4,525,400 tonnes per annum [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 [1], 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 [3]. Among the most important tilapia pathogens are Tilapia Lake virus (TiLV), infectious spleen and kidney necrosis virus (ISKNV), *Flavobacterium orientalis*, and *Streptococcus agalactiae* [4] [5] [6–8]. Effective diagnosis and monitoring of these pathogens are crucial for disease management and control [9]. 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 [9]. 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 [9,10] [11–16].

In 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 of even more interest 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 [17]. Proper integration of

multiplex PCR and ONT-based genotyping may therefore present a promising tool for the 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, sensitive, and cost-effective solution for the identification and characterization of multiple pathogens in a single reaction, with potential applications in disease diagnosis, surveillance, and control.

MATERIALS AND METHODS

Clinical samples and nucleic acid extraction

We analyzed archival clinical samples of fry, fingerling, juvenile and adult Nile tilapia and red tilapia 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 dual infection with more than one pathogen. 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 [18]. The diagnostic workflow from the point of sample collection on the farm to the final data analysis, which included nucleic acid extraction, multiplex PCR, library preparation, Nanopore sequencing, and data analysis is illustrated in Figure 1.

Primers used in this study: previous and newly designed

We obtained the primers for the four selected target pathogens from Bio Basic (Singapore), and the primer sequences are summarized in Supplemental_Table1. While primers for ISKNV, *F. orientalis*, and *S. agalactiae* were reported in previous studies [19] [20], primers for TiLV were newly designed based on the genomic sequence of segment 9. Expected amplicon sizes were 190, 203, 351 and 137 bp, respectively. Specificity of each primer pair was assessed *in silico* using Primer-BLAST program (https://www.ncbi.nlm.nih.gov/tools/primer-blast/).

Plasmid positive controls

All plasmid positive controls except for *S. agalactiae* (constructed in this study) were obtained from previous studies (Supplemental_Table2). 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.

Multiplex PCR detection of clinical samples

Previous studies claimed that qPCR reagents usually contain PCR additives and enhancers that can improve amplification efficiency [21]. 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_Table3). We determined the optimal annealing temperature (Ta) using gradient PCR with Ta ranging between 55-65 °C (Supplemental_Figure2). To identify the best combination of PCR components, we added ammonium sulfate, BSA, dNTPs, and MgCl2 in different proportions. The 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 Ta for 30

sec, with a final extension step at 60 °C for 5 min (Supplemental_Table3). We assessed the sensitivity of the assay using 10-fold serially diluted control plasmids for all four target templates (Supplemental_Figure3). To simulate a clinical sample from a fish, we spiked 50 ng of RNA extracted from a healthy tilapia into each reaction. We then analyzed 10 µl of each product by electrophoresis on a 3.5% agarose gel stained with ethidium bromide (Supplemental_Figure1). 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.

qPCR assay

To assess the efficiency of the tested extraction methods, templates were subjected to a qPCR assay using fish EF primers. Each 20 µL qPCR reaction consisted of 1X master mix (Luna Universal One-Step RT-qPCR kit), 1 µL of each template, and 200 nM of each primer pair. Cycling conditions involved a reverse transcription step at 55°C for 10 min, followed by an initial denaturation at 95°C for 3 min and 40 cycles of 95°C for 10 s and 60°C for 1 min. We compared cycle quantification (Cq) values to evaluate the efficiency of the tested extraction methods.

Nanopore sequencing

The present study employed amplicons generated from both single and multiplex PCR reactions (Table 1, samples marked with *) 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. 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 [22] 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 [23]. 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 [24]. 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 haplotype sequences

Only reads with unique alignments were used as the template for minimap2 and medaka variant calling based on the ARTIC pipeline [25] [26]. 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 [27], and only sequences without any ambiguous base were selected for subsequent haplotype generation. Haplotypes were generated using cd-hit v.4.8.1 [28], whereby sequences with 100% nucleotide identity and exact sequence length were clustered into the same haplotype. The haplotype sequences were compared against the NCBI Blast nt database (accessed on 25 April 2023) to identify the top 5 BLAST hits for each sequence.

RESULTS

Singleplex PCR condition optimization and detection limit

Single PCR amplification reaction using each primer pair and its respective target template (Supplemental Tables 1 and 2) generated correct predicted amplicon size (Supplemental Figure 1). Detection sensitivity of the new assay in the presence of spiked host RNA was shown to be 10^4 copies/reaction for all pathogens. However, in the presence of 10^3 copies of each template, the assay could detect *S. agalactiae,* TiLV and ISKNV but not *F. orientalis* (Supplemental Figure 3).

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. After thorough optimization of cycling parameters and PCR reaction mixture compositions, the assay was capable of detecting each pathogen efficiently in tilapia clinical samples, as confirmed by the presence of distinct PCR bands (Supplemental Table 3 and Supplemental Figure 2). The assay was able to detect TiLV in heavily infected samples (3-7) but failed to detect it in samples (8-12) with low levels of TiLV. Similarly, we observed high amplification of ISKNV in one sample (#26), and F. orientalis in four samples (#13-16) but only a weak amplification for *S. agalactiae* in one sample (#39) (Figure 2). 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).

Substantial read loss following multiple filtering steps

A total of 246,756 reads were successfully demultiplexed from three separate Flongle runs, each containing different sample types (Table 1). 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).

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

In the first run, 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 virus 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 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. In the second run, 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, a significant percentage (>5%) of reads aligned to other pathogens were observed among samples spiked with different amounts of pure viral nucleic acid materials (Figure 3). For example, among the TILV samples (TILV05, 15, 30, 50), the lowest amount of spiked viral nucleic acid (TILV05) resulted in weaker PCR bands 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, 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 synthetic PCR products as the template for all samples in run 1, 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 synthetic samples from run 3 (Figure 4 and Supplemental Table 5) that also consists of samples derived from tilapia organs (Table 1).

Successful recovery of pathogen haplotype from infected and synthetic samples

Given the small sampling size and possibly highly conserved nature of some of the pathogen gene segments, one haplotype was generated per pathogen gene segment except for the TilV gene segment whereby two haplotypes were recovered (Table 2). All recovered haplotypes exhibited 100% nucleotide identity to at least one publicly available sequence in the NCBI database. In addition, samples that were derived from the same synthetic plasmid were all clustered together in their respective haplotypes. In the case of the TILV amplicons, we observed two unique haplotypes 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 haplotype was detected in samples r2_NT_F1 to r2_NT_F5, which were collected from the same sampling site at the same time (Table 1).

DISCUSSIONS

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 [4,5,29] . 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 [22] 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 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 completely 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.

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 haplotype 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.

In summary, our study describes an attractive approach for detecting and verifying multiple tilapia pathogens, which has broader applicability for key pathogens in other livestock, fish, or crustaceans, provided that the pathogen genetic information is publicly available. While there are limitations to the current version of the pipeline, we are optimistic about the potential for future improvements in Nanopore technology to further enhance the accuracy and scalability of our approach. We believe that our workflow can be a valuable tool for disease surveillance and management in aquaculture and beyond, and we hope that our findings will spur further research and development in this area.

REFERENCES

[1] FAO, The state of world fisheries and aquaculture 2020. Sustainability in action, Food and Agriculture Organization of the United Nations. (2020). https://www.fao.org/3/ca9229en/ca9229en.pdf

[2] M. Wang, M. Lu, Tilapia polyculture: a global review, Aquaculture Research. 47 (2016) 2363–2374. https://doi.org/10.1111/are.12708.

[3] S.C. Debnath, J. McMurtrie, B. Temperton, J. Delamare-Deboutteville, C.V. Mohan, C.R. Tyler, Tilapia aquaculture, emerging diseases, and the roles of the skin microbiomes in health and disease, Aquacult Int. (2023). https://doi.org/10.1007/s10499-023-01117-4.

[4] O.L.M. Haenen, H.T. Dong, T.D. Hoai, M. Crumlish, I. Karunasagar, T. Barkham, S.L. Chen, R. Zadoks, A. Kiermeier, B. Wang, E.G. Gamarro, M. Takeuchi, M.N.A. Azmai, B. Fouz, R. Pakingking Jr., Z.W. Wei, M.G. Bondad-Reantaso, Bacterial diseases of tilapia, their zoonotic potential and risk of antimicrobial resistance, Reviews in Aquaculture. 15 (2023) 154–185. https://doi.org/10.1111/raq.12743.

[5] V.I. Machimbirike, M.D. Jansen, S. Senapin, P. Khunrae, T. Rattanarojpong, H.T. Dong, Viral infections in tilapines: More than just tilapia lake virus, Aquaculture. 503 (2019) 508–518. https://doi.org/10.1016/j.aquaculture.2019.01.036.

[6] M. Mabrok, S. Elayaraja, P. Chokmangmeepisarn, W. Jaroenram, N. Arunrut, W. Kiatpathomchai, P.P. Debnath, J. Delamare-Deboutteville, C.V. Mohan, A. Fawzy, C. Rodkhum, 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 (2021) 735780. https://doi.org/10.1016/j.aquaculture.2020.735780.

[7] S. Alathari, D.L. Chaput, L.M. Bolaños, A. Joseph, V.L.N. Jackson, D. Verner-Jeffreys, R. Paley, C.R. Tyler, B. Temperton, A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia, Viruses. 15 (2023) 965. https://doi.org/10.3390/v15040965.

[8] M. Kawasaki, J. Delamare-Deboutteville, R.O. Bowater, M.J. Walker, S. Beatson, N.L. Ben Zakour, A.C. Barnes, 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). https://doi.org/10.1128/AEM.00859-18.

[9] H.T. Dong, T. Chaijarasphong, A.C. Barnes, J. Delamare-Deboutteville, P.A. Lee, S. Senapin, C.V. Mohan, K.F.J. Tang, S.E. McGladdery, M.G. Bondad-Reantaso, From the basics to emerging diagnostic technologies: What is on the horizon for tilapia disease diagnostics? Reviews in

Aquaculture. 15 (2023) 186–212. https://doi.org/10.1111/raq.12734.

[10] P. Kralik, M. Ricchi, A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything, Frontiers in Microbiology. 8 (2017). https://www.frontiersin.org/articles/10.3389/fmicb.2017.00108 (accessed May 12, 2023).

[11] E. Soto, K. Bowles, D. Fernandez, J.P. Hawke, Development of a real-time PCR assay for identification and quantification of the fish pathogen *Francisella noatunensis* subsp. *orientalis*, Diseases of Aquatic Organisms. 89 (2010) 199–207. https://doi.org/10.3354/dao02204.

[12] P. Liamnimitr, W. Thammatorn, S. U-thoomporn, P. Tattiyapong, W. Surachetpong, Non-lethal sampling for Tilapia Lake Virus detection by RT-qPCR and cell culture, Aquaculture. 486 (2018) 75–80. https://doi.org/10.1016/j.aquaculture.2017.12.015.

[13] P. Waiyamitra, P. Tattiyapong, K. Sirikanchana, S. Mongkolsuk, P. Nicholson, W. Surachetpong, A TaqMan RT-qPCR assay for tilapia lake virus (TiLV) detection in tilapia, Aquaculture. 497 (2018) 184–188. https://doi.org/10.1016/j.aquaculture.2018.07.060.

[14] A. López-Porras, C. Elizondo, A.J. Chaves, A.C. Camus, M.J. Griffin, K. Kenelty, S. Barum, E. Barquero-Calvo, E. Soto, 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 (2019) 645–658. https://doi.org/10.1111/jwas.12576.

[15] J.G. Ramírez-Paredes, R.K. Paley, W. Hunt, S.W. Feist, D.M. Stone, T.R. Field, D.J. Haydon, P.A. Ziddah, M. Nkansa, J. Guilder, J. Gray, S. Duodu, E.K. Pecku, J.A. Awuni, T.S. Wallis, D.W. Verner-Jeffreys, First detection of infectious spleen and kidney necrosis virus (ISKNV) associated with massive mortalities in farmed tilapia in Africa, Transboundary and Emerging Diseases. 68 (2021). https://doi.org/10.1111/tbed.13825.

[16] S. Taengphu, P. Kayansamruaj, Y. Kawato, J. Delamare-Deboutteville, C.V. Mohan, H.T. Dong, S. Senapin, Concentration and quantification of Tilapia tilapinevirus from water using a simple iron flocculation coupled with probe-based RT-qPCR, PeerJ. 10 (2022) e13157. https://doi.org/10.7717/peerj.13157.

[17] J. Delamare-Deboutteville, S. Taengphu, H.M. Gan, P. Kayansamruaj, P.P. Debnath, A. Barnes, S. Wilkinson, M. Kawasaki, C. Vishnumurthy Mohan, S. Senapin, H.T. Dong, Rapid genotyping of tilapia lake virus (TiLV) using Nanopore sequencing, Journal of Fish Diseases. 44 (2021) 1491–1502. https://doi.org/10.1111/jfd.13467.

[18] W. Meemetta, J.A. Domingos, H.T. Dong, S. Senapin, Development of a SYBR Green quantitative PCR assay for detection of *Lates calcarifer* herpesvirus (LCHV) in farmed barramundi, Journal of Virological Methods. 285 (2020) 113920. https://doi.org/10.1016/j.jviromet.2020.113920.

[19]   H.T. Dong, W. Gangnonngiw, K. Phiwsaiya, W. Charoensapsri, V.V. Nguyen, P. Nilsen, P.J. Pradeep, B. Withyachumnarnkul, S. Senapin, C. Rodkhum, Duplex PCR assay and *in situ* hybridization for detection of *Francisella* spp. and *Francisella noatunensis* subsp. *orientalis* in red tilapia, Dis Aquat Organ. 120 (2016) 39–47. https://doi.org/10.3354/dao03021.

[20]   W.J. Leigh, R.N. Zadoks, A. Jaglarz, J.Z. Costa, G. Foster, K.D. Thompson, Evaluation of PCR primers targeting the *groEL* gene for the specific detection of *Streptococcus agalactiae* in the context of aquaculture, J Appl Microbiol. 125 (2018) 666–674. https://doi.org/10.1111/jam.13925.

[21]   H. Karunanathie, P.S. Kee, S.F. Ng, M.A. Kennedy, E.W. Chua, PCR enhancers: Types, mechanisms, and applications in long-range PCR, Biochimie. 197 (2022) 130–143. https://doi.org/10.1016/j.biochi.2022.02.009.

[22]   M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet.Journal. 17 (2011) 10–12. https://doi.org/10.14806/ej.17.1.200.

[23]   H. Li, Minimap2: pairwise alignment for nucleotide sequences, Bioinformatics. 34 (2018) 3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

[24]   W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation, PLOS ONE. 11 (2016) e0163962. https://doi.org/10.1371/journal.pone.0163962.

[25]   Core Pipeline - artic pipeline, (n.d.). https://artic.readthedocs.io/en/latest/minion/ (accessed April 27, 2023).

[26]   sars-cov-2-ont-artic-variant-calling/COVID-19-ARTIC-ONT, (n.d.). https://workflowhub.eu/workflows/111 (accessed May 12, 2023).

[27]   A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: quality assessment tool for genome assemblies, Bioinformatics. 29 (2013) 1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

[28]   Y. Huang, B. Niu, Y. Gao, L. Fu, W. Li, CD-HIT Suite: a web server for clustering and comparing biological sequences, Bioinformatics. 26 (2010) 680–682. https://doi.org/10.1093/bioinformatics/btq003.

[29]   A.P. Shinn, A. Avenant-Oldewage, M.G. Bondad-Reantaso, A.J. Cruz-Laufer, A. García-Vásquez, J.S. Hernández-Orts, R. Kuchta, M. Longshaw, M. Metselaar, A. Pariselle, G. Pérez-Ponce de León, P.K. Pradhan, M. Rubio-Godoy, N. Sood, M.P.M. Vanhove, M.R. Deveney, A global review of problematic and pathogenic parasites of farmed tilapia, Reviews in Aquaculture. 15 (2023) 92–153. https://doi.org/10.1111/raq.12742.

[30]   H.T. Dong, S. Jitrakorn, P. Kayansamruaj, N. Pirarat, C. Rodkhum, T. Rattanarojpong, S. Senapin, V.

Saksmerprome, Infectious spleen and kidney necrosis disease (ISKND) outbreaks in farmed barramundi (*Lates calcarifer*) in Vietnam, Fish Shellfish Immunol. 68 (2017) 65–73. https://doi.org/10.1016/j.fsi.2017.06.054.

[31] Y. Kawato, P.G. Mohr, M.S.J. Crane, L.M. Williams, M.J. Neave, D.M. Cummins, M. Dearnley, S. Crameri, C. Holmes, J. Hoad, N.J.G. Moody, Isolation and characterisation of an ISKNV-genotype megalocytivirus from imported angelfish *Pterophyllum scalare*, Diseases of Aquatic Organisms. 140 (2020) 129–141. https://doi.org/10.3354/dao03499.

[32] C.G. Yang, X.L. Wang, J. Tian, W. Liu, F. Wu, M. Jiang, H. Wen, Evaluation of reference genes for quantitative real-time RT-PCR analysis of gene expression in Nile tilapia (*Oreochromis niloticus*), Gene. 527 (2013) 183–192. https://doi.org/10.1016/j.gene.2013.06.013.

[33] A. Paria, J. Dong, S.B. Padinhate Purayil, M. Marappan, A. Chaudhari, A.R. Thirunavukkarasu, C. Purushothaman, R. Kv, 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 (2016) 597–605.

[34] Y. Thawornwattana, H.T. Dong, K. Phiwsaiya, P. Sangsuriya, S. Senapin, P. Aiewsakun, Tilapia lake virus (TiLV): Genomic epidemiology and its early origin, Transboundary and Emerging Diseases. 68 (2021) 435–444. https://doi.org/10.1111/tbed.13693.

[35] *Francisella noatunensis* subsp. orientalis, an emerging bacterial pathogen affecting cultured red tilapia (*Oreochromis* sp.) in Thailand - Nguyen - 2016 - Aquaculture Research - Wiley Online Library, (n.d.). https://onlinelibrary.wiley.com/doi/abs/10.1111/are.12802 (accessed April 26, 2023).

[36] H.T. Dong, G.A. Ataguba, P. Khunrae, T. Rattanarojpong, S. Senapin, Evidence of TiLV infection in tilapia hatcheries from 2012 to 2017 reveals probable global spread of the disease, Aquaculture. 479 (2017) 579–583. https://doi.org/10.1016/j.aquaculture.2017.06.035.

[37] V.V. Nguyen, H.T. Dong, S. Senapin, W. Gangnonngiw, N. Pirarat, P. Kayansamruaj, T. Rung-ruangkijkrai, C. Rodkhum, Transmission of *Francisella noatuensis* subsp. *orientalis* from subclinically infected hybrid red tilapia broodstock (*Oreochromis* sp.) to their offspring, Microbial Pathogenesis. 136 (2019) 103670. https://doi.org/10.1016/j.micpath.2019.103670.

[38] P. Kerddee, N. Dinh-Hung, H.T. Dong, I. Hirono, C. Soontara, N. Areechon, P. Srisapoome, P. Kayansamruaj, 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 (2021) 3061–3074. https://doi.org/10.1007/s00705-021-05207-7.

[39] V.V. Nguyen, H.T. Dong, S. Senapin, N. Pirarat, C. Rodkhum, *Francisella noatunensis* subsp. *orientalis*, an emerging bacterial pathogen affecting cultured red tilapia (*Oreochromis* sp.) in Thailand, Aquaculture Research. 47 (2016) 3697–3702. https://doi.org/10.1111/are.12802.

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

Code	Sample_name	Life stage	Country	Year Isolation	Previous Diagnosis	mPCR	Refs/comments
1	Healthy NT 1	juvenile	Thailand	2020	-	-	
2	Healthy NT 2	juvenile	Thailand	2020	-	-	
3	*NT F1**	fingerling	Thailand	2020	TiLV (++)	TiLV*	PeerJ [16]
4	*NT F2**	fingerling	Thailand	2020	TiLV (++)	TiLV*	PeerJ [16]
5	*NT F3**	fingerling	Thailand	2020	TiLV (++)	TiLV*	PeerJ [16]
6	*NT F4**	fingerling	Thailand	2020	TiLV (++)	TiLV*	PeerJ [16]
7	*NT F5**	fingerling	Thailand	2020	TiLV (++)	TiLV*	PeerJ [16]
8	NT fingerling pond C1	fingerling	?		TiLV (+)	-	
9	NT fingerling pond C2	fingerling	?		TiLV (+)	-	
10	NT female brood	adult	?	2020	TiLV (+)	-	
11	NT male brood	adult	?		TiLV (+)	-	
12	NT juvenile 4	juvenile	?		TiLV (+)	-	
13	RT F1	?	?		FnO (+)	-	Nguyen et al. 2015 [35]
14	RT F2	?	?	?	FnO (+)	-	Nguyen et al. 2015 [35]
15	RT F3	?	?		FnO (+)	-	Nguyen et al. 2015 [35]
16	RT F4	?	?		FnO (+)	-	Nguyen et al. 2015 [35]
17	NT 1.2	?	?		TiLV (+)	-	Dong et al. 2017b [36]
18	NT 1.3	?	?		TiLV (+)	-	Dong et al. 2017b [36]
19	NT 2.1	?	?	?	TiLV (+)	-	Dong et al. 2017b [36]
20	NT 2.3	?	?		TiLV (+)	-	Dong et al. 2017b [36]
21	NT 3.1	?	?		TiLV (+)	-	Dong et al. 2017b [36]
22	NT 3.3	?	?		TiLV (+)	-	Dong et al. 2017b [36]
23	NT Gh 21	?	Ghana	2019	ISKNV (++)	ISKNV*	
24	NT Gh 22	?	Ghana	2019	ISKNV (++)	ISKNV*	
25	*NT Gh 23**	juvenile	Ghana	2019	ISKNV (++)	ISKNV*	
26	*NT Gh 24**	adult	Ghana	2020	ISKNV (++)	ISKNV*	
27	*NT Gh 25**	juvenile	Ghana	2021	ISKNV (++)	ISKNV*	
28	RT EX 01	fry	Thailand		FnO (++)	FnO#	
29	RT EX02	fry	Thailand		FnO (++)	FnO#	
30	RT EX03*	fry	Thailand		FnO (++)	FnO#	
31	RT FM1a*	adult	Thailand	2019	FnO (++)	FnO#	Nguyen et al, 2019 [37]
32	RT M1*	adult	Thailand		FnO (++)	FnO	Nguyen et al, 2019 [37]
33	RT FM1b*	adult	Thailand		FnO (++)	FnO	Nguyen et al, 2019 [37]
34	NT PV F8	?	?		Not done	*	
35	NT PV F9	?	?	2020	Not done	#	Nguyen et al, 2019 [37]
36	NT PV F10	?	?		Not done	#	
37	*NT Group B**	fry	Thailand		SAG (++)	SAG	
38	*NT Zone A**	fry	Thailand	2019	SAG (++)	SAG*	
39	*NT Zone B**	fry	Thailand		SAG (++)	SAG	
40	*r1 TILV 1**	fingerling (RT)	Thailand	2018-07	TILV	TILV	Thawornwattana et al, 2021 [34]
41	*r1 ISKNV 1**	fry (AS)	Thailand	2018-11	ISKNV	ISKNV	Kerddee et al, 2021 [38]
42	*r1 FNO 1**	fingerling (RT)	Thailand	2015	FNO	FNO	Nguyen et al, 2016 [39]
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: NT, Nile tilapia, RT, red tilapia, -, 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 orientalis*; SAG, *Streptococcus agalactiae*; 4PAT: mixture of PCR amplicons 4 pathogens(TILV, ISKNV, FNO, SAG). Samples sequenced highlighted in yellow.

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

Pathogen	Samples	Hap	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 FNO371	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

FIGURE LEGENDS

Figure 1. Simplified workflow illustration of diseased fish sampling on farm to Nanopore sequencing results. The entire process takes 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. Equal amounts of each PCR product were combined into lane C to imitate the expected multiplex PCR amplicon products. The band product of *S. agalactiae groEL* was used for positive plasmid construction. -ve, no template control, +ve, mix positive control plasmids; M, 100 bp DNA marker (New England BioLabs).

Figure 3. Percentage of reads mapping to each pathogen after primer trimming and the total filtered read count for each sample. 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 different mapping strategies. 4pat_Pass, Reads that were directly assigned to one of the four pathogens using the reported stringent alignment approach; 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; 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; 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; 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 Figure1. Agarose gel electrophoresis (3.5% gel) of amplified products from single PCR amplification on plasmids of T (TiLV), I (ISKNV), F (*F. orientalis*) and S (*S. agalactiae*). 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 Figure2. 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 Figure3 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 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 Table4. Summary of qPCR testing for the detection and quantification of pathogens in tilapia samples.

Supplemental Table5. 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.

Code	Sample_name
1	Healthy NT 1
2	Healthy NT 2
3	*NT F1**
4	*NT F2**
5	*NT F3**
6	*NT F4**
7	*NT F5**
8	NT fingerling pond C1
9	NT fingerling pond C2
10	NT female brood
11	NT male brood
12	NT juvenile 4
13	RT F1
14	RT F2
15	RT F3
16	RT F4
17	NT 1.2
18	NT 1.3
19	NT 2.1

Code	Sample_name
20	NT 2.3
21	NT 3.1
22	NT 3.3
23	NT Gh 21
24	NT Gh 22
25	*NT Gh 23**
26	*NT Gh 24**
27	*NT Gh 25**
28	RT EX 01
29	RT EX02
30	*RT EX03**
31	*RT FM1a**
32	*RT M1**
33	*RT FM1b**
34	NT PV F8
35	NT PV F9
36	NT PV F10
37	*NT Group B**
38	*NT Zone A**
39	*NT Zone B**

TiLV	13	11	12	10	12	ud	ud	ud	37	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			23	14	18	13	ud	ud	ud
SAG	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	ud	24	25	33

| 52 | 53 | 42 | 41 | 43 | 40 | 3 | 4 | 5 | 6 | 7 | 25 | 26 | 27 | 30 | 31 | 33 | 32 | 37 | 38 | 39 | 51 | 50 | 49 | 48 | 47 | 46 | 45 | 44 | 54 |