Targeted gene sequencing with Nanopore enables rapid and accurate confirmatory diagnostic of Tilapia lake virus

Jerome Delamare-Debouttevillle[1*], Suwimon Taengphu[2], Han Ming Gan[3], Pattanapon Kayansamruaj[4], Partho Pratim Debnath[5], Andrew Barnes[6], Shaun Wilkinson[7,8], Minami Kawasaki[6], Chadag Vishnumurthy Mohan[1], Saengchan Senapin[2,9], Ha Thanh Dong[10*]

[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]GeneSEQ Sdn Bhd, Rawang 48300, Selangor, Malaysia

[4]Center of Excellence in Aquatic Animal Health Management, Faculty of Fisheries, Kasetsart University, Bangkok, Thailand

[5]WorldFish, Khulna, Bangladesh

[6]The University of Queensland, School of Biological Sciences and Centre for Marine Science, Brisbane, Queensland, 4072, Australia

[7]School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand,

[8]Wilderlab NZ Ltd., Wellington, New Zealand

[9]National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani, Thailand

[10]Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok, Thailand

Abstract

Infectious diseases represent one of the major challenges to sustainable aquaculture production. Rapid and accurate diagnosis of emerging pathogens during early-suspected disease cases is critical to facilitate timely response to deploy adequate control measures and prevent or reduce spread. Currently, most laboratories use PCR to amplify partial pathogen genomic regions combined with sequencing of PCR amplicon(s) using conventional Sanger sequencing services for confirmatory diagnosis. The main limitation of this approach is the uncertain turnaround time. Here, we report an innovative approach using Oxford Nanopore Technologies (ONT)-based amplicon sequencing. Using clinical samples from infected fish, we applied this approach for the rapid confirmation (in less than 12 h) of tilapia lake virus (TiLV), a disease-causing virus affecting global tilapia aquaculture. Despite the known high error rate of Nanopore sequencing at the single read level, the obtained consensus sequences after polishing exhibit strikingly high identity to their respective Illumina and Sanger-verified references (99.83-100 % identity). This study suggests that ONT-based amplicon sequencing is a promising platform to deploy in regional aquatic animal health diagnostic laboratories in low and medium income countries, for fast and accurate confirmation and phylogenetic/genotyping of emerging infectious pathogens from field samples within a single day.

Keywords: Nile tilapia (*Oreochromis niloticus*), Red tilapia (*Oreochromis* spp.), Tilapia lake virus, semi-nested RT-PCR, Oxford Nanopore Technologies, diagnostic by sequencing, bioinformatics and phylogeny analyses

1. Introduction

Aquaculture is one of the fastest growing food production sectors and is of increasing importance to global food security, particularly in low income, food deficit countries, where it plays a significant role in livelihood and subsistence. However, the sustainability and expansion of the sector is hampered by disease epidemics. Endemic and emerging infectious diseases (Brummett et al., 2014), pose major animal health issues and economic losses, affecting millions of smallholders (Subasinghe et al., 2019), (FAO, 2020).

Tilapia are the second most important aquaculture species (in volume) produced globally, with an industry value of $9.8 billion annually (FAO, 2020). Intensification of tilapia production has driven the emergence of diseases through the translocation of asymptomatically infected animals (Rodgers et al., 2011), (Jansen et al., 2019), (Dong, Ataguba, et al., 2017).

Rapid and accurate diagnostic capacity of aquatic pathogens is a central pillar to any successful national aquatic animal health strategy, helping key aquaculture value chain actors to select disease-free fish broodstock, conduct pathogen surveillance, confirm the aetiological agent of disease outbreaks and prevent their further spread to neighbouring farms, regions and countries. This is especially important for viruses considering the lack of completely effective prophylactic treatments and vaccines for most viral pathogens of fish (Crumlish, 2017), (Ninawe et al., 2017), (Assefa & Abunna, 2018).

On suspicion of viral disease, the first recommended procedure is to demonstrate clinical pathology via simple observations of abnormal behaviours and external/internal clinical signs, then to confirm the presence of viable viral particles in clinical samples using susceptible cell line culture, which can take days to weeks for a confirmatory diagnosis.

For farmed aquatic animals, molecular techniques (e.g. PCR) to confirm the presence of viral nucleic acids (DNA/RNA) are preferred because they yield much faster presumptive diagnoses. These specific viral PCRs followed by Sanger sequencing are routine in many laboratories, usually taking a few days from sample to sequence results due to scarcity of facilities, with associated transport and queueing times. Unfortunately, in many low and middle-income countries (LMICs), clinical samples from disease outbreaks have to be sent overseas due to lacking of locally available sequencing capacity or limited access to specialist laboratories.

While Sanger sequencing remains the current preferred sequencing platform to produce accurate short read sequence data, it is time consuming and depends on the availability and accessibility

of Sanger's sequencing machine where needed. In addition, its analysis is somewhat laborious and may require manual inspection of the chromatogram. Second and third generation sequencing platforms such as Ion Torrent, Illumina and PacBio are extremely powerful for genomic analysis of aquatic pathogens, but require substantial capital investment and major laboratory infrastructure. Nevertheless, they have been used to study viruses affecting global fish aquaculture, such as TiLV, piscine reovirus (PRV), piscine myocarditis virus (PMCV), salmonid alphavirus (SAV), infectious salmon anaemia virus (ISAV) (Gallagher et al., 2018), (Nkili-Meyong et al., 2016).

The MinION/Flongle sequencing platform from Oxford Nanopore Technologies offers a simple low-cost portable device for generating real-time sequence data. The low equipment cost, and particularly the lack of requirement for a well-equipped laboratory facility, makes MinION particularly attractive for sequence data-driven management and control of aquatic pathogens in remote locations in LMIC. In this study, we explored ONT-based amplicon sequencing for rapid confirmation and genotyping of TiLV, the causative agent of syncytial hepatitis of tilapia, a disease affecting tilapia aquaculture in over 16 countries (Taengphu et al., 2020).

TiLV is an enveloped, negative-sense, single-stranded RNA virus that contains 10 genomic segments ranging from 465 to 1641 bp, with a total genome size of 10,323 bp (Bacharach et al., 2016), encoding 14 predicted proteins (Acharya et al., 2019). The virus was recently re-classified as a new genus *Tilapinevirus*, the sole genus under the new family *Amnoonviridae* in the order *Articulavirales* (International Committee on Taxonomy of Viruses (ICTV, 2019). The latter order also contains another family, *Orthomyxoviridae*, with several members including influenza viruses and Infectious Salmon Anaemia virus (ICTV, 2019).

When new viruses (such as TiLV) emerge in aquaculture, non-validated PCR and RT-PCR methods appear very quickly after first detection of the viral diseases. Several TiLV PCR detection assays have been developed, including RT-PCR (Eyngor et al., 2014), nested RT-PCR (Kembou Tsofack et al., 2017), semi-nested RT-PCR (Taengphu et al., 2020) (Dong, Siriroob, et al., 2017), (Castañeda et al., 2020), RT-qPCR (Kembou Tsofack et al., 2017), (Tattiyapong et al., 2018), (Waiyamitra et al., 2018) and RT-LAMP (Phusantisampan et al., 2019), (Yin et al., 2019). However with no validated OIE approved assays, sequencing of amplicons can provide robust supporting evidence that the disease has been detected. For this study, we selected a semi-nested RT-PCR method (Taengphu et al., 2020) targeting TiLV segment 1, for its sensitivity reported to

be 100 times more than a previous TiLV segment 3-based protocol (Dong, Siriroob, et al., 2017), and because it can be used for preliminary genotyping of TiLV (Taengphu et al., 2020).

Here, we report successful use of that semi-nested RT-PCR coupled with Nanopore sequencing of amplicons for rapid confirmatory diagnostic and genotyping of TiLV. We also discuss the range of possible practical applications and implications of Nanopore sequencing, as a portable platform for robust molecular field diagnostics investigations into the origin and spread of other aquaculture pathogens of economic significance.

2. Materials and Methods

2.1. Workflow. The diagnostic workflow from sample collection from farmed moribund fish, extraction of nucleic acid, semi-nested RT-PCR, library preparation, Nanopore sequencing and data analysis is described in Figure 1.

2.2. RNA samples and reference sequences

We used five archived RNA templates extracted from clinical Nile tilapia (*Oreochromis niloticus*) and red tilapia (*Oreochromis* spp.) specimens and from TiLV infected E-11 cell culture (Table 1). All five samples were previously confirmed to be TiLV positive. The samples were originally isolated from specimens collected in Thailand (BC01 and BC03), Bangladesh (BC02), and Peru (BC04 and BC05) as described in previous reports (Debnath et al., 2020) (Taengphu et al., 2020). Table 1 also includes eleven full-length TiLV segment 1 reference sequences retrieved from NCBI. The NCBI reference sequences originated from tilapia specimens collected from Thailand, Bangladesh, Peru, Ecuador, Israel and the USA between 2011 and 2018, and were used for sequence alignment and phylogenetic analysis with the amplicon consensus sequences generated in this study.

2.3. Semi-nested RT-PCR

Partial regions of the TiLV segment 1 genome were amplified by semi-nested RT-PCR as described previously (Taengphu et al., 2020). Five microliters of the second round PCR products were analyzed by electrophoresis on a 1% agarose gel stained with ethidium bromide solution. The remaining 20 µL reaction volume from the second round PCR was purified for each sample on a NucleoSpin Gel and PCR Clean□up column (Macherey-Nagel) and eluted with 20 µL of the kit elution buffer (5 mM Tris-HCl, pH 8.5). The purified products were quantified using Qubit dsDNA Broad Range kit (Qiagen) with a Qubit 3.0 fluorometer prior to Nanopore multiplex library preparation.

2.4. Library preparation of TiLV PCR products for Nanopore sequencing. To prepare the TiLV library, the ligation sequencing kit (SQK-LSK109) and the native barcoding expansion 1-12 kit (EXP-NBD104) were used according to the Oxford Nanopore Technologies (ONT) standard protocols adapted for the Flongle flow cell. We used 250 ng PCR products for each sample (BC01-BC05), one unique barcode (BC) per sample and washed samples with the Short Fragment Buffer (SFB) just before the elution step at the end of the protocol. DNA concentrations were determined between each step using the Qubit assay. TiLV libraries were

loaded as per the standard protocol onto a Flongle flow cell (FLO-FLG106) with 29 active pores fitted to a Flongle adapter (FLGIntSP) for MinION.

2.5. Data acquisition and basecalling. Control of the MinION and high accuracy base-calling data acquisition were performed offline in real-time using the stand-alone MinIT (MNT-001): a pre-configured compute module with MINKNOW software version (19.05.02). The raw Fast5 files were subsequently re-basecalled and demultiplexed using the latest Guppy version (v.4.4.1) in high accuracy mode to improve some homopolymer base-calling errors.

2.6. Bioinformatics analyses for TiLV amplicons consensus sequences generation. The base-called and demultiplexed FastQ files were individually assessed using NanoStat. Raw reads were aligned to a primer-trimmed TILV Segment 1 gene region (Accession Number: MN687685.1) using Minimap2 v2.17 (-ax map-ont –secondary=no). Reads with alignment length of more than 550 bp were subsequently filtered out for consensus generation, since they were assumed to have been generated from the sequencing of the 620 bp PCR product. Briefly, the filtered reads were re-aligned to the reference sequence using Minimap v2.17 followed by one round of polishing with RACON v1.4.20 (-m 8 -x -6 -g -8 -w 250) and then Medaka_consensus v1.1.3 (-m r941_min_high_g360). To examine the effect of read depth and sequencing time on consensus accuracy for fast confirmation of PCR products, raw reads aligned to the 274 bp reference region (semi-nested PCR target fragment), were filtered and subsampled randomly for 1000, 500, 100, and 50 reads with seqtk v1.2 followed by the same consensus generation steps as described above. Pair-wise nucleotide similarity of the consensus sequences against their respective reference sequences were calculated using NCBI BlastN. To obtain the number of reads sequenced over time, ‘Sequencing start time’ was extracted from every sequence identifier using grep and cut commands. The extracted data were used to generate histograms representing the number of reads generated every 5 min.

2.7. Alignment of TiLV segment 1 amplicon consensus sequences to public references for phylogenetic analysis. A total of 24 TiLV segment 1 sequences, five consensus sequences derived from semi-nested products (274bp), and five consensus sequences derived from the first round products (620bp) were aligned with 14 full-length TiLV segment 1 reference sequences retrieved from GenBank database (Table 1) in Jalview (Waterhouse et al., 2009) using the web service Muscle v3.8.31 (web service) defaults parameters (Edgar, 2004). The 5’ and 3’ non-aligned regions and primer binding sites were trimmed (577 and 231 bp, respectively) and the

trimmed alignment was subsequently used for phylogenetic tree construction. The phylogenetic tree was inferred by Maximum Likelihood in IQ-TREE (v.1.6.12) based on the nucleotide alignment of consensus TiLV gene fragments (577 and 231 bp) generated in this study and TiLV Segment 1 gene reference sequences (1560 bp) from previous studies (Table 1).

3. Results

3.1. TiLV positive clinical samples confirmed by PCR

The segment 1 semi-nested PCR assay confirmed that the five samples used in this study (Table 1) were TiLV positive (Figure S1). Bands at 620 bp are the product of the first round RT-PCR, and bottom bands at 274 bp are the product of the second round semi-nested PCR. A band at 1-1.1 kb is the product of cross-hybridization of the amplified products and was strongly amplified in heavily infected samples (BC01 and BC03), less apparent in medium infected samples (BC02 and BC05) and absent from BC04 sample collected from apparently healthy Nile tilapia with low viral load (Figure S1).

3.2. Sequencing output

The sequencing run on the Flongle flow cell generated 174.69K reads with 114.99 Mb of estimated bases and 93.53 Mb base called. Depending on the sample, 517 to 964 reads were generated in the first 5 min of the run (Figure S2). Those numbers gradually decreased with reduction of available active sequencing pores to drop on average below 116 reads per sample after 4 h, 15 reads per sample after 5h and no more reads produced past 6 h of the sequencing run (Figure S2). The number of reads sequenced over time will vary depending on flow cell type (Minion vs Flongle), flow cell pore count, library quality, and amplicon size. Histograms of the read length distribution—for all five samples—indicate two main peaks at 620 bp and 274 bp (Figure S3). BC01 and BC02 had a higher peak at 620 bp and BC03, BC04 and BC05 at 274 bp. Our PCR results and sequence data both confirmed the semi-quantitative nature of this (ONT)-based amplicon sequencing approach that can differentiate between heavily, medium and low TiLV infected samples (Figure S1 and S3).

3.3. Raw Nanopore error rate

Before the trimming, filtering and polishing steps, the average percentage identity of raw Nanopore reads against Sanger TiLV segment 1 references varied between 92.5 and 93.2%, which correspond to 6.8 and 7.5% base-calling error rate, respectively (Table S1 and Table S2).

3.4. Accurate consensus generation for TiLV identification

All consensus sequences (577 and 231 bp) generated in this study have been successfully identified as TiLV (Table 2). Out of the five samples, only the Thai BC03 and Peruvian BC04 had their full length TiLV segment 1 region (1560 bp) previously Sanger sequenced: TH-2018-N and PE-2018-F3-4, respectively (Table 1). We confirmed 100% nucleotide identity of the 577 bp amplicon of the Thai BC03 and Peruvian BC04 consensus to their original references, (Table 2A). The Thai BC01 (from E-11 cell culture with Nile tilapia infected tissue) was also 100% identical to BC03 (isolated from red tilapia) but BC01 came from a different farm seven months later, suggesting that this variant is capable of infecting multiple species in different farming areas of Thailand.

The 577 bp BC02 Bangladeshi consensus sequence was 99.83% identical to BD-2017-181 (Table 2A). The single SNP (A instead of G) in position 334 (Figure 2A) was further assessed in Integrative Genomics Viewer (IGV) using BC02.medaka.bam file (read depth) with final BC02.medaka.fasta sequence. The SNP was confirmed to be amplicon-specific, partitioned between 274 and 620 amplicons (Figure 2C). Full summary of sequencing statistics for mixed amplicons (274 and 620 bp) derived from NanoStat (De Coster et al., 2018) can be found in Table S1.

A BlastN analysis of the Peruvian 577 bp BC05 consensus sequence returned 99.83% identity to PE-2018-F3-4 (Table 2A), with alignment of BC05 and PE-2018-F3-4 showing only one SNP (A instead of a G) in position 347 (Figure 2B). This SNP was confirmed in IGV, which revealed consistent base call of an Adenine (A) in the majority of the reads (BC05.consenus.bam file) with only one Guanine (G) corresponding to a homopolymer base calling error (Figure 2D). While both BC04 and BC05 were collected in 2018, they came from different farms. This indicates the presence in Peru of at least two TiLV variants at the time of sampling.

3.5. Optimal sequencing coverage

Random subsampling of 1000, 500, 100 and 50 reads from the pool of 274 bp amplicons did not affect percentage identity of consensus sequences with Sanger references, which remained 100% identical in all cases (Table 2B). NanoStat summary of sequencing output for 274 bp and sub-sampling analysis are presented in Table S2.

3.6. Phylogenetic analysis of TiLV segment 1 amplicon consensus

The 577 bp and 231 bp consensus sequences from this study were compared with 14 full-length (1560 bp) TiLV segment 1 sequences retrieved from NCBI (Table 1). Our consensus sequences clustered into two separate clades, namely C1 and C2, also known as the Thai and Israeli 2011 clades, respectively (Figure 3). Clade C1 was divided into two sub-clades, C1a and C1b. Clade C1a contains BC01 and BC03 Thai isolates both clustering closely with TH-2018-N. Clade C1b includes BC02 which is most similar to BD-2017-181. The Israeli 2011 clade (C2) comprises BC04 and BC05 Peruvian isolates clustering with PE-2018-F3-4. IL-2011-Til-4-2011 forms a monophyletic clade outside of "Israel 2011" clade (Figure 3).

4. Discussion

Semi-nested PCR methods developed for TiLV segment 1 (Taengphu et al., 2020), TiLV segment 3 (Dong, Siriroob, et al., 2017) (Castañeda et al., 2020), koi herpesvirus (Bergmann et al., 2010) and scale drop disease virus (SDDV) (Charoenwai et al., 2019) (Kerddee et al., 2020) have been considered semi-quantitative where one, two or three bands on the gel can be indicative of low, medium, and high viral load, respectively.

In this study, the base-calling error rate of raw Nanopore reads was below 8%. The base-calling accuracy of MinION data remains a disadvantage of the platform when compared against Sanger or Illumina sequencing. However, it is improving with advances in flow cell chemistry and base-calling software, for example, five years ago, the MinION base-calling error rate was between 20-30% (Laver et al., 2015).

While we cannot ascertain if the SNP identified in Bangladeshi BC02 is real, it could be genuine variation from the viral population sequenced. BC02 was collected on the same farm at the same time but not from the same diseased fish that was used to derive the whole genome of BD-2017-181: one of the only four publicly available TiLV segment 1-reference sequences from Tilapia in Bangladesh (Debnath et al., 2020). We know that viral RNA-dependent RNA polymerases are error-prone, with misincorportation of a wrong nucleotide estimated every 10,000-1,000,000 nucleotides polymerized depending of viral species (Sanjuán et al., 2010). This high rate of mutation comes from the lack of proofreading ability in RNA polymerases (Steinhauer et al., 1992). Given the size of the TiLV RNA genome of 10,323 bases, a mutation rate of 1 in 10,000

would mean an average of 1 mutation in every replicated genome. If a single tilapia cell is infected with TiLV and produces 10,000 new viral particles, this mutation frequency means in theory that about 10,000 new TiLV variants have been produced. This incredible high mutation rate explains why RNA viruses evolve so quickly. Viral populations even in a single infection are not homogeneous and will be mixed at any point in time during the infection. What is sequenced from the PCR is usually an amplification of the most populous variant at the time sampled with the additional stochastic effect of which templates amplify in the first few rounds of the PCR, plus the possible (but rare) misincorporation of a dNTP by the PCR polymerase early in the amplification. The SNP in BC02 could be real but would need to be confirmed by Sanger sequencing.

For rapid diagnostic of TiLV, one would need at least 50 X 274 bp reads to enable accurate TiLV identification. Real-time analysis of base-called and demultiplexed Nanopore barcoded reads will allow estimation of the minimum sequencing time (or number of reads) required to achieve a positive identification, which should be occurring in just a few minutes depending on the number of samples being sequenced, flow cell pore occupancy, library preparation quality and computing capability. A study sequencing the complete genome of salmonid alphavirus (SAV1) by Nanopore reported similar low sequence coverage to generate highly accurate consensus (Gallagher et al., 2018), where authors needed as little as 20 X coverage to get a consensus 99% similar to Sanger reference, while 1,000 X coverage lonely led to 99.97% similarity.

Our tree is in agreement with the literature, since it classifies the Thai and Bangladeshi consensus (BC01, BC02 and BC03) into the correct “Thai” clade/genotype and the Peruvian isolates (BC04 and BC05) into the “Israeli 2011” clade with IL-2012-AD-2016 remaining a monophyletic group referred as the “Israeli 2012” clade (Debnath et al., 2020) (Taengphu et al., 2020). The difference observed with IL-2011-Til-4-2011 forming a monophyletic clade outside the previously described “Israel 2011” clade could be explained by the different sequence lengths used between studies. Here we only used short amplicons (231 and 577 bp) as opposed to the full-length segment 1 sequence (1560 bp) as used in the other two studies (Debnath et al., 2020) (Taengphu et al., 2020).

While short amplicons seem suitable for preliminary TiLV genotyping, a recent study analyzed each individual TiLV genome segment separately, resulting in different phylogenetic trees with high estimation uncertainties (Chaput et al., 2020). The authors' suggested exercising caution when using phylogenetic analysis to infer geographic origin and track the movement of TiLV, and recommend using whole genomes for phylogeny wherever possible. To avoid having to sequence complete viral genomes, further sequencing data may be enough to identify regions of the genome that are descriptive—similar to multi-locus sequence typing scheme used to identify prokaryote lineages. Another good example on the need for complete genomic sequences has recently been described in a study conducted by (Thawornwattana et al., 2020), which looked at eight TiLV complete genomes from Thailand collected between 2014 and 2019. Those genomes were analyzed by Bayesian inference allowing for estimation of virus evolutionary timescales, rates and global population dynamics since the early origin of TiLV. This was only possible using complete genomic sequences.

The inherent nature of segmented virus such as TiLV limits one of the main benefits of Nanopore sequencing, which is to generate a complete viral genome with a few small overlapping PCR amplified regions. Salmonid alphavirus (SAV), a ~12 kb non-segmented, single-stranded, positive-sense RNA virus is the only fish virus genome successfully sequenced by Nanopore and was confirmed for assembly accuracy against Sanger verified reference sequence (Gallagher et al., 2018). To date, the 19 complete genomes of TiLV have been sequenced by Sanger (Debnath et al., 2020), (Thawornwattana et al., 2020) and, Illumina (Chaput et al., 2020), (Subramaniam et al., 2019), (Al-Hussinee et al., 2018), but none have been sequenced by Nanopore. To achieve this will require amplifying all 10 segments individually by RT-PCR using different primer pairs and cycling conditions and we accept that this process may be time-consuming and possibly challenging given the relatively high nucleotide divergence among TiLV strain from different lineages.

The choice of whether to sequence short amplicons, entire segments or the whole genome of TiLV will depend on the specific need. For simple and rapid confirmatory diagnosis with some phylogeny inferences, we have shown that using 274 and 620 bp amplicons from TiLV segment 1 works very well but for high-resolution epidemiological and evolutionary analyses a whole genome approach would be required.

In conclusion, PCR-MinION combined with our optimized bioinformatics pipeline is a rapid method to generate accurate consensus sequences for TiLV identification and genotyping. This method currently takes less than 12h from clinical samples to sequence results. We show that low read depth (or coverage) does not affect the accuracy of 274 bp consensus generation, hence the possibility to further reduce sequencing time. In the hands of trained and skilled molecular diagnosticians and aquaculture technicians, this device with the specific sample preparation protocols and our analytical workflow will enable point-of-care testing and sequencing in remote locations, helping teams of governmental and supra-national institutions during disease outbreak investigations. Such application of Nanopore has been successfully applied to study human epidemics such as Ebola virus in remote areas of West Africa (Hoenen et al., 2016), the Zika virus in hard to reach regions of Brazil (Faria et al., 2016). More recently, the technology was used to sequence and identify SARS-CoV-2, the virus causing the COVID-19 pandemic (Wang et al., 2020).

In aquatic animal production systems, our approach can offer a rapid deployable mobile solution for early identification of TiLV and other newly emerging infectious diseases of economics importance. This will help competent authorities to empower legislative decision making on biosecurity protocols, define precise movement controls of aquatic animals, and provide recommendations to farmers to take appropriate actions. Our multiplex approach offers scalability and opportunity to reduce per-sample costs even further. Deployed to a broader scale in national reference and regional laboratories in LMIC, this platform can be used to screen clinical samples from routine surveillance programs and disease outbreak investigations helping to minimize the introduction, and spread of TiLV and other infectious diseases of farmed aquatic animals, contributing to both economic and food security.

Tables

TABLE 1 Metadata of archived RNA samples (No. 1-5) extracted from clinically sick tilapia specimens and from TiLV infected E-11 cell culture (all previously confirmed to be TiLV positive). Those that came from Thailand, Bangladesh and Peru (BC01-BC05) were used in this study for TiLV segment 1 RT-PCR amplification and Nanopore sequencing. Eleven TiLV full length segment 1 reference sequences: samples (No. 6-19) from diseased tilapia were retrieved from NCBI. Note that sample No.3 originated from same specimen as reference TH-2018-N (No. 11) and sample No. 4 same specimen as reference PE-2018_F3-4 (No. 13).

No.	Sample code	Collection date	Origin	Fish host	NCBI Accession no.	References
1	BC01 Cell line day 4	2019	Thailand	Nile tilapia	Not done	This study
2	BC02 Ti Bang 176-1	2017	Bangladesh	Nile tilapia	Not done	
3	BC03 S1-18	2018	Thailand	RT fingerling	TH-2018-N (MN687745.1)	
4	BC04 m Peru 2018 F3-4	Feb 2018	Peru	Nile tilapia	PE-2018_F3-4 (MK425010.1)	
5	BC05 O Peru 2018 F4-5	Feb 2018	Peru	Nile tilapia	Not done	
6	IL-2011-Til-4-2011	May 2011	Israel	Tilapia	KU751814.1	(Eyngor et al., 2014) (Bacharach et al., 2016)
7	IL-2012-AD-2016	Aug 2012	Israel	Hybrid tilapia	KU552131.1	NCBI
8	TH-2016-TV7	May 2016	Thailand	Nile tilapia	KX631936.1	(Surachetpong et al., 2017)
9	EC-2012	Jul 2012	Ecuador	Nile tilapia	MK392372.1	(Subramaniam et al., 2019)
10	TH-2018-K	Aug 2018	Thailand	NT juvenile	MN687755.1	(Thawornwattana et al., 2020)
11	TH-2018-N	Jul 2018	Thailand	RT fingerling	MN687745.1	
12	TH-2019	Feb 2019	Thailand	NT fingerlings	MN687765.1	
13	PE-2018-F3-4	Feb 2018	Peru	Nile tilapia	MK425010.1	(Pulido et al., 2019)
14	BD 2017	Jul 2017	Bangladesh	Nile tilapia	MN939372.1	(Chaput et al., 2020)
15	BD-2017-181	2017	Bangladesh	Nile tilapia	MT466437.1	(Debnath et al., 2020)
16	BD-2019E1	2019	Bangladesh	Nile tilapia	MT466447.1	
17	BD-2019-E3	2019	Bangladesh	Nile tilapia	MT466457.1	
18	USA-2019-WVL19054	2019	USA	Nile tilapia	MN193523-1	(Al-Hussinee et al., 2018)
19	USA-2019-WVL19031	Nov 2018	USA	Nile tilapia	MN193513.1	

BC: barcode from Nanopore barcoding expansion 1-12 kit (EXP-NBD104); country codes: TH, Thailand, IL, Israel, EC, Ecuador, PE, Peru, and BD, Bangladesh. Animal codes: RT: red tilapia (*Oreochromis* spp.); NT: Nile tilapia (*Oreochromis niloticus*); Hybrid tilapia (*Oreochromis niloticus x Oreochromis aureus*)

TABLE 2 BlastN results of consensus sequences using (A) full set of mixed amplicons (620 and 274 bp); (B) full set and sub-sampling (1k, 500, 100, 50 reads) of 274 bp amplicons - results were the same between the full set and sub-samples and are hence shown once only. Percentage (%) identity of consensus sequences to their closest Sanger reference sequences. *query length of medaka consensus sequences with their primer-binding sites trimmed; † samples not previously Sanger sequenced.

(A)

	*Query length (bp)	Identities (bp/bp)	% Identity	Gaps	Sanger references length (bp)	Sanger references	NCBI Accession no.
BC01 †	577	577/577	100	0/577	1560	TH-2018-N	MN687745.1
BC02 †	577	575/576	99.83	0/576	1560	BD-2017-181	MT466437.1
BC03	577	577/577	100	0/577	1560	TH-2018-N	MN687745.1
BC04	577	577/577	100	0/577	1560	PE-2018-F3-4	MK425010.1
BC05 †	577	576/577	99.83	0/577	1560	PE-2018-F3-4	MK425010.1

(B)

	*Query length (bp)	Identities (bp/bp)	%Identity	Gaps	Sanger references length (bp)	Sanger references	NCBI Accession no.
BC01†	231	231/231	100	0/231	1560	TH-2018-N	MN687745.1
BC02†	231	230/230	100	0/230	1560	BD-2017-181	MT466437.1
BC03	231	231/231	100	0/231	1560	TH-2018-N	MN687745.1
BC04	231	230/230	100	0/230	1560	PE-2018-F3-4	MK425010.1
BC05†	231	230/230	100	0/230	1560	PE-2018-F3-4	MK425010.1

Figures legends

FIGURE 1 Overall workflow from sample collection of diseased fish on farm to sequence results. The entire process takes less than 12 hours. * DNA repair, end-preparation, multiplex native barcode and adapter ligation

FIGURE 2 (A and B) Identification of single nucleotide polymorphisms (SNPs) using sequences alignment of TiLV segment 1 medaka consensus sequences (this study) with their closest Sanger verified references. A. Bangladeshi BC02 consensus (576 bp) aligned with BD-2017-181 (MT466437.1) showing SNP in position 334 (red arrow); B. Peruvian BC05 consensus (577 bp) aligned with PE-2018-F3-4 (MK425010.1) with SNP in position 347 (green arrow). (C and D) SNPs examination in Integrative Genomics Viewer (IGV) (version 2.8.10). C. Read depth (BC02.medaka.bam file) aligned with final medaka consensus sequence (BC02.medaka.fasta file) showing the SNP is partitioned between 274 and 620 bp amplicons; D. Read depth (BC05.medaka.bam file) aligned with final medaka consensus sequence (BC05.medaka.fasta file) confirming the SNP is real since it is identical in 97% of the reads, except for 1 homopolymer base-call error (G).

FIGURE 3 Maximum likelihood tree constructed in IQ-TREE based on the nucleotide sequences alignment of short TiLV consensus (620 bp and 274 bp) with their primer binding sites trimmed (577 bp and 231 bp) and full-length TILV Segment 1 gene reference sequences (1560 bp) retrieved from NCBI (Table 1). Tree rooted with IL-2011-Til-4-2011 as outgroup. The branch lengths indicate the number of substitutions per site, and node labels indicate bootstrap support values in percentage.

Note: figures supplied as separate files

Supplementary Material

Supplementary tables legends

TABLE S1 NanoStat summary statistics of analysis of mixed amplicons (620 and 274 bp) for each sample (BC01-05) using the full set without sub-sampling. *read depth after clustering and filtering steps; **mean percent identity of Nanopore raw reads to each sample specific reference; ***during basecalling; BC: barcode

TABLE S2 NanoStat summary statistics of analysis of 274 bp amplicons for each sample (barcode01-05) using the full set and with sub-sampling (sub1K (1000), 500, 100, or 50 reads)

Note: supplementary tables supplied as separate files

Supplementary Figures legends

FIGURE S1 (A) Original 1% agarose gels showing detection of TiLV segment 1 from five samples used in this study: 1. BC01 (Thailand), 2. BC02 (Bangladesh), 3. BC03 (Thailand), 4. BC04 (Peru), 5. BC05 (Peru); other samples (a to i) were not included in this study. Gels were stained with ethidium bromide solution. M, 2-Log DNA marker (New England Biolabs); Ng, negative control. Expected band size of 620 bp and 274 bp represent amplicons from respective first and semi-nested PCR, respectively, with lanes marked +++ for heavy infection, ++ for medium infection and + for a light infection. The band marked with # on the right side of gels arose from cross hybrization of the amplified products. (B) Composite image based on snapshots taken from (A) only showing the five samples used in this study.

FIGURE S2 Histograms of sample read counts generated every 5 min for 6 h. A. BC01, B. BC02, C. BC03, D. BC04, E. BC05.

FIGURE S3 Distribution of sequence lengths over all sequences between 50-1399 bp from the mixed amplicons analysis (620 and 274 bp). A. BC01, B. BC02, C. BC03, D. BC04, E. BC05. Blue arrows indicate the peaks for 274 bp amplicons and green arrows the peaks for 620 bp amplicons. Sequence length distribution obtained in FastQC High Throughput Sequence QC Report (v 0.11.9) using final medaka.bam file as inputs. Figure produced in GraphPad Prism 9.0.2; bp: base-pair

Note: supplementary figures supplied as separate files

Acknowledgments

This work was undertaken as part of the CGIAR Research Program on Fish Agri-Food Systems (FISH) led by WorldFish, and the CGIAR Big Data Platform Inspire Challenge 2019 led by CIAT and IFPRI. These programs are supported by contributors to the CGIAR Trust Fund. The funders provided support in the form of salary for authors [J.D.D; P.P.D; C.V.M], travels, laboratory consumables and analytical costs, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization, J.D.D., S.S., H.T.D.; investigation, S.T., J.D.D., P.P.D., H.T.D., S.S.; formal analysis, J.D.D., H.M.G., P.K.; methodology, S.S., H.T.D., S.T., J.D.D., H.M.G.; P.K.; supervision; S.S., H.T.D., J.D.D.; writing - original draft, J.D.D., H.T.D.; review & editing, S.T., P.K., H.M.G., P.P.D., M.K., A.B., S.W., C.V.M., S.S. All authors have read and agreed to the current version of the manuscript.

Data Availability Statement

The data that support the findings of this study are available at the following links: demultiplexed FastQ files for all five samples can be found under Bioproject PRJNA703741, sequence read archive (SRA) SRR13766240, and BioSample accessions SAMN18024369 (BC01), SAMN18024370 (BC02), SAMN18024371 (BC03), SAMN18024372 (BC04), SAMN18024373 (BC05). The intermediate bioinformatics files (medaka.bam; medaka.bam.stats) and final consensus sequences (medaka.fasta) from partial TiLV segment 1 amplicons combined analysis (620 bp and 274 bp) and random 274 bp analysis with subsamples for 1000, 500, 100, and 50 reads, with reference fasta sequences used for both analyses can be found under Zenodo.org dataset DOI 10.5281/zenodo.4556414.

Conflict of interest statement

The authors declare no conflict of interest.

References

Acharya, V., Chakraborty, H. J., Rout, A. K., Balabantaray, S., Behera, B. K., & Das, B. K. (2019). Structural Characterization of Open Reading Frame-Encoded Functional Genes from Tilapia Lake Virus (TiLV). *Molecular Biotechnology*, *61*(12), 945–957. https://doi.org/10.1007/s12033-019-00217-y

Al-Hussinee, L., Subramaniam, K., Ahasan, M. S., Keleher, B., & Waltzek, T. B. (2018). Complete Genome Sequence of a Tilapia Lake Virus Isolate Obtained from Nile Tilapia (*Oreochromis niloticus*). *Genome Announcements*, *6*(26), e00580-18, e00580-18. https://doi.org/10.1128/genomeA.00580-18

Assefa, A., & Abunna, F. (2018). *Maintenance of Fish Health in Aquaculture: Review of Epidemiological Approaches for Prevention and Control of Infectious Disease of Fish* [Review Article]. *Veterinary Medicine International*; Hindawi. https://doi.org/10.1155/2018/5432497

Bacharach, E., Mishra, N., Briese, T., Zody, M. C., Kembou Tsofack, J. E., Zamostiano, R., Berkowitz, A., Ng, J., Nitido, A., Corvelo, A., Toussaint, N. C., Abel Nielsen, S. C., Hornig, M., Del Pozo, J., Bloom, T., Ferguson, H., Eldar, A., & Lipkin, W. I. (2016). Characterization of a Novel Orthomyxo-like Virus Causing Mass Die-Offs of Tilapia. *MBio*, *7*(2), e00431-16, /mbio/7/2/e00431-16.atom. https://doi.org/10.1128/mBio.00431-16

Bergmann, S. M., Riechardt, M., Fichtner, D., Lee, P., & Kempter, J. (2010). Investigation on the diagnostic sensitivity of molecular tools used for detection of koi herpesvirus. *Journal of Virological Methods*, *163*(2), 229–233. https://doi.org/10.1016/j.jviromet.2009.09.025

Brummett, R. E., Alvial, A., Kibenge, F., Forster, J., Burgos, J. M., Ibarra, R., St-Hilaire, S., Chamberlain, G. C., Lightner, D. V., Khoa, L. V., Hao, N. V., Tung, H., Loc, T. H.,

Reantaso, M., Wyk, P. M. V., Chamberlain, G. W., Towner, R., Villarreal, M., Akazawa, N., … Nikuli, H. L. (2014). *Reducing disease risk in aquaculture. Agriculture and environmental services discussion paper 9. World Bank report number 88257-GLB, 119 pp.* http://documents.worldbank.org/curated/en/110681468054563438/Reducing-disease-risk-in-aquaculture

Castañeda, A. E., Feria, M. A., Toledo, O. E., Castillo, D., Cueva, M. D., & Motte, E. (2020). Detection of tilapia lake virus (TiLV) by semi nested RT-PCR in farmed tilapias from two regions of Peru. *Revista de Investigaciones Veterinarias del Perú (RIVEP)*, *31*(2). https://www.cabdirect.org/cabdirect/abstract/20203285327

Chaput, D. L., Bass, D., Alam, Md. M., Al Hasan, N., Stentiford, G. D., van Aerle, R., Moore, K., Bignell, J. P., Haque, M. M., & Tyler, C. R. (2020). The Segment Matters: Probable Reassortment of Tilapia Lake Virus (TiLV) Complicates Phylogenetic Analysis and Inference of Geographical Origin of New Isolate from Bangladesh. *Viruses*, *12*(3). https://doi.org/10.3390/v12030258

Charoenwai, O., Meemetta, W., Sonthi, M., Dong, H. T., & Senapin, S. (2019). A validated semi-nested PCR for rapid detection of scale drop disease virus (SDDV) in Asian sea bass (*Lates calcarifer*). *Journal of Virological Methods*, *268*, 37–41. https://doi.org/10.1016/j.jviromet.2019.03.007

Crumlish, M. (2017). Bacterial Diagnosis and Control in Fish and Shellfish. In *Diagnosis and Control of Diseases of Fish and Shellfish* (pp. 5–18). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119152125.ch2

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

Debnath, P. P., Delamare□Deboutteville, J., Jansen, M. D., Phiwsaiya, K., Dalia, A., Hasan, M. A., Senapin, S., Mohan, C. V., Dong, H. T., & Rodkhum, C. (2020). Two-year surveillance of tilapia lake virus (TiLV) reveals its wide circulation in tilapia farms and hatcheries from multiple districts of Bangladesh. *Journal of Fish Diseases*, *43*(11), 1381–1389. https://doi.org/10.1111/jfd.13235

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

Dong, H. T., Siriroob, S., Meemetta, W., Santimanawong, W., Gangnonngiw, W., Pirarat, N., Khunrae, P., Rattanarojpong, T., Vanichviriyakit, R., & Senapin, S. (2017). Emergence of tilapia lake virus in Thailand and an alternative semi-nested RT-PCR for detection. *Aquaculture*, *476*, 111–118. https://doi.org/10.1016/j.aquaculture.2017.04.019

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Eyngor, M., Zamostiano, R., Kembou Tsofack, J. E., Berkowitz, A., Bercovier, H., Tinman, S., Lev, M., Hurvitz, A., Galeotti, M., Bacharach, E., & Eldar, A. (2014). Identification of a novel RNA virus lethal to tilapia. *Journal of Clinical Microbiology*, *52*(12), 4137–4146. https://doi.org/10.1128/JCM.00827-14

FAO. (2020). *The State of World Fisheries and Aquaculture 2020: Sustainability in action*. FAO. https://doi.org/10.4060/ca9229en

Faria, N. R., Sabino, E. C., Nunes, M. R. T., Alcantara, L. C. J., Loman, N. J., & Pybus, O. G. (2016). Mobile real-time surveillance of Zika virus in Brazil. *Genome Medicine*, *8*. https://doi.org/10.1186/s13073-016-0356-2

Gallagher, M. D., Matejusova, I., Nguyen, L., Ruane, N. M., Falk, K., & Macqueen, D. J. (2018). Nanopore sequencing for rapid diagnostics of salmonid RNA viruses. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-34464-x

Hoenen, T., Groseth, A., Rosenke, K., Fischer, R. J., Hoenen, A., Judson, S. D., Martellaro, C., Falzarano, D., Marzi, A., Squires, R. B., Wollenberg, K. R., de Wit, E., Prescott, J., Safronetz, D., van Doremalen, N., Bushmaker, T., Feldmann, F., McNally, K., Bolay, F. K., … Feldmann, H. (2016). Nanopore Sequencing as a Rapidly Deployable Ebola Outbreak Tool. *Emerging Infectious Diseases*, *22*(2), 331–334. https://doi.org/10.3201/eid2202.151796

International Committee on Taxonomy of Viruses. (2019) Taxonomy, Virus Taxonomy: 2018 Release. [Online] Retrieved from https://ictv.globa l/taxonomy (accessed on 13 March 2021)

Jansen, M. D., Dong, H. T., & Mohan, C. V. (2019). Tilapia lake virus: A threat to the global tilapia industry? *Reviews in Aquaculture*, *11*(3), 725–739. https://doi.org/10.1111/raq.12254

Kembou Tsofack, J. E., Zamostiano, R., Watted, S., Berkowitz, A., Rosenbluth, E., Mishra, N., Briese, T., Lipkin, W. I., Kabuusu, R. M., Ferguson, H., del Pozo, J., Eldar, A., & Bacharach, E. (2017). Detection of Tilapia Lake Virus in Clinical Samples by Culturing and Nested Reverse Transcription-PCR. *Journal of Clinical Microbiology*, *55*(3), 759–767. https://doi.org/10.1128/JCM.01808-16

Kerddee, P., Dong, H. T., Chokmangmeepisarn, P., Rodkhum, C., Srisapoome, P., Areechon, N., Del-Pozo, J., & Kayansamruaj, P. (2020). Simultaneous detection of scale drop disease virus and *Flavobacterium columnare* from diseased freshwater-reared

barramundi *Lates calcarifer*. *Diseases of Aquatic Organisms*, *140*, 119–128. https://doi.org/10.3354/dao03500

Laver, T., Harrison, J., O'Neill, P. A., Moore, K., Farbos, A., Paszkiewicz, K., & Studholme, D. J. (2015). Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomolecular Detection and Quantification*, *3*, 1–8. https://doi.org/10.1016/j.bdq.2015.02.001

Ninawe, A. S., Hameed, A. S. S., & Selvin, J. (2017). Advancements in diagnosis and control measures of viral pathogens in aquaculture: An Indian perspective. *Aquaculture International*, *25*(1), 251–264. https://doi.org/10.1007/s10499-016-0026-9

Nkili-Meyong, A. A., Bigarré, L., Labouba, I., Vallaeys, T., Avarre, J.-C., & Berthet, N. (2016). Contribution of Next-Generation Sequencing to Aquatic and Fish Virology. *Intervirology*, *59*(5–6), 285–300. https://doi.org/10.1159/000477808

Phusantisampan, T., Tattiyapong, P., Mutrakulcharoen, P., Sriariyanun, M., & Surachetpong, W. (2019). Rapid detection of tilapia lake virus using a one-step reverse transcription loop-mediated isothermal amplification assay. *Aquaculture*, *507*, 35–39. https://doi.org/10.1016/j.aquaculture.2019.04.015

Pulido, L. L. H., Mora, C. M., Hung, A. L., Dong, H. T., & Senapin, S. (2019). Tilapia lake virus (TiLV) from Peru is genetically close to the Israeli isolates. *Aquaculture*, *510*, 61–65. https://doi.org/10.1016/j.aquaculture.2019.04.058

Rodgers, C. J., Mohan, C. V., & Peeler, E. J. (2011). The spread of pathogens through trade in aquatic animals and their products. *Revue Scientifique et Technique (International Office of Epizootics)*, *30*(1), 241–256. https://doi.org/10.20506/rst.30.1.2034

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., & Belshaw, R. (2010). Viral Mutation Rates. *Journal of Virology*, *84*(19), 9733–9748. https://doi.org/10.1128/JVI.00694-10

Steinhauer, D. A., Domingo, E., & Holland, J. J. (1992). Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*, *122*(2), 281–288. https://doi.org/10.1016/0378-1119(92)90216-C

Subasinghe, R., Delamare-Debouteville, J., Mohan, C., & Phillips, M. (2019). Vulnerabilities in aquatic animal production. In *Revue Scientifique et Technique de l'OIE* 38(2), 423–436. https://doi.org/10.20506/rst.38.2.2996

Subramaniam, K., Ferguson, H. W., Kabuusu, R., & Waltzek, T. B. (2019). Genome Sequence of Tilapia Lake Virus Associated with Syncytial Hepatitis of Tilapia in an Ecuadorian Aquaculture Facility. *Microbiology Resource Announcements*, *8*(18). https://doi.org/10.1128/MRA.00084-19

Surachetpong, W., Janetanakit, T., Nonthabenjawan, N., Tattiyapong, P., Sirikanchana, K., & Amonsin, A. (2017). Outbreaks of Tilapia Lake Virus Infection, Thailand, 2015–2016. *Emerging Infectious Diseases*, *23*(6), 1031–1033. https://doi.org/10.3201/eid2306.161278

Taengphu, S., Sangsuriya, P., Phiwsaiya, K., Debnath, P. P., Delamare-Debouteville, J., Mohan, C. V., Dong, H. T., & Senapin, S. (2020). Genetic diversity of tilapia lake virus genome segment 1 from 2011 to 2019 and a newly validated semi-nested RT-PCR method. *Aquaculture*, *526*, 735423. https://doi.org/10.1016/j.aquaculture.2020.735423

Tattiyapong, P., Sirikanchana, K., & Surachetpong, W. (2018). Development and validation of a reverse transcription quantitative polymerase chain reaction for tilapia lake virus detection in clinical samples and experimentally challenged fish. *Journal of Fish Diseases*, *41*(2), 255–261. https://doi.org/10.1111/jfd.12708

Thawornwattana, Y., Dong, H. T., Phiwsaiya, K., Sangsuriya, P., Senapin, S., & Aiewsakun, P. (2020). Tilapia lake virus (TiLV): Genomic epidemiology and its early origin.

Transboundary and Emerging Diseases, tbed.13693. https://doi.org/10.1111/tbed.13693

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

Wang, M., Fu, A., Hu, B., Tong, Y., Liu, R., Liu, Z., Gu, J., Xiang, B., Liu, J., Jiang, W., Shen, G., Zhao, W., Men, D., Deng, Z., Yu, L., Wei, W., Li, Y., & Liu, T. (2020). Nanopore Targeted Sequencing for the Accurate and Comprehensive Detection of SARS-CoV-2 and Other Respiratory Viruses. *Small*, *16*(32), 2002169. https://doi.org/10.1002/smll.202002169

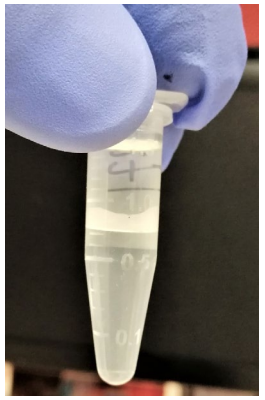
Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Yin, J., Wang, Q., Wang, Y., Li, Y., Zeng, W., Wu, J., Ren, Y., Tang, Y., Gao, C., Hu, H., & Bergmann, S. M. (2019). Development of a simple and rapid reverse transcription–loopmediated isothermal amplification (RT-LAMP) assay for sensitive detection of tilapia lake virus. *Journal of Fish Diseases*, *42*(6), 817–824. https://doi.org/10.1111/jfd.12983

1 Samples collection on farm

0.5 h

2 RNA extraction

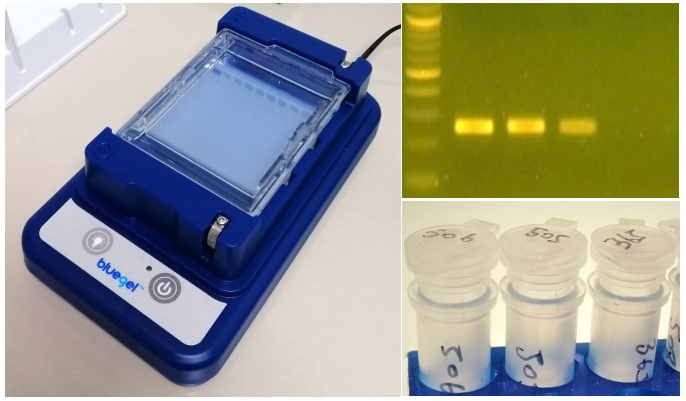


1 h

3 TiLV-specific PCR

2 h

4 Gel electrophoresis & DNA clean up



1 h

0.5 h

5 Standardize amplicons

1 h

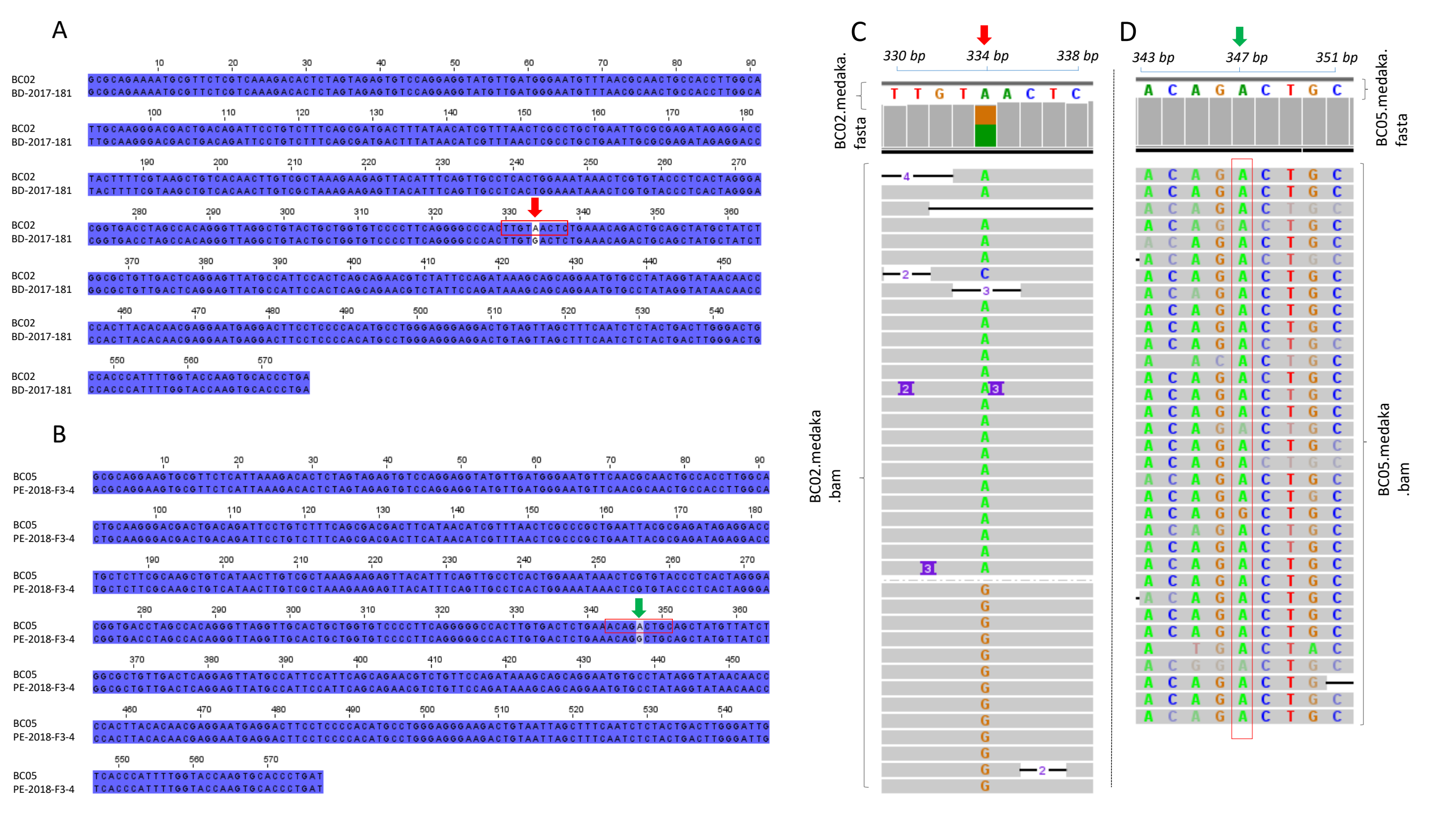
6 Nanopore multiplex Library preparation*

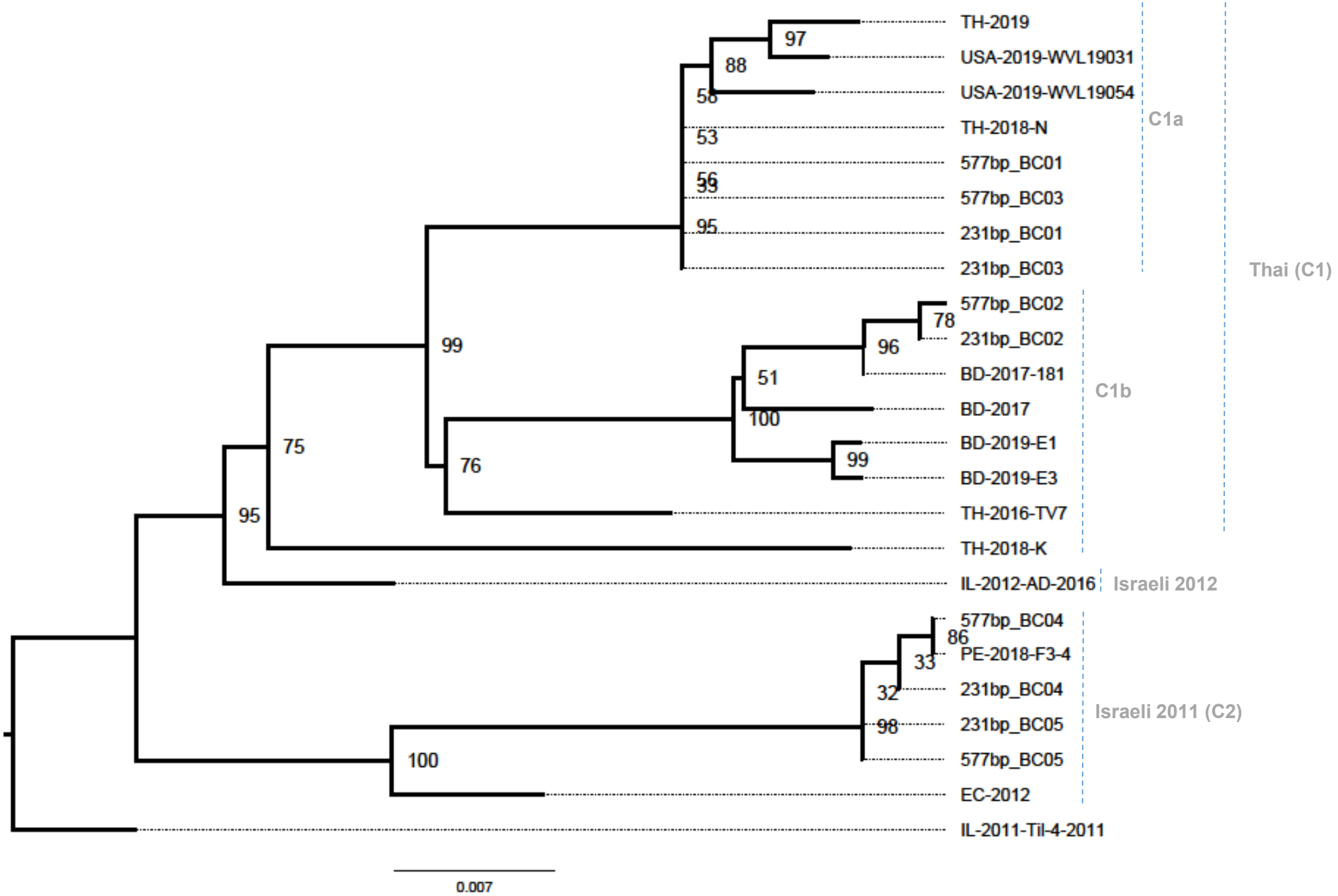
15 min - 5 h

7 MinION sequencing

0.5-1 h

8 Mapping to reference


A
BC02
BD-2017-181
B
BC05
PE-2018-F3-4
C
330 bp
334 bp
338 bp
BC02.medaka.
fasta
BC02.medaka
.bam
D
343 bp
347 bp
351 bp
BC05.medaka.
fasta
BC05.medaka
.bam


TH-2019
USA-2019-WVL19031
USA-2019-WVL19054
TH-2018-N
577bp_BC01
577bp_BC03
231bp_BC01
231bp_BC03
577bp_BC02
231bp_BC02
BD-2017-181
BD-2017
BD-2019-E1
BD-2019-E3
TH-2016-TV7
TH-2018-K
IL-2012-AD-2016
577bp_BC04
PE-2018-F3-4
231bp_BC04
231bp_BC05
577bp_BC05
EC-2012
IL-2011-Til-4-2011
97
88
56
53
56
55
95
78
96
51
100
99
99
76
75
95
86
33
32
98
100
C1a
C1b
Thai (C1)
Israeli 2012
Israeli 2011 (C2)
0.007