Developing Field-based Methods for Real-time Genomic Surveillance of Viral Outbreaks in Tilapia

Submitted by **Shayma Abdulmuttaleb Kareem Alathari**, to the University of Exeter as a thesis for the degree of ***Doctor of Philosophy in Biological Sciences*** in January 2024.

Signature: 

Abstract

Viral pathogens are having a significant impact on Tilapia health, threatening global food security. It is critical to increase biosecurity measures, diagnostic testing, and sequencing to better understand their evolution and spread in aquaculture. Yet, this is not feasible or adapted in many countries, as numerous fish farms are situated in resource limited regions with reduced access to diagnostic labs. Delays in turn-around time for results reduces the capacity for timely interventions. One way to avoid devastating outcomes of viral outbreaks in aquaculture is real-time whole genome sequencing (WGS) of isolates, using in-field tools to rapidly resolve chains of transmission and implement control measures.

Here, we investigate the feasibility of a tiled PCR method for genomic surveillance of Infectious spleen and kidney necrosis virus (ISKNV), a powerful tool successfully adapted for genomic surveillance of important human pathogens. We perform a complete phylogeographic analysis on isolates collected from Lake Volta, Ghana, since 2018. ISKNV, a slow evolving double stranded DNA virus has been causing mass mortalities in Ghana, and is capable of infecting a wide range of marine and freshwater fish. In addition, an in-field, non-destructive water sampling method was developed to monitor viruses in tilapia fish cages, by concentrating viral fraction on filters. Challenges were encountered when applying these methods in the field in Ghana, and the technical and economic issues are discussed. We evaluate our developed method on an equally important single stranded RNA virus, known as Tilapia Lake Virus (TiLV), which has been affecting the growth of tilapia for more than a decade. Finally, we highlight the challenges faced during a field -based genomic surveillance campaign performed on Lake Volta. This work is the first of its kind to develop and test a valuable in-field, routine monitoring and detection tool for viral outbreaks impacting the prosperity of aquaculture.

Table of Contents

List of Figures and Tables:

a. List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

an additional short read sequenced genome of a sample collected from the same farm in Bangladesh.

Chapter 5

Figure 1. *A MiniPCR machine connected to a mobile phone, performing a tiled PCR on ISKNV isolates collected from Lake Volta, Ghana.*
Figure 2. *In-field sequencing, in a hotel and near the fish cages on Lake Volta, Ghana.*
Figure 3. *Training a member of the commission of fisheries.*

b. List of Tables

Definitions and abbreviations

ddPCR: Droplet digital PCR

RT-PCR: Reverse transcription polymerase chain reaction

RT-qPCR: reverse transcriptase quantitative polymerase reaction

ISKNV: Infectious spleen and kidney necrosis virus

TiLV: Tilapia Lake virus

SNP: Single nucleotide polymorphism

iSNV: Intra-host single nucleotide variant

Hz: Hertz

µG: Microgram

gm: Gram

µL: Microlitre

mg: Milligram

MCP: Major capsid protein

GIFT: Genetically improved farmed tilapia

dsDNA: Double stranded DNA

(s/n/i): substitutions per nucleotide per infection.

CPE: Cytopathic effects

LMIC: Low- and Middle-Income Countries

TEM: Transmission electron microscopy

Cefas: Centre for Environment, Fisheries and Aquaculture Science

FAO: Food and Agriculture Organisation

OIE: Office International des Epizooties

µm: Micrometre

µM: Micrometre

mL: Millilitre

nM: Nanomolar

LAMP: Loop-mediated isothermal amplification

WGS: Whole genome sequencing

ssRNA: Single stranded RNA

ORF: Open reading frame

Covid Impact Statement

I had completed five months of my PhD when the Covid-19 pandemic and lockdowns hit. At that point I was learning how to work independently and have just started in the laboratory. During the lockdown period I was a mother of two young children (6 and 9) with online schooling and single parenting, as my husband could not get back to us. We did not see him for seven months. Due to lockdown restrictions implemented after phased re-opening of the laboratories, I had no access to ISKNV samples from Cefas or from Ghana. I was unable to start processing samples that were already at Cefas until late October/November 2020.

Following training for cell culture, I re-initiated the cell culture lab, where during lockdown I (in-person) collected cell lines from Cefas to propagate the virus understudy (ISKNV), but after the continued lockdown in Cefas, I had no access to the virus, and had to terminate the cell lines.

I tried to play my role as a molecular biologist and assisted in the ramping up for covid testing in Somerset (Southwest Pathology Laboratories) and spent some weekends in assisting with vaccinating the elderly, during the initial role out of the vaccines for Covid-19.

Further delay and impact occurred in early 2021 when further lockdowns and restrictions resulted in delays to proposed fieldwork to collect samples. The lockdowns and restrictions resulted in changes to my proposed fieldwork and research which were essential for Chapters 2 and 3, and we were in search for new viruses affecting tilapia aquaculture.

My research and analysis were further impacted in February 2021 when there was a fire in our laboratory, resulting in a 6-week lab closure and extensive entry restrictions upon re-opening that continued for over 6 months. The effects of the fire impacted our laboratory, and therefore sample processing and analysis, well into 2022.

Covid and the laboratory fire were both extreme circumstances that impacted my research data collection and analysis, but I was able overcome these challenges and they have made me a more resilient and robust scientist, searching for

methods to overcome the restrictions, which led to creating a lab from home, using portable devices to continue my training as a scientist.

Acknowledgements

I would like to express my immense gratitude to my supervisors, Professor Ben Temperton and Professor Charles Tyler for their continued support and guidance since the day I began the journey of my return to science. You have helped me reach my dreams and become part of a world filled with opportunities. I am extremely grateful for this opportunity and to be part of the research community at the University of Exeter.

Thanks to Cefas for choosing me for this project, especially Richard Paley and Andrew Joseph. this project wouldn't have been complete without you. I would also like to thank the whole team of the Commission of Fisheries in Ghana for allowing me to collect samples from their country, and explore their culture. This trip was the highlight of my PhD.

I would also like to thank the following people, without whom I would not have been able to complete this research, Aaron Jeffries, Karen Moore, David Studholme and Eduarda Santos. Your doors have always been open to my questions and concerns.

Thank you to my colleagues and long-life friends, Vicky Jackson, Imogen Worth, Lara Pinheiro, Katie Deakin, Georgie Savage, Liv Goode, and Lauren Adams. You have showered me with all sorts of kindness and support and made me feel cherished, giving me a sense of belonging, when I was thousands of miles away from home. You've made this journey a joyous one, and I wouldn't have wanted it any other way. Stephanie Andrews, you guided me through so many hopeless days and believed in my strength. Having you as my office buddy has made the biggest difference to my final year, and I am forever grateful. I thought I was alone in my journey until I met you all.

A huge thanks to Anke Lange, Luis Bolanos, Julie Fletcher, Lisa Bickly and Jo Warwick-Dugdale. You have answered my many questions, and guided me through the days that I was lost, with so much patience and kindness.

Thanks to Sariqa Wagley. You have provided me with encouragement and never failed to find ways to assist me in any way you could. Your kindness has no limit, and I am forever in your debt.

A huge thank you to my mentor, friend and rescuer, Dominique Chaput, who made me the scientist I am today.

For dad. Your love for science has been ingrained in me since the days you were a student and took me into the lab. I always aspired to be just like you (in a lab coat mixing things). Mum, I am forever grateful for your love and support and taking some of the parenting role throughout the Summer. You have both pushed me to be a better version of myself.

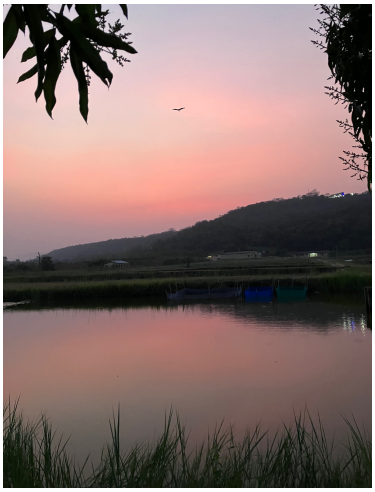
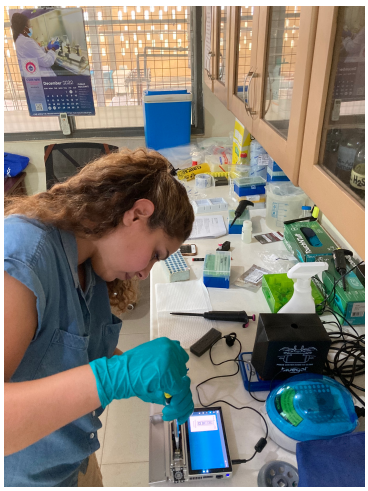
To my mother-in-law and father-in-law, Abdulrazzak Alsaidi. He was and always will be my biggest supporter. *'May his soul rest in peace'*

I would like to thank my brother Haider and his wife Sara for their unconditional support to me and my children throughout my studies. Words cannot express my gratitude. For my sister Lamees, that always guided me and warned me about how tough this was going to be. I cherish all of the phone calls on the way back from work, listening to my triumphs and not so great days. For my niece Jenine, and nephews Ameen, Hamza, and Tal, you have always been encouraging and motivating.

Finally, this work would not have been possible without the love and support of my husband Faik, my amazing and supportive kids Maryam and Ibrahim, you have motivated me every single day. Thank you for your understanding and patience. You all have struggled from being miles apart from each other for me to complete my dream.

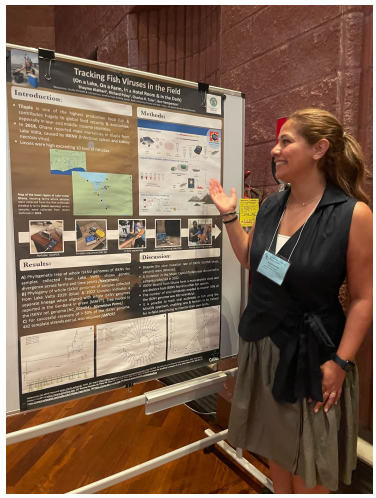
This PhD is dedicated to my lovely homeland,

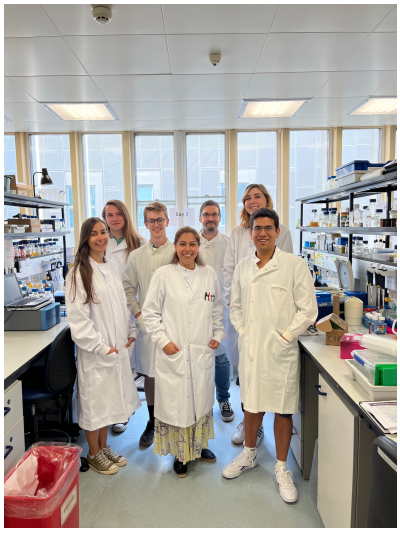
Iraq






Tracking Fish Viruses in the Field
Introduction:
Methods:
Results:
Discussion:




PINT OF SCIENCE
2022


On a Lake, On a Farm, In a Hotel Room & In the Dark!

1. General Introduction:

Today, 811 million people suffer from hunger and three billion cannot afford healthy diets. The United Nations has listed Zero Hunger as one of the global sustainable development goals to end extreme poverty by 2030 (Boykin et al. 2018). As global populations continue to grow, we aim to find ways to meet demand for food security, and aquaculture could potentially play an important role if its growth remains sustainable (FAO 2022). According to the Food and Agriculture Organisation (FAO), the total global fisheries and aquaculture production reached 178 million tonnes in 2020, with a total first sale value estimated at USD 406 billion, of which USD 265 billion was from aquaculture production alone. Aquaculture production remains the main driver of the growth of total fisheries production since the late 1980s, and continues to expand, albeit at a slower rate in the last two years (FAO 2022). New strategies have been suggested, such as Blue Transformation to enhance the role of aquatic feeding systems, by providing the legal, policy and technical frameworks required to sustain growth and innovation (FAO 2022). Blue Transformation builds on existing successes while providing a framework to overcome sustainability challenges, to maximise the contribution of aquatic (both marine and inland) food systems to food security (FAO 2022). Despite an increase in output, all forms of production are limited by infectious diseases, causing direct production loss and closure of aquaculture facilities. Recognizing the capacity of aquaculture for further growth demands new sustainable aquaculture development strategies. The priority should be to further develop aquaculture in Low- and Middle-Income Countries (LMICs), including Africa and in other regions where population growth will challenge food systems most (FAO 2020).

Tilapia are important for the sustainability of ecological systems and are the second most important group of farmed fish worldwide (FAO 2020). Eleven tilapia species are currently farmed in Africa, compared to only three species in 1980 (El-Sayed and Fitzsimmons 2023). In particular, Nile tilapia (*Oreochromis niloticus*) is a key fish species for freshwater aquaculture and is the most widely cultured tilapia species, with a global production estimated at 4,525,400 tons (El-Sayed and Fitzsimmons 2023; FAO 2020). This genus is hugely important for providing employment, as well as domestic and export earnings to large

populations worldwide (Machimbirike et al. 2019; Eyngor et al. 2014). Associated production has almost doubled over the past decade, due to their relative ease of farming, marketability and stable market prices (Wang and Lu 2016). Additionally, their fast growth rates, tolerance to extreme environmental conditions, high resistance to stress and diseases, trophic plasticity, high adaptation to tropical, subtropical and temperate environments, and their ability to reproduce in captivity has made them an ideal candidate for aquaculture all over the globe (El-Sayed and Fitzsimmons 2023).

Despite the fact that Nile tilapia (*Oreochromis niloticus*) are African freshwater fish, they have been introduced into many countries, where their journey started during the second half of the 20th Century, especially in Southeast Asia and the Americas, for aquaculture and fisheries enhancement. As a result, aquaculture of Nile tilapia became well established and has been steadily expanding in many countries (El-Sayed and Fitzsimmons 2023). In Africa, tilapia production is still dominated by Egypt (Yacout et al. 2016), but has become increasingly important in several other countries, boosting the local economy as an affordable source of animal protein for human consumption. Tilapia aquaculture provides an important source of nutrition, especially for populations that are otherwise dependent on a narrow range of staple foods. Tilapia have also been used to biologically control vectors of disease, such as malaria, Zika and bilharzia through predation on the hosts of the parasites (e.g., mosquitoes) (FAO 2018a).

In Ghana, Nile tilapia production has risen rapidly from only 954 tonnes in 2005 (FAO 2018b; Verner-Jeffreys et al. 2018; Ramírez-Paredes et al. 2021) to 52,470 tonnes in 2016 due to the government/World Bank-funded fisheries program through Ministry of Fisheries and Aquaculture Development (MOFAD) (Amenyogbe et al. 2018). Ghana, with its prosperous environment; full of rivers, seas, dams, dugouts, and suitable topography, climate, authority support, and high demand for fish, has all made aquaculture practicable and established countrywide (Amenyogbe et al. 2018). Tilapia is the preferable species for fish farming and consumption in Ghana, creating economic opportunities by employing thousands of people and improving the livelihoods and the general economy (Asiedu et al. 2017). Most production in Ghana is conducted under high density stocking in floating cage systems, as occurs in Lake Volta (Figure 1).

Most Ghanaian production of tilapia is centred around Lake Volta, (Figure 2), in cage culture units, with hatcheries predominantly located besides the River Volta, below the dam to the lake. This lake area provides an ideal location to support further expansion of the industry as it has a very large surface area, with an ideal temperature profile (28–30°C) for year-round production of tilapia (Verner-Jeffreys et al. 2018).

Figure 1. A tilapia farm on Lake Volta in Ghana, showing floating cage culture units.

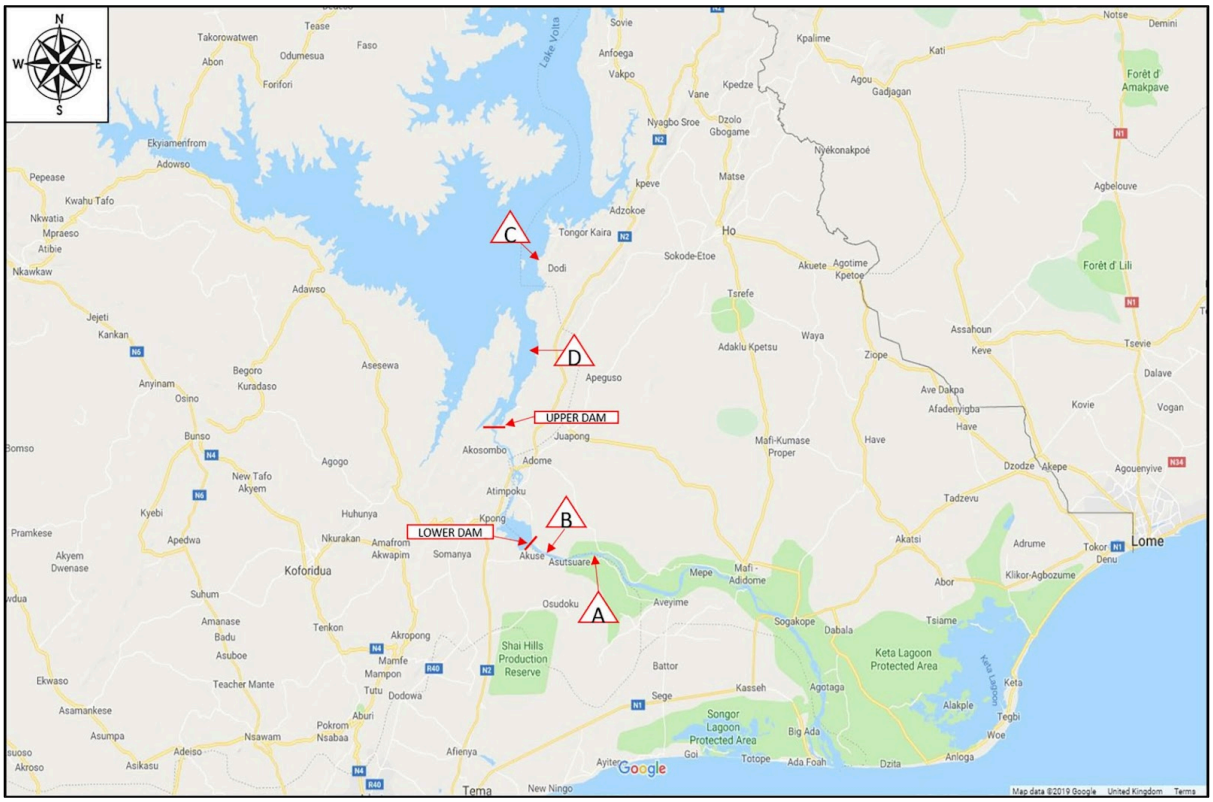


Figure 2. A map of the lower region of Lake Volta in Ghana, West Africa. Red triangles indicate the regions in chronological order (A to D) where the outbreaks of mortality occurred (Ramírez-Paredes et al. 2021).

Despite Ghana making rapid advances in aquaculture development, the sector faces several challenges related to management and production problems, such as limited knowledge of modern aquaculture techniques, inadequate supplies of improved seed, lack of continuing aquaculture policy direction, and inadequate funding for research (Amenyogbe et al. 2018). Furthermore, high density production of fish makes tilapia more susceptible to infectious diseases. Initially the major diseases threatening intensively cultured tilapia farms have been identified as predominantly bacterial infections, such as streptococcal infections (Dong et al. 2015). However, there is an increasing number of emerging viral infections that affect farmed tilapia worldwide. Most recently, global attention has been focused on the emergence of a new virus, tilapia lake virus (TiLV), due to the widespread outbreaks of the disease in three continents, signalling a serious threat to tilapia aquaculture globally (Dong, Siriroob, et al. 2017; Machimbirike et al. 2019; Jansen, Dong, and Mohan 2018). Meanwhile, several other viral infections with equal importance have been neglected by the scientific community. These viruses require systematic investigation, as they have been

reportedly associated with relatively high mortalities (20–100%) in several occurrences of natural disease outbreaks or laboratory challenges. Examples of some of these viruses described in tilapia include Bohle iridovirus (BIV, Ranavirus), betanodavirus, tilapia larvae encephalitis virus (TELV), infectious pancreatic necrosis virus (IPNV), nervous necrosis virus (VNN, Betanodavirus), and infectious spleen and kidney necrosis virus (ISKNV) (Machimbirike et al. 2019; Shlapobersky et al. 2010).

In late 2018, unusual patterns of very high mortality in the Asutsuare region (see Figure 2) were reported in intensive tilapia cage culture systems across Lake Volta in Ghana (https://goo.gl/LmqbG2). Polymerase chain reaction (PCR) confirmation, quantitative polymerase chain reaction (qPCR) and DNA sequencing showed that all samples were strongly positive for the presence of the ISKNV virus. Samples collected from the same farms had all tested negative for the virus the previous year. A week following the first report, a second farm located in the Akuse region suffered the same problem. By the end of 2018, despite the attempts to reduce losses by increasing the production of fingerlings, or treatment with antibiotics, most tilapia farmers in Lake Volta were not able to contain these mortalities. By mid-October to late November, the Dodi region and the Asikuma region reported massive mortalities (Ramírez-Paredes et al. 2021). Tilapia production on lake Volta, accounting for over 90% of Ghana's aquaculture output, has severely suffered due to the ISKNV disease outbreaks. This has led to the closure of more than 50 farms, loss of over 400 jobs, and the disruption of the livelihoods of the communities along the lake (Okai 2021).

1.1 Infectious Spleen and Kidney Nicrosis Virus

ISKNV is a species of the genus *Megalocytivirus*, and is one of five genera within the Iridoviridae family of large, enveloped, double stranded DNA viruses (Mohr et al. 2015), that appeared in ornamental fish in the late 1980s until the early 1990s (Go et al. 2016). These five genera include: Iridovirus, Chloriridovirus, Ranavirus, Lymphocystivirus, and Megalocytivirus (Mahardika et al. 2009). Recently, *Megalocytiviruses* were divided into three genotypes: infectious spleen kidney necrosis virus (ISKNV), red seabream virus (RISV) and turbo reddish body iridovirus (TRBIV). This division was based on phylogeny of the conserved major capsid protein and the ATPase genes (Fu et al. 2011; Shiu et al. 2018). In general, RSIV, ISKNV, and TRBIV represent three genotypes, with similar viral properties to ISKNV species. However, phenotypic diversity still exists among the different isolates, even within the same genotype or subgenotype, with concomitant differences in virus replication, virulence, and host range (Fu et al. 2023). *Megalocytivirus* have been identified as pathogens that cause fatal systemic infections, leading to mass mortalities of numerous fish species, attracting attention with significant impact on ecological and economical wild and cultured fish (Xu et al. 2008; Sukenda et al. 2020). ISKNV infects invertebrates and poikilothermic vertebrates, including insects, fish, amphibians, and reptiles (Shi et al. 2010).

The characterizations of the ISKNV genome by molecular cloning and physical mapping have been reported, and the complete genome sequence has been determined from mandarin fish (*Siniperca chuatsi)* samples collected in 1998. The genome was 111,362bp in length and contained 124 putative ORFs (Figure 3) (He et al. 2001).

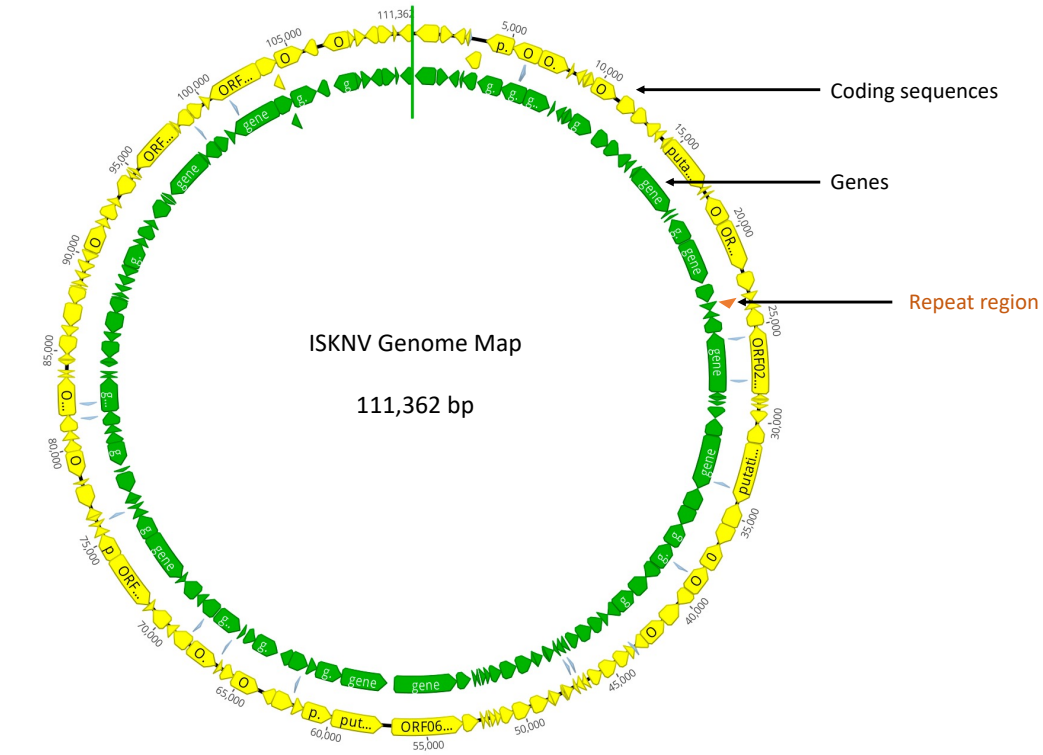


Figure 3. An ISKNV genome map created in Geneious Prime (v. 2022.1.1), showing the repeat region; the inner ring showing the genes (124 ORFs); and the outer ring showing the coding sequences (CDS).

ISKNV virions are icosahedral, around 150 nm in diameter, and show an electron lucent core under transmission electron microscopy (TEM) (Figure 4) (Ramírez-Paredes et al. 2021). The core contains a single linear dsDNA molecule, whose structure is highly methylated at cytosines in the CpG and circularly permuted upon infection (Fu et al. 2011).

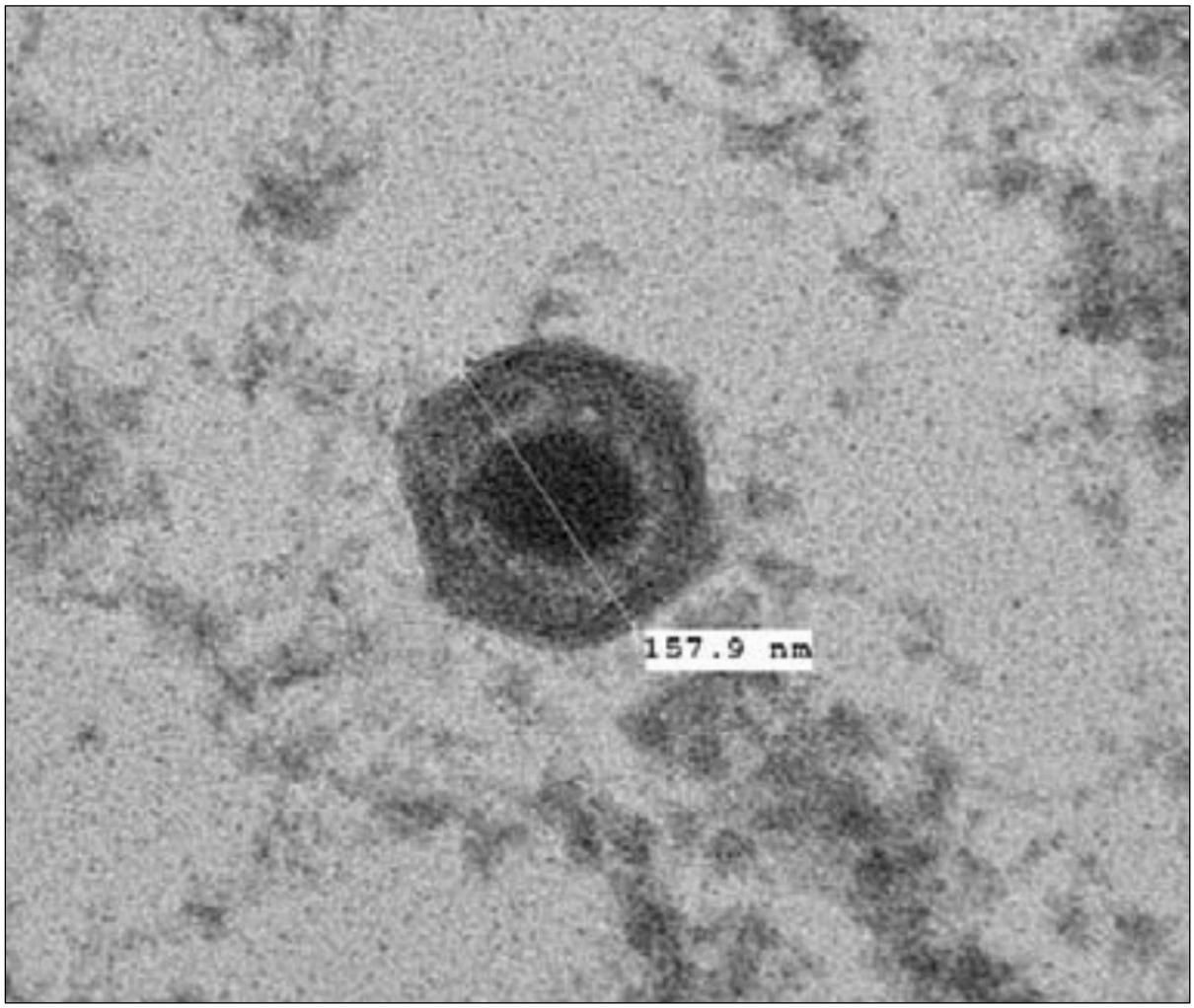


Figure 4. Mature icosahedral ISKNV virion from Lake Volta; showing the outer capsid and inner membrane with central electron lucent core (Ramírez-Paredes et al. 2021).

The structure of the virions possess a lipid membrane that lies between the viral DNA core and the capsid, and plays a role in infection of the host through the involvement in virus entry to the cell (Williams et al. 2005). The major capsid protein (MCP) is also considered an important structural component that mediates virus entry into the host cell, as it is involved in the process of ISKNV virus infection by interacting with caveolin-1-protein (Cav 1) of the host cell to induce the caveolin endocytosis, shown in Figure 5 (Islam et al. 2023; Throngnumchai et al. 2021). Although enveloped virions of ISKNV were observed in Nile tilapia (*Oreochromis niloticus*) infected with ISKNV in Ghana (Ramírez-Paredes et al. 2021), Iridoviruses can be either enveloped or nonenveloped, depending upon whether they are released from the cell by lysis, or bud from the plasma membrane. The envelope is not essential for cell entry and naked virions can also be infectious (Williams et al. 2005). While both enveloped and naked

virions are infectious, the infectivity of the enveloped virions is higher, suggesting that one or more viral envelope proteins play an important role in virion entry (Williams et al. 2005).

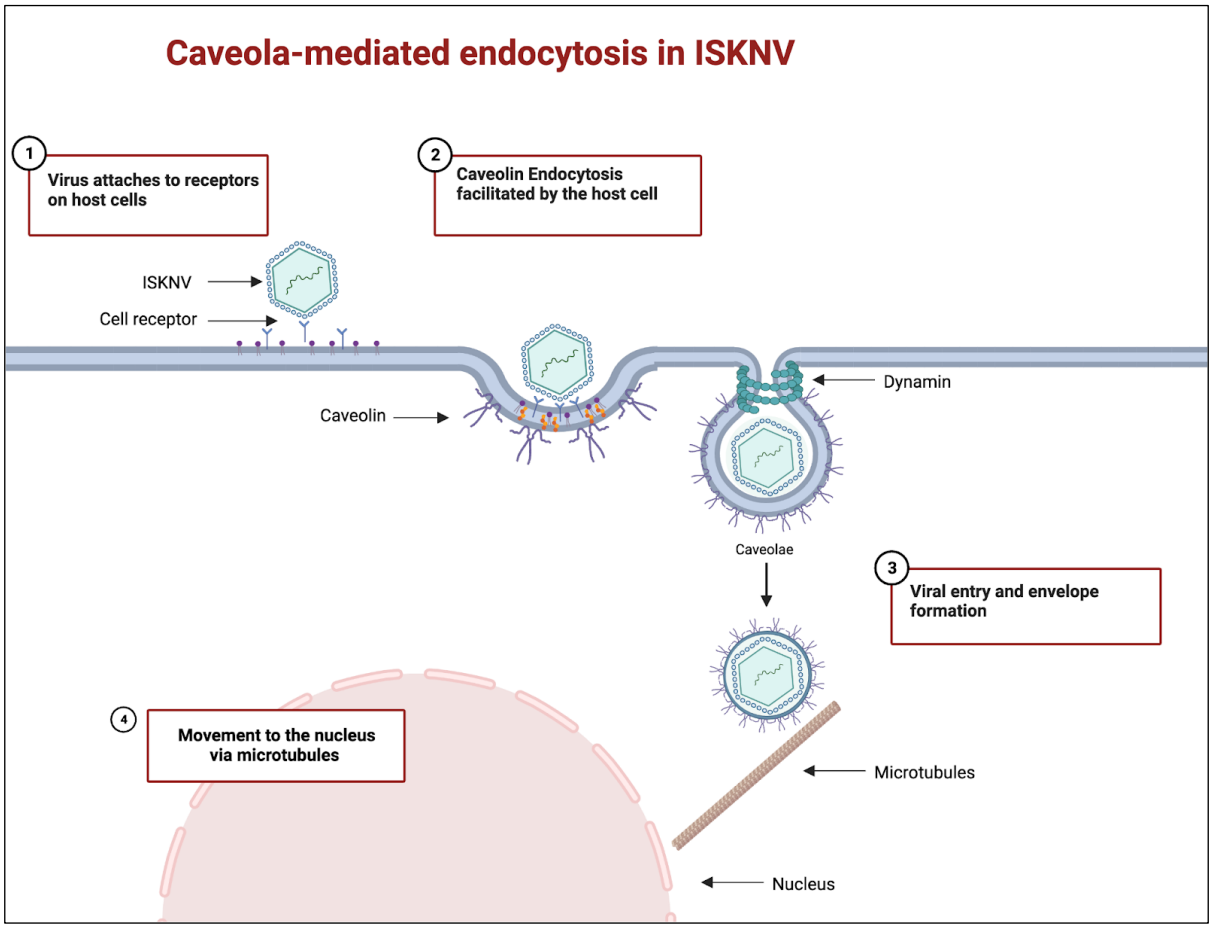


Figure 5. An illustration of the cellular entry of ISKNV; showing viral entry via Caveola- mediated endocytosis, facilitated by host receptors on the cell membrane.

The MCP (ORF006) is one of the major ISKNV immunogenic proteins, along with ORF054 (transmembrane protein), ORF055 (transmembrane protein), ORF101 (transmembrane protein), ORF117 (transmembrane protein), and ORF125 (ankyrin repeat protein). Moreover, ORF086 (helicase protein) is presumed to be a helicase vital to virus replication (Dong et al. 2013; Fu et al. 2015), while Zeng et al. (2021) stated that ORF022 is a virulent gene (Zeng et al. 2021). Their proof was based on a study which involved testing a live attenuated gene-deleted vaccine candidate, ΔORF022L, in mandarin fish. They reported 100% survival of the ΔORF022L-infected fish challenged with ISKNV infection, inducing the response of an anti-ISKNV-specific antibody.

1.2 Clinical Symptoms of ISKNV:

The typical external and internal signs of *megalocytivirus* disease are lethargy, anorexia, darkening of skin tissue, distended body cavity (coelomic distension due to ascites) (Figure 6). Other reported signs include: erratic swimming, increased ventilation, ulceration, haemorrhages (including pinpoint haemorrhages on the skin and gills), pale gills/anaemia, fin erosion, white faeces, and heavy mortalities (Jung-Schroers et al. 2016; Subramaniam et al. 2016; Yanong and Waltzek 2010). In the outbreak of ISKNV affecting tilapia in Ghana between the year 2018 and 2019, fish were observed swimming away from their school with erratic behaviour, and displayed a range of clinical signs, including skin nodules, frayed fins, loss of eyes, opaque eyes, loss of scales, exophthalmia, anorexia, discoloration or darkened skin, excess of mucous, skin haemorrhages and distended abdomen (Ramírez-Paredes et al. 2021). During the 2023 outbreak, samples were observed with similar symptoms as the above, along with white lips and tail erosion, seen in Figure 6 (a).

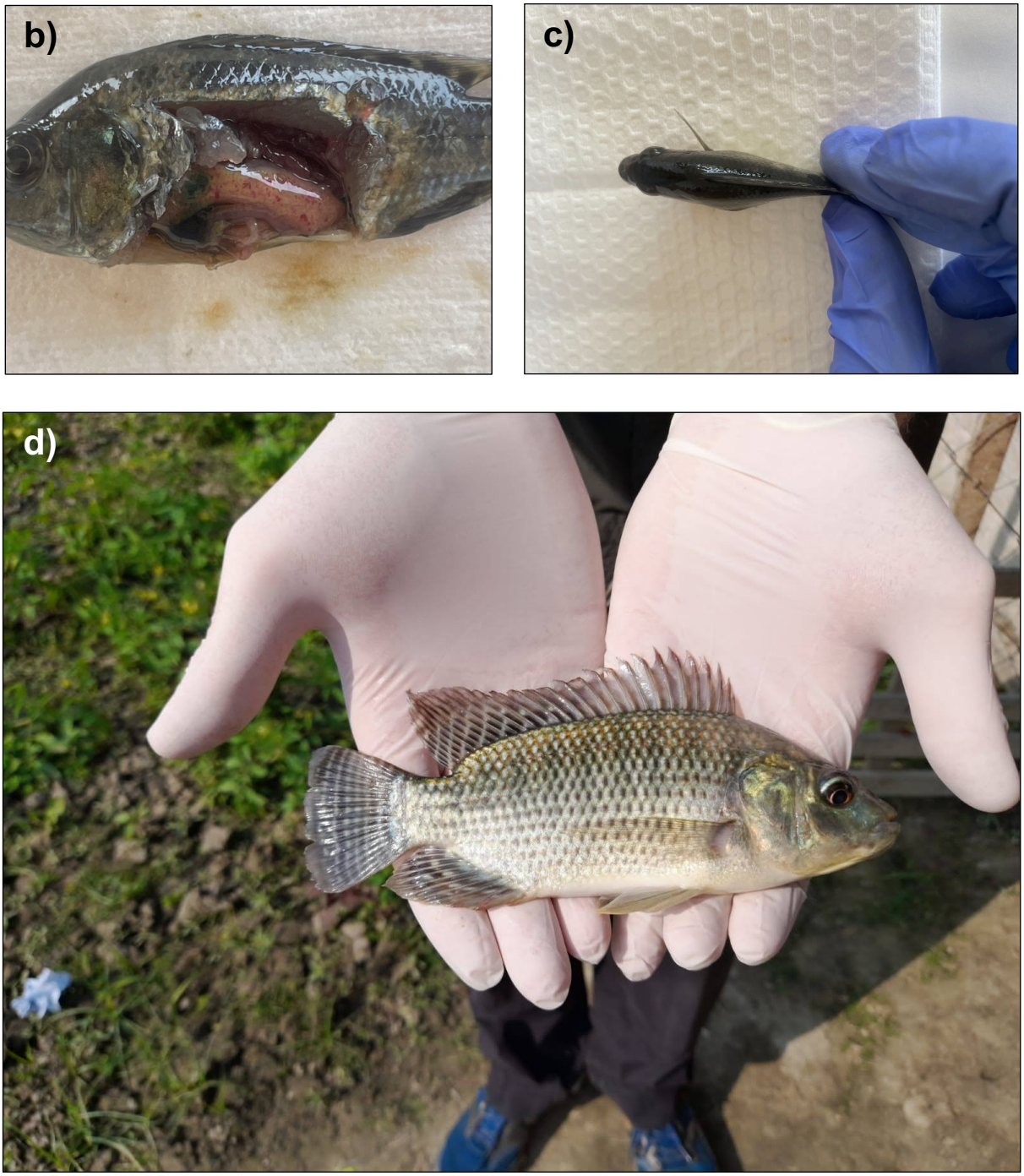


Figure 6. Moribund Tilapia fish (*Oreochromis niloticus)* collected from Lake Volta, Ghana (2023); a) distended body cavity and tail erosion; **b)** friable, pale, and haemorrhages liver; **c)** skin darkening and bulging eyes **d)** a healthy Nile tilapia collected from Khulna, Bangladesh, for comparison. *Fig 5.d Photo credit: to Jamie McMurtrie*.

Histopathological examination of fish infected with ISKNV reveals a severe systemic abundance of intravascular megalocytes that are especially prominent in the gills, kidney, spleen, liver, and intestinal submucosa. Enlarged cells could be observed by light microscopy due to the formation of inclusion body-bearing cells (IBCs) (hypertrophied cells containing large foamy or granular basophilic inclusions) as well as necrotized cells, allowing virus propagation within the intracytoplasmic virus assembly site (VAS). During necropsy, fish are presented with enlarged and haemorrhagic organs including the spleen, heart, brain, gills, but most notably liver and kidney. Other organs and tissues, including muscles, gonads, heart, gills, and the gastrointestinal tract, may also be affected. Some fish may have amber coloured haemorrhagic fluid visible within the body cavity (Mahardika, Muzaki, and Suwirya 2009; Yanong and Waltzek 2010; Subramaniam et al. 2016; Ramírez-Paredes et al. 2021). Additionally, use of cell lines can provide a method for demonstrating the presence of infectious ISKNV in a sample by examining the cytopathic effects (CPE) of ISKNV infected cell lines when examined under a light microscope. This can be done by using different cell-lines such as Grunt fin (GF), Bluegill fry (BF-2) and Snake head (E-11) fish cell lines (Ramírez-Paredes et al. 2021; OIE 2023). Diagnosis post inoculation, is achieved by observing altered cellular morphology, including rounding cells, as well as cell dropout and monolayer disruption (Figure 7).

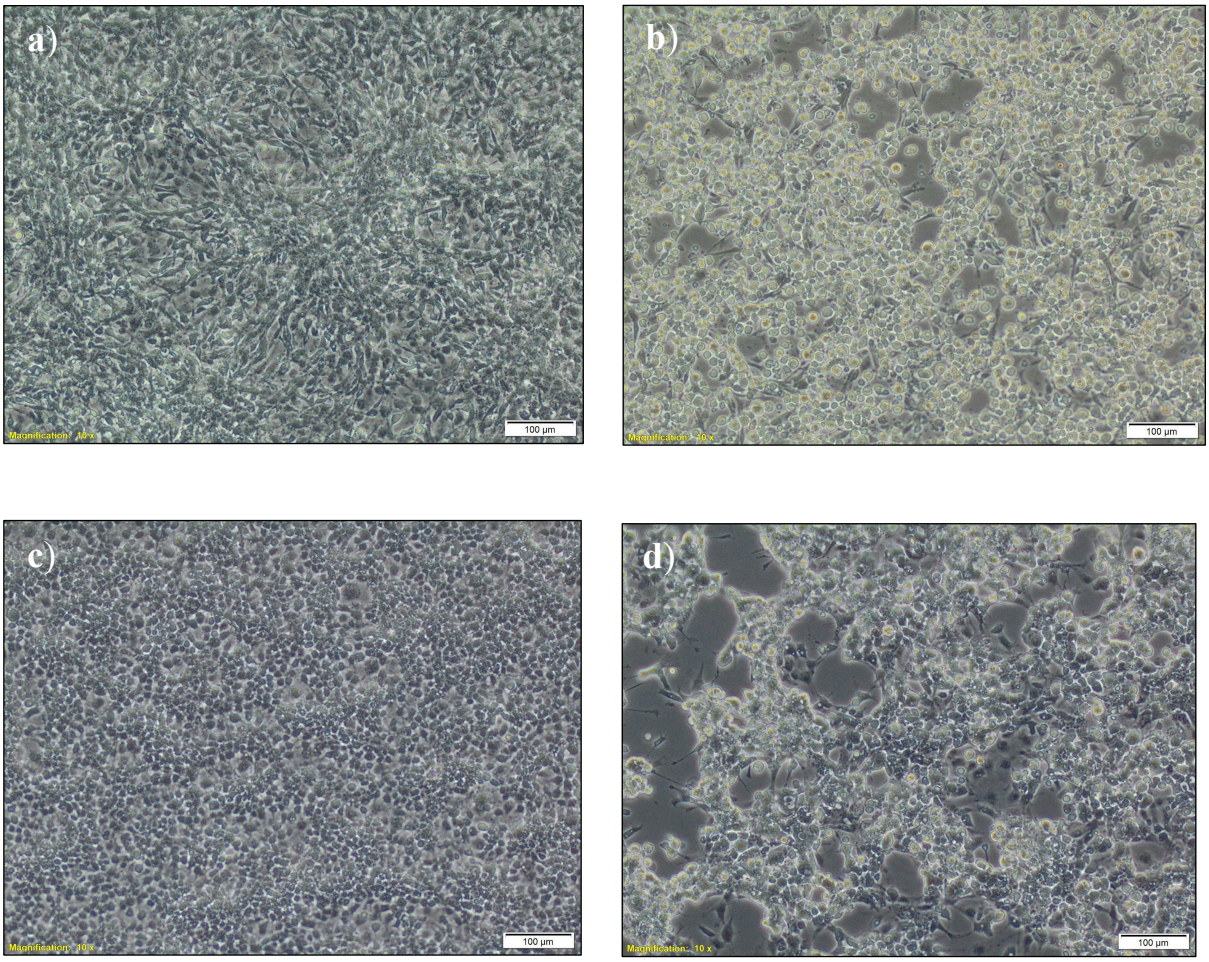


Figure 7. Microscopic examination of the Cytopathic effect (CPE) of ISKNV in cell lines. a) Blue-gill fry (BF-2) cell line control post 48 hr inoculation; **b)** infected BF-2 with ISKNV strain (PM382596), showing enlarged and refractile cells; **c)** Control flask of E-11 cell lines at 48 hr; **d)** infected BF-2 with ISKNV (strain PM382596) showing enlarged and refractile cells. Obvious detachment of the cell line monolayer. Scale bar = 100 µm.

1.3 Transmission of ISKNV:

In many forms of aquaculture, farmed fish share the same water column with wild aquatic animals and thus will experience the same viral challenges. Viruses carried by wild aquatic animals are often not sufficiently abundant to sustain the natural transmission cycle, but can be facilitated by the high density of hosts in aquaculture which, with associated chronic stress, provides opportunities for the emergence of viral diseases (Kibenge 2019). Spread of viral diseases in aquaculture is due to many factors such as fish migration, and the return of the

migratory species to the spawning sites, carrying both host and viral pathogens long distances. Human activities have also played a key role in the spread of viral diseases in farmed fish, such as industrial activities and mining, through the introduction of new fish species in the area. An increased viral transmission has also been related to an increase in the number of ships entering a farm, as well as a shorter distance between farms (Mugimba et al. 2021). Additionally, the expanding global trade in live aquatic animals and their products has been accompanied by long distance geographical redistribution of aquatic animal species and their viruses, causing a continuous emergence of viral diseases in aquaculture (Kibenge 2019). ISKNV can infect both freshwater and marine fishes, in both cultured and wild stocks, and is potentially spread through fish trading from Asia and other Southeast Asian countries (Sukenda et al. 2020; Yanong and Waltzek 2010). Affected stages are usually adult female (ovary), fertilised eggs, fry and fingerlings (Machimbirike et al. 2019). Susceptibility in juvenile fish is generally higher than that in adults. ISKNV is capable of inflicting mass mortalities especially at early stages (fry and juvenile stages). Since infections with this virus have the potential to be transmitted by both horizontal (transmission among the same generation) and vertical (transmission from parent to offspring) sources, there is an obvious risk to commercial aquaculture (Suebsing et al. 2016; Figueiredo et al. 2020).

1.4 Emergence and geographic distribution of ISKNV:

Megalocytiviruses are causative agents of severe disease accompanied by high mortality in multiple species of marine and freshwater fish (Kurita and Nakajima 2012). ISKNV has been reported in most continents such as Asia (Fusianto et al. 2021), Africa (Ramírez-Paredes et al. 2021), North America (Shahin et al. 2021), South America (Figueiredo et al. 2022), and Australia (Go and Whittington 2006) (Table 1).

ISKNV, which is diagnosed by a characteristic histopathology and electron microscopically studies, was first identified to infect Chinese mandarin fish 1994 (Table 1), and has resulted in significant economic losses in many fishponds in China. In 2001, mandarin fish was the only species affected by ISKNV in natural outbreaks. In 2002, infection trials on 20 other teleosts cultured in China showed

that large-mouth bass, *Micropterus salmoides*, were also highly susceptible to ISKNV, yet the other 18 fish under examination were resilient, including tilapia. (He et al. 2002; Jung-Schroers et al. 2016). Subsequently, outbreaks of ISKNV were reported in many parts of the world with an increasing host range (Table 1). Mortality events due to ISKNV were reported in North America, South America, Europe, Australia, Africa, Southeast Asia. The first observation of ISKNV in tilapia species was on a tilapia farm in the United States in 2012 that reported a 50-75% mortality rate over a two-month period (Howell 2019).

Table 1. Literature reports of different species of fish infected with ISKNV.

SPECIES AFFECTED (HOST)	***Year***	***MAIN FINDINGS***	***SITE***	***IMPORT***	***SOURCE***
Mandarin fish (Siniperca chuatsi)	1994-2001	Restricted to Asian countries	China	NA	(He et al. 2002)
	2012	Transmission: contaminated water/feeding	Japan	China	(Subramaniam et al. 2016)
Zebrafish (Danio rerio)	2012		Malaysia		(Subramaniam et al. 2014)
Ornamental fish (Trichogaster leeri)	2008	Induced mortality up to 70%	Korea	Singapore China	(Jeong et al. 2008)
Nile tilapia (Oreochromas niloticus)	2012		USA	NA	(Howell 2019)
	2016	Recurring outbreak in same facility	USA	Latin America and Asia	(Subramaniam et al. 2016)
	2018-2019	levels of morbidity & mortality (60-90%)	Ghana	Africa, Asia and South America	(Ramírez-Paredes et al. 2021).
	2022	levels of morbidity & mortality (60-90%)	Ghana	NA	(Alathari et al. 2023)

	2020		Brazil	NA	(Figueiredo et al. 2020)
Angel fish (Pterophyllum scalare)	2014	Transmission: through water	Germany	Colombia	(Jung-Schroers et al. 2016)
molly (Poecilia sphenops) & Angel fish (Pterophyllum scalare)	2018-2019		India		(Pattanayak, Paul, and Sahoo 2020)
Ornamental fish (Platys xiphophorus)	2012	97 of 111 imported fish were infected	Australia	Singapore, Malaysia, Sri Lanka	(Mohr et al. 2015)
	2013-2014				
Dwarf gourami Trochogaster lalius	2000		Australia		(Go and Whittington 2006)
	2004		Korea		(Jeong et al. 2008)
	2000		Singapore		(Kurita and Nakajima 2012)
Asian sea bass barramundi (Lates calcarifer)	2000		Malaysia		(Kurita and Nakajima 2012)
	2001-2009		Taiwan		(Huang 2011)
	2012-2014	1st documented infection is Asian sea bass	Vietnam	NA	(Dong et al. 2017)
	2017		Indonesia		(Thanasaksiri et al., 2021)

Orange-spotted grouper (Epinephelus coioides)	2000	Philippines	(Kurita and Nakajima 2012)
	2001-2009	Malaysia	(Razak, Ransangan, and Sade 2014)
	2004	China	(Kurita and Nakajima 2012)
Three-Spot gourami Trichogaster trichopterus	2015	Brazil	(De Lucca Maganha et al, 2018)
Paradisefish Macropodus opercularis	2015	Brazil	(De Lucca Maganha et al, 2018)
Giant gourami Osphronemus goramy)	2020	Indonesia	(Sukenda et al. 2020)
Siamese fighting fish Gourami	2016-2018	Thailand	(Baoprasertkul and Kaenchan, 2019)
Pearl gourami Trichogaster leeri	2004	Korea	(Jeong et al. 2008)
	2012	Malaysia	(Subramaniam et al., 2014)
Murray cod Maullochella peelii	2003		(Lancaster et al, 2003)
Flatted grey mullet Mugil cephalus	1999-2000	Singapore	(Gibson-Kueh et al 2004)

1.5 Containing ISKNV:

Few controls or mitigation measures currently exist for viral diseases in aquaculture, which continue to negatively impact aquaculture significantly (Kibenge 2019). Identification of persistent infection with low prevalence is challenging because viral load may be below limits of detection, especially in asymptotic carriers from natural populations. Thus, it is important to develop a sensitive tool that can help early diagnosis (Suebsing et al. 2016). Too often, a prolonged period from the first observation of mortality in fish to the identification and reporting of the causative agent occurs, delaying the application of appropriate control and risk management measures. Unless a paradigm shift occurs in dealing with aquaculture biosecurity risks, this sector will remain vulnerable to new and emerging diseases (FAO 2020). The use of pathogen-free tilapia for breeding programmes, could increase producers' chances in avoiding viral infections. Developing therapeutics and implementing a vaccination programme would also be beneficial as part of a long-term development strategy. Unfortunately, vaccination is usually ineffective for these viruses, as they infect fish at early stages (i.e. larvae and fingerlings), when the fish do not have a fully developed immune system (Mondal and Thomas 2022). A commercial vaccine designed for RSIV did not offer protection against other genotypes, such as ISKNV (Dong et al. 2017). Recently, a new study has presented an inactivated ISKNV-I vaccine, which is a formalin-killed cell (FKC) vaccine generated from an ISKNV-I isolate, that could confer almost complete protection against RSIV-I and RSIV-II as well as ISKNV-I, belonging to the genus megalocytivurus (Fu et al. 2023). However, such vaccines are costly, and challenging to administer to individual fish, especially in remote farms.

Some simple but effective preventative methods were reported, such as the use of disinfectants in a study done by Fusianto and et al (Fusianto et al. 2019), while another method was listed to assess the potential of seaweed compounds to block viral entrance by inhibiting the MCP (Islam et al. 2023). Incorporating disease resistance and feed efficiency traits into fish breeding has been suggested, and could create genetically improved fish strains. In Southeast Asia the introduction of genetically improved farmed tilapia (GIFT) has led to improved

productivity, ranging from 18% to 58% in China and Bangladesh with each generation yielding 7–10% gains in productivity (Ragasa et al. 2022).

1.6 ***Diagnosis & Monitoring ISKNV***

1.6.1 ***Polymerase Chain Reaction:***

PCR has been used as a confirmatory diagnostic method for detecting ISKNV, making it useful for surveillance in support of aquaculture biosecurity. Targeted organs for ISKNV PCR assays are spleen, kidney, or liver, with subsequent purification of nucleic acid. The first PCR assay developed for ISKNV in fish samples was described by Kurita and et al, and is recommended by the OIE (Office International des Epizooties) as the 'gold standard' reference PCR assay for its detection (Kurita et al.1998; OIE 2019). Other molecular techniques have been described for the detection of ISKNV such as, nested PCR, qPCR, and LAMP assay (Xu et al. 2008; Suebsing et al. 2016; Pattanayak, Paul, and Sahoo 2020). Although conventional PCR is a useful tool for *megalocytivirus* detection, because of its high specificity and sensitivity, its results are semi-quantitative and therefore unable to quantify the viral load precisely. qPCR is considered to have more advantages compared to conventional PCR, including quantitative measurement, minimal standardisation using a standard curve and easy data analysis (Lin et al. 2017). qPCRs are now frequently used to quantify viral pathogens in aquaculture, such as fish and shrimp, with the TaqMan assay being more specific than the SYBR Green assay. Lin et al have shown that using TaqMan quantitative real-time PCR for detecting ISKNV Genomic DNA was 10,000 times higher than that of conventional PCR (Lin et al. 2017). In general, all the methods mentioned above usually lack the specificity for ISKNV, and may fail to reliably differentiate ISKNV from other subgroups of *Megalocytiviruses*, i.e. RSIV and TRBIV (Pattanayak, Paul, and Sahoo 2020).

Furthermore, droplet digital PCR (ddPCR) has been established as a novel, sensitive, accurate and absolute quantitation method that does not require a standard curve (a calibration curve which is specifically constructed for each pathogen). Standard curves that mimic the samples well enough for accurate results are difficult to create, and batch-to-batch differences must be accounted

for. Additionally, small differences in efficiency in the lower ranges of the standard curve may further bias quantitative results (Kiselinova et al. 2014; Maar and Prantner 2020). In a recent study (Lin et al. 2020), a sensitive ddPCR protocol was developed to rapidly detect and quantify ISKNV DNA. This method proved to be highly specific to ISKNV and does not cross-react with other iridoviruses. This study also showed that the sensitivity of the ddPCR assay was 20-fold higher than that of the qPCR assay, and the positive detection rate of ddPCR (65.22%) was higher than that of qPCR (30.43%), showing superiority for detection in samples with low ISKNV viral loads, enabling the surveillance of sources and transmission routes of ISKNV (Lin et al. 2020).

Despite this method being reliable for the detection of ISKNV, important differences between genotypes and further subgroups within the ISKNV species could be overlooked. Conventional clinical tests such as PCR and serology, are being continually optimised, but are restricted to specific gene targets. Simple visualisation of PCR products can provide a measure of presence/absence of a particular region of interest and variation outside of these regions remains undetected (Kiselev et al. 2020). Therefore, viral diagnosis based on gene sequencing, using single nucleotide polymorphisms (SNPs) could differentiate isolates from different host species, country of origin, and time of collection (Fusianto et al. 2023).

1.6.2 ***Genomic Surveillance:***

To mitigate the effects and spread of viral diseases in aquaculture, it is critical to achieve rapid detection of the causative agent, understand their epidemiology; and to disseminate the information efficiently to raise awareness (Assefa and Abunna 2018). Until recently, comprehensive surveillance systems relied on case counting and simple genotyping techniques (Ghosh et al. 2012), but surveillance has been markedly improved through recent advances in genomics. In human health, genome sequencing has revolutionised our ability to track infectious disease outbreaks, from initial detection to understanding factors that contribute to its geographical spread. It has emerged as a critical tool in real-time response to these outbreaks, by providing insights into how viruses transmit, spread and evolve (Gardy, Loman, and Rambaut 2015; Quick et al. 2017).

Virus genomics have been used to investigate infectious disease outbreaks for several decades. Each time a virus replicates, errors in the genome sequence can occur that generate genetic variation. At the population level, the rate at which this occurs is related to the number of transmission events and error-correcting capability of the virus (with ssRNA viruses mutating two orders of magnitude faster than dsDNA viruses) (Peck and Lauring 2018). Variants (viruses that differ in their genomic sequence) can emerge with different phenotypic characteristics (Lauring and Hodcroft 2021). Genomic surveillance can assist in detecting these new variants, and allow the understanding of epidemiological and emergence dynamics from virus genomes sampled and sequenced over short epidemic timescales. The science of using genomics and associated epidemiological analyses is known as 'Genomic epidemiology'. Early virus sequencing from an outbreak could uncover the identity and geographic location of the reservoir host, and determine its evolutionary rate to help predict its future course (Gardy, Loman, and Rambaut 2015; Grubaugh, Ladner, et al. 2019; CDC 2021).

As shown by the recent emergence of variants of concern (VOC) for SARS-CoV-2; (a recent term used to describe variants that have potentially enhanced transmission, pathogenicity, immune escape, or a combination of all three) viruses can evolve, rapidly gaining fitness advantages. An example is the new lineages of SARS-CoV-2 that have been associated with elevated rates of viral spread, and decreased sensitivity to natural and/or vaccine acquired immunity (Lythgoe et al. 2021; Altmann, Boyton, and Beale 2021; Naveca 2021). Genetic diversity increases as an outbreak progresses, due to the accumulation of genetic changes in the viral genomes at each round of replication (Grubaugh et al. 2019), increasing the probability of new variants emerging. The greater the prevalence of the disease, the greater the likelihood of VOCs emerging due to the increased number of replication events. This gives rise to an urgent need for rapid and reliable diagnostics.

Targeted assays only cover specific genes and new VOCs could be missed. In contrast, whole genome sequencing captures all variation across the full genome, providing greater resolution of emerging diversity. With the information obtained

from genomic surveillance (using whole genome sequencing data) and epidemiological data, it is possible to reconstruct chains of transmission. The branching patterns of phylogenetic trees can be used to predict linkages between infected hosts, identify super spreaders and detect putative routes of transmission. In addition to tracking the spread of a virus in the present, it can reconstruct the processes that drove their spread in the past, determining when it first arose within a population (Grubaugh, Ladner, et al. 2019; Quick et al. 2020). Viral gene sequences were used in 2006 to reconstruct the spread of foot-and-mouth disease virus (FMDV) in the United Kingdom, and its transmission was traced from farm to farm (Cottam et al. 2006). Genomic data also played an important role in understanding the West African Ebola outbreak, as it identified unconventional transmissional chains and showed that most outbreaks were linked to persistently infected survivors, demonstrating sexual transmission of the virus (Arias et al. 2016; Diallo et al. 2016; Grubaugh, Ladner, et al. 2019).

Most phylogenetics-based transmission chain analyses have focused on mutations observed in viral consensus genomes, which represent the dominant variants within infected hosts. For slow-evolving viruses, this is performed under the assumption that a single individual will only be infected by a single strain of the virus. Some RNA viruses, however, evolve sufficiently rapidly to produce multiple new variants within a single round of infection. Knowledge of within-host diversity (containing intra-host single nucleotide variants (iSNVs)) of the virus at the population level, and how frequently this is transmitted, is critical for determining rates of adaptation and patterns of transmission. Additionally, newer methods incorporating viral iSNVs may increase the resolution of transmission chain analyses as multiple variants are transmitted between hosts (Lythgoe et al. 2021; Grubaugh et al. 2019).

Phylogeographic methods, which assess how an epidemic may unfold through time and space, have been transformed by genomic epidemiology. Using simple stochastic models, it provides location estimates for every ancestral node in a virus phylogeny, by reconstructing a detailed spatial history of virus spread from the origin of an outbreak (Figure 8). These analyses are enabled by the integration of virus genomics and diverse metadata sets, and are dependent on the timeliness of data generation and open data sharing (Grubaugh et al. 2019).

Phylogeographic analyses, for example, have revealed that multiple introductions of Zika virus were responsible for sustaining the 2016 outbreak in Florida (Grubaugh et al. 2017). Additionally, these analyses shed light on the factors driving the viral spread, such as geographic distances and population size. Critically, such analyses may only be capable of elucidating partial pictures of outbreak spread, and sampling biases may severely affect these analyses, especially when locations are poorly represented and lineages are missed (Quick et al. 2020). Moreover, pathogen genome sequences are of limited utility when viewed in isolation. They must be examined in the context of a constantly updated database of comparator strains, and the associated epidemiological and surveillance data (Gardy et al. 2015).

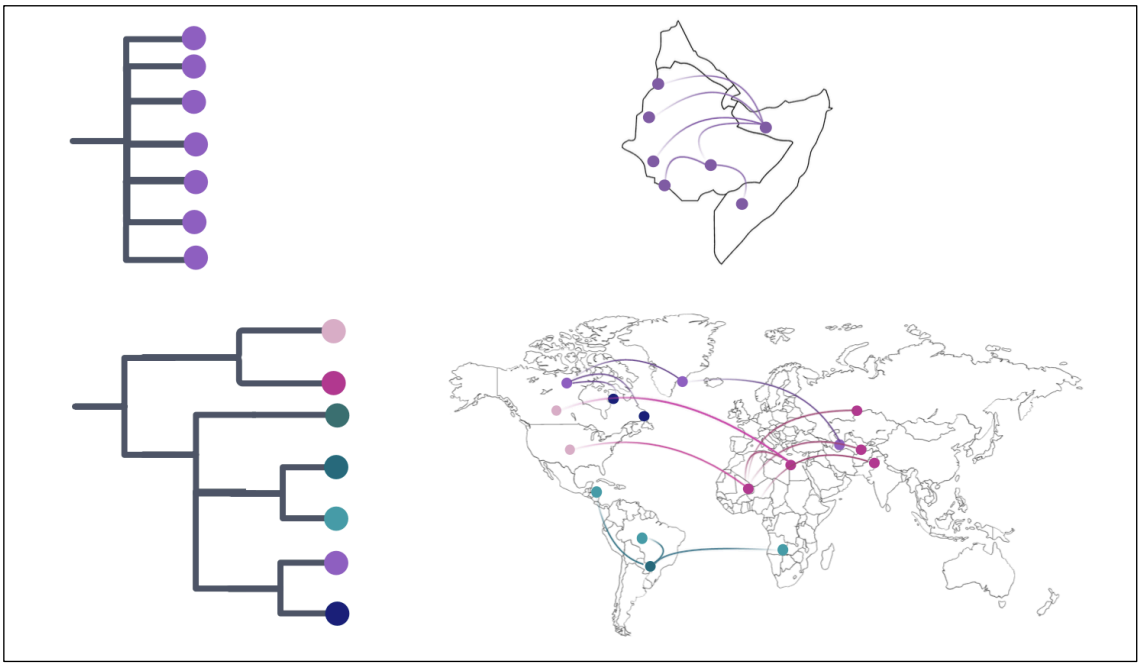


Figure 8. A diagram showing how phylogeographic maps support genomic analysis as an epidemic tracking tool; displaying the phylogenetic relationships between isolates and their putative transmissions on a map.

Once the outbreak has been brought under control or resolved, phylogenetic analyses can give insight into evolutionary patterns during inter-epidemic periods by comparing virus genome sequences sampled across different outbreaks. In aquaculture, this can answer the question of whether the virus in question was able to persist in the fish between outbreaks; whether each new outbreak has arisen from an endemically circulating lineage, or whether they represent

independent events from a different reservoir. Inter-epidemic analyses can also be used to reveal the nature of virus evolution and spread in reservoir species, and adaptation to its host, as well as assisting in vaccine development and drug design (Faria et al. 2016; Grubaugh, Gangavarapu, et al. 2019; Quick et al. 2020).

Genomic surveillance of aquaculture provides several benefits to fish farmers. Whole genome studies of ISKNV and other megalocytiviruses will enable genomic epidemiology and will provide information to enhance disease control in aquaculture (Fusianto et al. 2023). First, they can identify putative sources of infection by comparing the variant in their farms to existing datasets of known variants at other sites. Analysis of phylogeny with associated metadata (e.g. transportation logs; introduction of new fish etc; patterns of water movement) can identify sources of transmission. Second, altered phenotypic properties such as increased virulence, viral load, virion stability outside the host etc. can be linked to genomic variants of concern. When these new VOCs are detected in future events, farmers can act appropriately e.g. increasing distance between cages, culling stock to avoid transmission or moving affected cages outside of the main water course to avoid downstream transmission. Finally, since the absence of cases cannot directly indicate the absence of the virus, surveillance and control programs must remain active during inter-epidemic periods. With the availability of a sufficiently sensitive test, the continuous monitoring of water samples from the fish cages, in search of ISKNV variants present in the farms and the environment, could also be used as a precautionary method to monitor background prevalence and emergence of new variants without destructive sampling of the fish.

Effective genomic surveillance requires virus genome data sharing and standardisation of approaches (such as variant calling and phylogenomic approaches) during aquaculture outbreaks to provide relevant information to both farmers and decision makers. Setting up a national or regional information exchange between farmers and responsible parties should be both compulsory and centrally managed and funded to enable this (Assefa and Abunna 2018). Currently, the speed, nature and extent of virus genome data sharing is inconsistent, sometimes resulting in confusion over choosing the best practice (Gardy et al. 2015).

1.6.3 ***Whole Genome Surveillance Using a Tiling Amplicon Scheme:***

Although shotgun metagenomics (the process of sequencing a random subsample of total nucleic acid content in a mixed community sample), has been successfully applied to both virus discovery and diagnostics, direct metagenomic sequencing from clinical samples becomes challenging as genome coverage may be low or absent if viruses are present at low abundance in a sample with high levels of host nucleic acid background (Quick et al. 2017). Other challenges include requiring a complex procedure of sample preparation, expensive equipment and advanced bioinformatic training (Kiselev et al. 2020). To overcome these issues, and to generate complete viral genomes from clinical samples in an economic manner, target enrichment is often required. A multiplex PCR approach (tiling amplicon scheme), which produces amplicons that span the viral genome has been utilised for targeted enrichment of viral genomes from samples containing very few genome copies per reaction, shown in Figure 9 (Quick et al. 2017). Quick and et al, have developed a web-based primer design tool known as the 'Primal Scheme' to produce efficient multiplex primer schemes. By adapting this method for different viral targets, the Artic-network has developed a pipeline that is an end-to-end system for processing samples from viral outbreaks to generate real-time epidemiological information that is feasible for deployment and interpretation in resource-limited settings. Additionally, the exponential nature of PCR makes the technique robust to a large range of input titres, and has been successfully used by several groups studying viral outbreaks in humans, such as Ebola and Zika virus. Recently, the pipeline was adapted to sequence samples from the SARS-CoV-2 outbreak, supporting early sequencing efforts in many countries as it can be established rapidly to monitor outbreaks (Quick et al. 2016, 2017; Tyson et al. 2020; Resende et al. 2020).

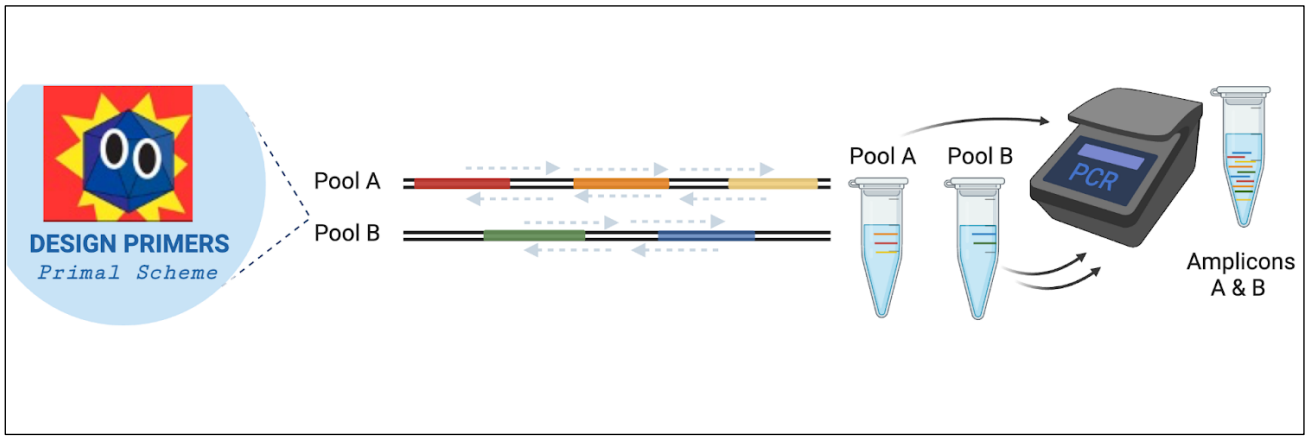


Figure 9. An illustration of multiplex tiling PCR and pooling; showing primers are designed using the PrimalScheme; primers in pool 'A' and 'B' overlap between but not within reactions; amplicons are generated by pool 'A' and 'B' and are pooled together.

Current sequencing approaches, based on second-generation sequencing platforms (e.g., Illumina and Ion Torrent), mean that only sequencing large batches of samples is economically viable, with a significant upfront cost for equipment and personnel. These approaches require a specialised laboratory and several days' turnover for the library preparation and sequencing, even under highly automated settings (Deeg et al. 2022). Additionally, these conventional sequencing technologies are difficult to deploy in Low- and middle-income countries, where availability of continuous power and cold chains, laboratory space, and trained personnel is restricted (Quick et al. 2016). The availability of the portable genome sequencer, MinION, developed by Oxford Nanopore Technologies (ONT), which is compact size and supports real-time data analysis, could facilitate the application of genomic sequencing in point-of-care testing for infectious diseases (Xu et al. 2018). The data throughput of MinION is high, allowing users to make more efficient use of the flow cell (and reduce cost) by multiplexing many samples in a single sequencing run, as ONT has developed PCR-free barcode sets (Native Barcode Expansion 96-PBC096), compatible with the R 9.4.1 flowcells, allowing multiplexing of up to 96 samples (Srivathsan et al. 2018). This has now been replaced with the most recent Rapid Barcoding Kit 96 (SQK-RBK110.96); compatible with the latest flow cell R10.4, achieving higher model read accuracy of over 99.1%, superior variation detection, lower false-discovery rate (FDR) in methylation calling, and comparable genome recovery rate (Ni et al. 2023). However, rapid changes in ONT platforms and chemistries

pose a challenge for low- and middle- income countries (LMIC) where access to the updated version is limited, and expensive to purchase all updated devices and reagents.

A challenge that remains with a PCR tiling approach to genome sequencing is that it has traditionally been used for viruses with small genomes, such as Ebola (19kbp) and SARS-CoV-2 (30kbp) (Quick et al. 2016; Itokawa et al. 2020). This is because the length of PCR products are limited by the efficiency of the DNA polymerase used during the extension step, and it is important to choose the correct, high accuracy polymerase capable of amplifying such long amplicons (Warr et al. 2021). Longer genomes require a larger number of 'tiles' to span the genome, which increases the likelihood of interference between primer pairs. In addition, increased viral genome length requires increasingly complex mechanisms for proof-reading to avoid over-accumulation of deleterious mutations. For instance, the genome of SARS-CoV-2 is unusually long for an RNA virus, encoding proof-reading machinery. RNA viruses mutate at a rate of 10^{-4} to 10^{-6} substitutions per nucleotide per infection (s/n/i) while the rate in DNA viruses is much slower (10^{-6} to 10^{-8} s/n/i) (Peck and Lauring 2018). RNA virus evolution involves error-prone polymerases, and a constant interplay of mutation and fitness-based selection. In contrast, DNA viruses are less error-prone and can perform error-correction provided by the complementary strand and can also encode specific proof-reading DNA polymerases (Choi 2012). Thus, outbreaks of dsDNA viral diseases accumulate mutations at a much slower rate, potentially limiting the utility of genomic surveillance, at least in the short-term as transmission rates could potentially outstrip mutation rates by several orders of magnitude. Nevertheless, the larger size of many DNA virus genomes allows for the accumulation of genetic variation in every round of viral replication, and hotspots of mutation such as hypervariable regions have been identified in viral genomes (Szpara and Van Doorslaer 2019).

Applying a tiled amplicon sequencing approach to segmented RNA viruses could create a further challenge. Frequent mutations would mean that a constant update and re-evaluation of primers selected will be required to mitigate mutation related impacts. Additionally, Influenza A and TiLV, both segmented viruses are well known for their ability to perform reassortment. Each segment of TiLV has

different degrees of nucleotide and amino acid variation (Chaput et al. 2020; Li and Chen 2014). This results in the failure of the phylogenetic analysis of one segment of the viral genome to give a complete picture for tracking and predicting their movement between countries. Therefore it is highly recommended to use whole genome sequencing wherever possible (Croville et al. 2023; Chaput et al. 2020), directly from samples, as culture and concentration of viruses could affect the sequencing results and should be avoided. Other challenges for tiled amplicon sequencing of segmented RNA viruses are that these segments are often short, and capturing a single tile on a segment would result in loss of enough coverage at the ends of the consensus sequences.

Limitations of tiled amplicon sequencing include its inability to discover new viruses or sequencing highly diverse or recombinant viruses, as primer schemes are highly specific. Amplicon sequencing is prone to coverage dropouts that may result in incomplete genome coverage, especially at lower abundance, in regions of high mutation rates, and the loss of both 5′ and 3′ regions that fall outside the outer primer binding positions (Quick et al. 2017). Several new viruses infecting aquatic organisms have been discovered through Next-generation sequencing (NGS) methods (Kibenge 2019). Advanced molecular diagnostics, such as those being designed in this thesis, will assist farmers in tracking these infections at different points in the tilapia life cycle once they have been identified (Howell 2019).

1.6.4 ***Water Sampling for Viruses***

Monitoring the spread of viral pathogens in water is crucial for an effective response, and understanding the viral lineages that constitute the infections can uncover the origins and transmission patterns of outbreaks, as well as detecting the emergence of novel variants before its detection in clinical samples. Due to the often high infectivity and rapid transmission of viruses, comprehensive screening of individuals is often challenging, particularly in cases with mild or no symptoms (Farkas et al. 2020; Child et al. 2023). In human viral genomic surveillance, wastewater sampling has been considered a vital source to understand mutations and infection dynamics at a population level, as well as an early indicator of new outbreaks.

Following the SARS-CoV-2 outbreak, methods have been developed for monitoring wastewater to be used to detect the arrival and subsequent decline of outbreaks and associated variants. For many years, research on the surveillance of viruses in wastewater has been considered a vital source to understand mutations and infection dynamics for human viruses, especially for detecting enteroviruses, such as Polio, Hepatitis A, and Retroviruses. This is usually done using accurate and validated methods, with subsequent risk analysis and modelling, which is paramount in understanding the dynamics of viral outbreaks (Farkas et al. 2020; Kittigul et al. 2000; Dharmadhikari et al. 2022). Viral concentration is usually followed by viral quantification, using amplification-based viral quantification. Culture-based analysis of viral infectivity is rarely performed on wastewater samples due difficulty to maintain the virus *in vitro* (Farkas et al. 2020). Next Generation Sequencing (NGS) platforms are then used to determine genomic variants, providing a comprehensive depiction of infection dynamics in the population (Dharmadhikari et al. 2022).

In aquaculture, water-transmitted viral pathogens are a significant threat, challenging fish welfare and the economy of this industry. They have been difficult to control due to an increased susceptibility among hosts and limited understanding of the transmission dynamics. Traditional sampling of fish is destructive, limiting farmer engagement, and relies on a costly, time-consuming, and resource-demanding approach based on routine sampling. Water sampling surveillance could reduce, to a great extent, the overall sacrifice of fish, as well as being a straightforward, cost-efficient, and timesaving, approach for detecting viruses in fish farms. All these factors make water monitoring a reliable tool to predict or prevent outbreaks in aquaculture, or even to ensure that the water is free from the causative agent before resuming farming following an outbreak (Haramoto et al. 2009; Bernhardt et al. 2021).

Some of the early methods used to concentrate fish viral pathogens in water, used Tangential Flow Filtration (TFF). This proved to be an effective method for concentrating Infectious hematopoietic necrosis virus (IHNV) and IPNV from large volumes of water (Watanabe et al. 1988). For the sensitive detection of viruses in wastewater, samples are usually concentrated before quantification, and are often centrifuged or filtered to eliminate debris, followed by

electronegative membrane filtration, ultrafiltration, polyethylene glycol (PEG) precipitation ultracentrifugation, or precipitation with ammonium sulphate, enabling a 20x-800x concentration (Farkas et al. 2020; Child et al. 2023). Recently, Bernhardt et al have carried out a study for the concentration and detection of Salmon alphavirus (SAV) by filtration through an electronegative membrane filter with subsequent rinsing of the filter with a lysis buffer (Bernhardt et al. 2021). Although concentration methods are usually inexpensive and easy to set-up, they can be time-consuming and difficult to perform with high sample throughput, especially when high turbidity samples are processed (Farkas et al. 2020; Child et al. 2023). Another disadvantage of these methods is the co-concentration of organic compounds, which often interfere with downstream virus detection such as the PCR assays and extraction. Concentration efficiency may vary among different samples, therefore, appropriate process controls should be added to the sample to estimate viral recoveries (Farkas et al. 2020; Kittigul et al. 2000).

1.6.5 ***In-field Sequencing of ISKNV:***

The need for accurate, rapid and on-site diagnosis of infectious disease grows as globalised human activity accelerates (Boykin et al. 2019). To achieve genomic surveillance for aquatic viruses affecting the growth of aquaculture, sequencing technologies are needed. Recent advances in sequencing platforms such as the Oxford Nanopore minION allow real-time sequencing on a pocket-sized portable sequencer that requires little library preparation, therefore enabling sequencing in remote locations (Deeg et al. 2022). These sequencers can be used with a portable miniPCR for amplicon generation, powered easily by connecting to a mobile phone. The portability of this technology enables training for scientists or veterinarians with little molecular knowledge in remote or resource limited regions of the world, eliminating not only the need for transporting samples for diagnostics, but also their travelling abroad to obtain the training. Using a portable lab could be possible in areas such as Ghana, using a Pelicase to transfer equipment and reagents to facilitate the movement of all the lab necessities.

For nanopore sequencing, real-time basecalling to fastq files can be achieved by connecting to a portable laptop or a GPU-enabled basecalling unit, such as the Nanopore computation unit MinIT or the MK1C. However, several technical hurdles to adapting Nanopore sequencing do exist. While Nanopore sequencing can yield extremely long reads, the number of sequencing pores and their loading rate is limited, resulting in low throughput when sequencing short reads such as amplicons. One of the major barriers to producing sequencing outputs in the field is the lack of a simple, quick and effective method to extract DNA from a sample without the need for laboratory equipment requiring mains power and space, items such as benchtop centrifuges, fridges, freezers and temperature- sensitive extraction kits which can be bulky and rely on traditional laboratory infrastructure (Boykin et al. 2019). An additional issue for using Nanopore sequencing is the low accuracy of the sequencing platform at the time of this project using the R9.4.1 flow cells, ranging from 5% to 15%, obscuring true variation, consequently, creating a challenge to determine VOC-defining SNPs (Deeg et al. 2022; Liu et al. 2022). This low accuracy requires high alignment coverage at SNP locations to ensure accurate SNP calling (Figure 10). However, newer Nanopore flow cells promise greatly increased accuracy due to a longer barrel and dual reader head in the pore protein and have recently become available, replacing the R9.4.1 flow cells. This updated flow cell, the R10.4 technology, is therefore expected to greatly improve sequencing accuracy and possibly allow the lowering of alignment thresholds for SNP calling, thereby increasing the throughput more than two-fold (Deeg et al. 2022). Benchmarking has confirmed R10.4 outperforming R9.4.1 for high read accuracy and variant detection, however, genome recovery rates of R10.4 and R9.4.1 are comparable, and the increased accuracy of R10.4 flow cells is coupled with a decreased yield (Ni et al. 2023). Problems related to the availability of ONT suppliers in certain countries, could also be a limiting factor, where even centralised laboratories have supply chain issues and lack MinION-specific sequencing skills (Wasswa et al. 2022). Undoubtedly, the advantages of adopting in-field sequencing for aquatic, infectious viruses has the potential to outweigh its limitations.

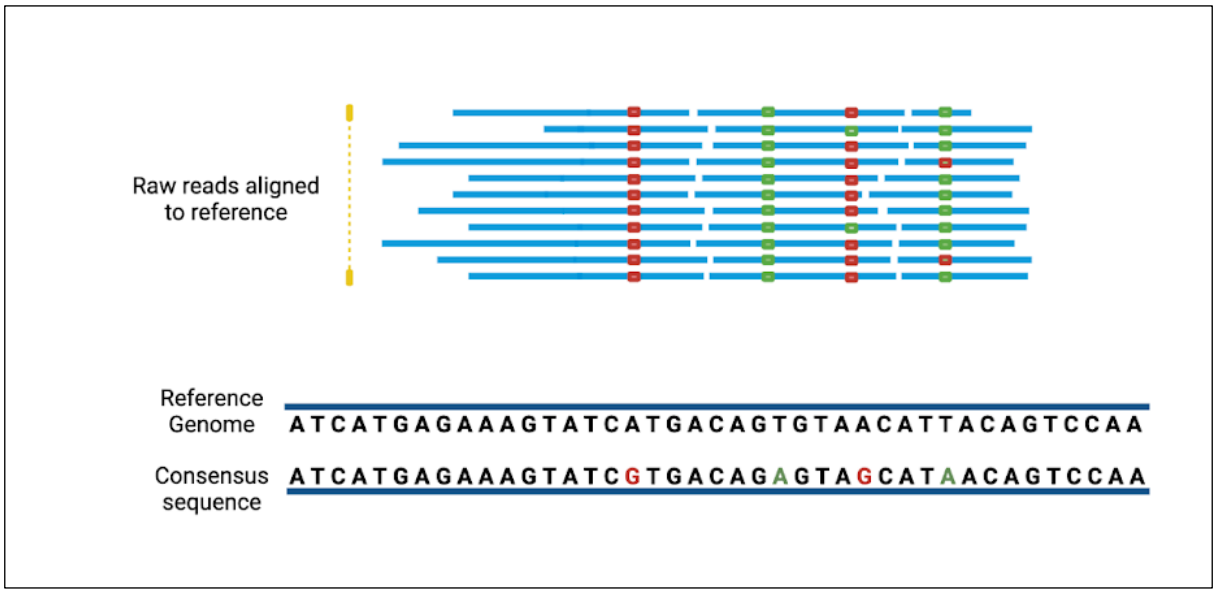


Figure 10. Single nucleotide polymorphisms (SNPs) called in genome sites where the consensus sequence has bases that differ from the reference genome. Error-prone reads can be corrected with high genome coverage when generating consensus sequences.

1.7 Tilapia Lake Virus: A segmented RNA virus affecting the growth of tilapia aquaculture

In the late 2000s, large losses of farmed tilapia were recorded throughout Israel. This was subsequently identified and termed Tilapia Lake Virus (TiLV) (Eyngor et al. 2014). Tilapia mortality associated with TiLV infections has also been described in Ecuador (Tsofack et al. 2017), Egypt (Fathi et al. 2017), Thailand (Dong et al. 2017) India, Malaysia (Amal et al. 2018) and the Philippines (OIE 2017). Mortality levels between 20% and 90% have been reported in farmed and wild tilapia populations (Jansen et al. 2018). In Egypt, during the summer months of 2015, TiLV mortalities indicated a potential economic impact of around USD 100 million, with 37% of fish farms being affected (Fathi et al. 2017). This has resulted in a huge impact on global food security and nutrition (Chaput et al. 2020; Kibenge 2019). Due to the international trade of tilapia for more than 50 years, TiLV may have been circulating worldwide through movement of live fish for aquaculture in the absence of knowledge of the existence of an associated risk (Dong, Ataguba, et al. 2017; Jansen, Dong, and Mohan 2018; Kibenge 2019).

TiLV, an Amnoonviridae, has a genome of 10 segments of linear negative sense single stranded RNA. Infections vary widely in severity (from asymptomatic to extremely lethal) for reasons that are currently unresolved, but reassortment may be a contributing factor. Furthermore, in other segmented RNA viruses, such as influenza, reassortment has been the cause of the sudden emergence of extremely virulent strains (Chaput et al. 2020).

Currently, there is no cure for viral diseases in aquaculture and while vaccines and selective breeding have proved successful in reducing the severity of some viral diseases, there are currently severe knowledge gaps relating to TiLV, one of the most significant emerging pathogens in tilapia aquaculture with no effective, affordable vaccines yet available (Jansen et al. 2018). Tracking TiLV's movement across borders is crucial for minimising its impact on farmed and wild fish populations, and methods described previously for ISKNV could be used to monitor its spread. As a segmented virus, this poses a challenge for tiled PCR approaches and its ability to reassort is an additional hurdle to identifying mutations for applying genomic surveillance.

Finally, effective disease control in aquaculture requires knowledge of the pathogens and their hosts and a detailed understanding of the epidemiology of the disease (FAO 2019). Progressive farming practices now enable discovery of emerging viruses through surveillance and laboratory diagnosis (Kibenge 2019). The best option for controlling the continuous emergence of viral diseases in aquaculture is ideally at the farm level, where better knowledge about the viral diseases and their improved diagnosis, inspection and surveillance programs translate into higher profits for the farmer and, therefore, motivation for a sustainable industry (Kibenge 2019).

The overall aims of this thesis work were to develop further understanding on the strains of ISKNV in outbreaks circulating Lake Volta, Ghana. This was conducted through the development of methods to track the phylogeography of ISKNV across Lake Volta through the application of sequencing methods, including for application in-field. The tiled PCR method developed for tracking ISKNV was further investigated on a viral outbreak in tilapia for the RNA virus TiLV. Collectively this work sought to appraise the tiled PCR method as an in-field tool for genome surveillance of both DNA and RNA viruses of major importance to aquaculture.

The specific thesis objectives were:

1) To develop a tiling PCR protocol that enables whole genome sequencing of ISKNV and investigate its capability to detect sufficient non-synonymous variation to classify slow-evolving dsDNA viruses into variants that are distinct between outbreaks, host species and/or geographical spread (**Chapter 2**).

2) To develop and apply methods to detect the presence of ISKNV in concentrated water samples from Lake Volta, and via amplicon sequencing enable monitoring of ISKNV variants present in a farm setting as a non-invasive alternative to the destructive sampling of fish (**Chapter 3**).

3) Investigate the feasibility of performing In-field direct concentration and detection measures of ISKNV in water and tilapia fish from cages on Lake Volta, Ghana. (**Chapter 3**).

4) Evaluation of a tiled PCR method for a segmented RNA virus - *Tilapia Lake Virus* affecting the growth of tilapia aquaculture for more than a decade **(Chapter 4).**

The thesis concludes with an evaluation on the key findings of the work presented, their significance and future avenues for research to better support genome surveillance methods for disease tracking, control and prevention in aquaculture (**Chapter 5 & 6).** An overview of the objectives of each chapter of the thesis are shown in Figure 11.

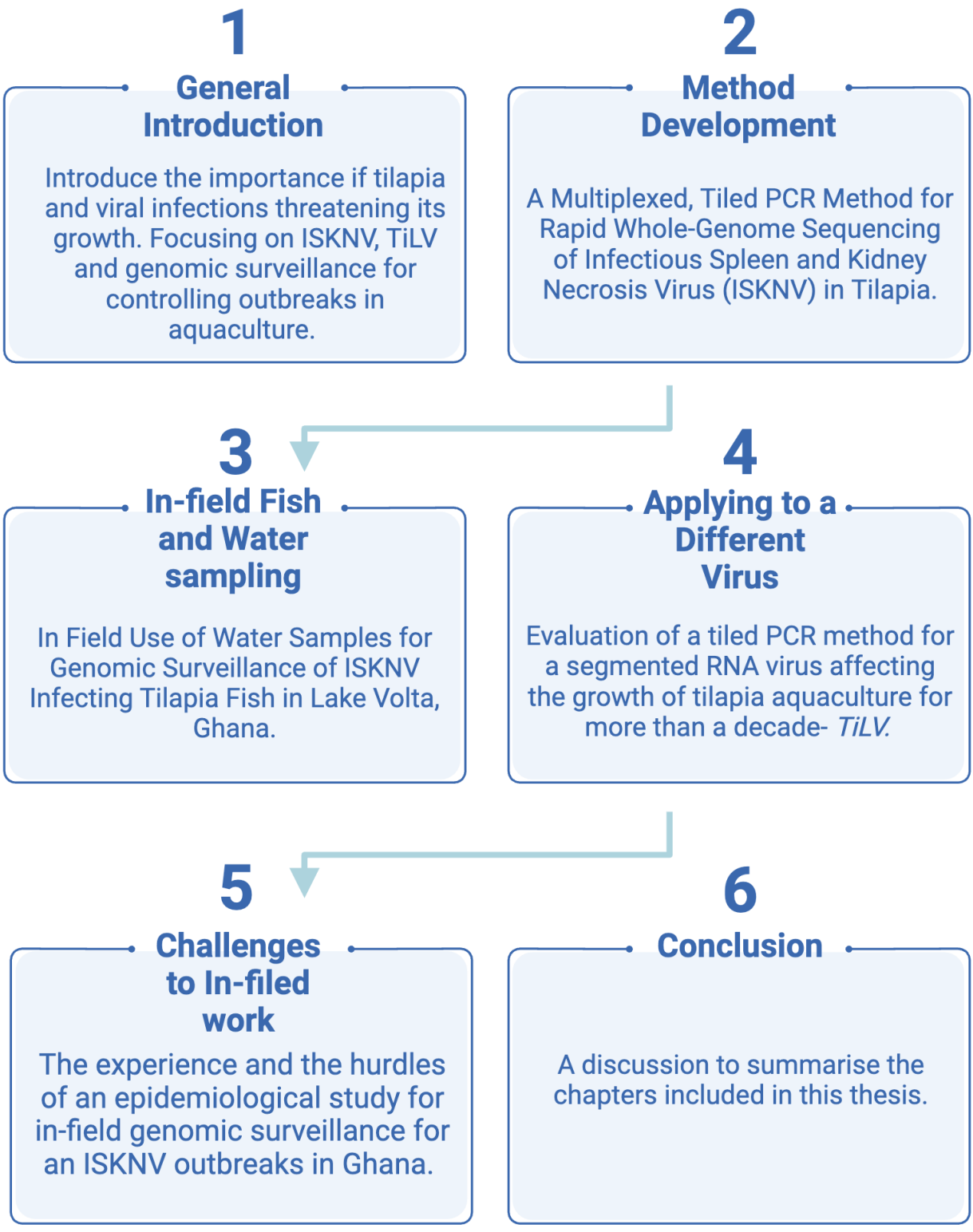


Figure 11. An overview of the chapters of this Thesis

References

Alathari, Shayma, Dominique L. Chaput, Luis M. Bolaños, Andrew Joseph, Victoria L. N. Jackson, David Verner-Jeffreys, Richard Paley, Charles R. Tyler, and Ben Temperton. 2023. “A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia.” *Viruses* 15 (4): 965.

Altmann, Daniel M., Rosemary J. Boyton, and Rupert Beale. 2021. “Immunity to SARS-CoV-2 Variants of Concern.” *Science*.

Amal, M. N. A., C. B. Koh, M. Nurliyana, M. Suhaiba, Z. Nor-Amalina, S. Santha, K. P. Diyana-Nadhirah, M. T. Yusof, M. Y. Ina-Salwany, and M. Zamri-Saad. 2018. “A Case of Natural Co-Infection of Tilapia Lake Virus and Aeromonas Veronii in a Malaysian Red Hybrid Tilapia (Oreochromis Niloticus×O. Mossambicus) Farm Experiencing High Mortality.” *Aquaculture* 485 (February): 12–16.

Amenyogbe, Eric, Gang Chen, Zhongliang Wang, Mingde Lin, Xiaoying Lu, Denicia Atujona, and Emmanuel D Abarike. 2018. “A Review of Ghanas Aquaculture Industry.” *Journal of Aquaculture Research & Development* 09 (08). https://doi.org/10.4172/2155-9546.1000545.

Arias, Armando, Simon J. Watson, Danny Asogun, Ekaete Alice Tobin, Jia Lu, My V. T. Phan, Umaru Jah, et al. 2016. “Rapid Outbreak Sequencing of Ebola Virus in Sierra Leone Identifies Transmission Chains Linked to Sporadic Cases.” *Virus Evolution* 2 (1): vew016.

Asiedu, Berchie, F. K. E. Nunoo, and Seidu Iddrisu. 2017. “Prospects and Sustainability of Aquaculture Development in Ghana, West Africa.” *Cogent Food & Agriculture* 3 (1): 1349531.

Assefa, Ayalew, and Fufa Abunna. 2018. “Maintenance of Fish Health in Aquaculture: Review of Epidemiological Approaches for Prevention and Control of Infectious Disease of Fish.” *Veterinary Medicine International* 2018 (February): 5432497.

Bernhardt, Lisa-Victoria, Atle Lillehaug, Lars Qviller, Simon Chioma Weli, Estelle Grønneberg, Hanne Nilsen, and Mette Myrmel. 2021. “Early Detection of Salmonid Alphavirus in Seawater from Marine Farm Sites of Atlantic Salmon Salmo Salar.” *Diseases of Aquatic Organisms* 146 (September): 41–52.

Boykin, Laura, Ammar Ghalab, Bruno Rossitto De Marchi, Anders Savill, James M. Wainaina, Tonny Kinene, Stephen Lamb, et al. 2018. “Real Time Portable Genome Sequencing for Global Food Security.” *F1000Research* 7 (1101): 1101.

Boykin, Laura M., Peter Sseruwagi, Titus Alicai, Elijah Ateka, Ibrahim Umar Mohammed, Jo-Ann L. Stanton, Charles Kayuki, et al. 2019. “Tree Lab: Portable Genomics for Early Detection of Plant Viruses and Pests in Sub-Saharan Africa.” *Genes* 10 (9). https://doi.org/10.3390/genes10090632.

CDC. 2021. “Cases, Data, and Surveillance.” March 20, 2021. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-

surveillance.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fmore%2Fscientific-brief-emerging-variant.html.

Chaput, Dominique L., David Bass, Md Mehedi Alam, Neaz Al Hasan, Grant D. Stentiford, Ronny van Aerle, Karen Moore, John P. Bignell, Mohammad Mahfujul Haque, and Charles R. Tyler. 2020a. "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.

Child, Harry T., Paul A. O'Neill, Karen Moore, William Rowe, Hubert Denise, David Bass, Matthew J. Wade, et al. 2023. "Optimised Protocol for Monitoring SARS-CoV-2 in Wastewater Using Reverse Complement PCR-Based Whole-Genome Sequencing." *PloS One* 18 (4): e0284211.

Choi, Kyung H. 2012. "Viral Polymerases." *Advances in Experimental Medicine and Biology* 726: 267–304.

Cottam, Eleanor M., Daniel T. Haydon, David J. Paton, John Gloster, John W. Wilesmith, Nigel P. Ferris, Geoff H. Hutchings, and Donald P. King. 2006. "Molecular Epidemiology of the Foot-and-Mouth Disease Virus Outbreak in the United Kingdom in 2001." *Journal of Virology* 80 (22): 11274–82.

Croville, Guillaume, Mathilda Walch, Laetitia Lèbre, Sonia Silva, Fabien Filaire, and Jean-Luc Guérin. 2023. "An Amplicon-Based Nanopore Sequencing Workflow for Rapid Tracking of Avian Influenza Outbreaks, France, 2020-2022." *BioRxiv*. https://doi.org/10.1101/2023.05.15.538689.

Deeg, Christoph M., Ben J. G. Sutherland, Tobi J. Ming, Colin Wallace, Kim Jonsen, Kelsey L. Flynn, Eric B. Rondeau, Terry D. Beacham, and Kristina M. Miller. 2022. "In-Field Genetic Stock Identification of Overwintering Coho Salmon in the Gulf of Alaska: Evaluation of Nanopore Sequencing for Remote Real-Time Deployment." *Molecular Ecology Resources* 22 (5): 1824–35.

Dharmadhikari, Tanmay, Vinay Rajput, Rakeshkumar Yadav, Radhika Boargaonkar, Dhawal Patil, Saurabh Kale, Sanjay P. Kamble, Syed G. Dastager, and Mahesh S. Dharne. 2022. "High Throughput Sequencing Based Direct Detection of SARS-CoV-2 Fragments in Wastewater of Pune, West India." *The Science of the Total Environment* 807 (Pt 3): 151038.

Diallo, Boubacar, Daouda Sissoko, Nicholas J. Loman, Hadja Aïssatou Bah, Hawa Bah, Mary Claire Worrell, Lya Saidou Conde, et al. 2016. "Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus Persistence in Seminal Fluid for More Than 500 Days." *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 63 (10): 1353–56.

Dong, Chuanfu, Xiaopeng Xiong, Yongwen Luo, Shaoping Weng, Qing Wang, and Jianguo He. 2013. "Efficacy of a Formalin-Killed Cell Vaccine against Infectious Spleen and Kidney Necrosis Virus (ISKNV) and Immunoproteomic Analysis of Its Major Immunogenic Proteins." *Veterinary Microbiology* 162 (2–4): 419–28.

Dong, H. T., S. Jitrakorn, P. Kayansamruaj, N. Pirarat, C. Rodkhum, T. Rattanarojpong, S. Senapin, and V. Saksmerprome. 2017. "Infectious Spleen and Kidney Necrosis Disease (ISKND) Outbreaks in Farmed Barramundi

(Lates Calcarifer) in Vietnam." *Fish & Shellfish Immunology* 68 (September): 65–73.

Dong, H. T., S. Siriroob, W. Meemetta, W. Santimanawong, W. Gangnonngiw, N. Pirarat, P. Khunrae, T. Rattanarojpong, R. Vanichviriyakit, and S. Senapin. 2017. "Emergence of Tilapia Lake Virus in Thailand and an Alternative Semi-Nested RT-PCR for Detection." *Aquaculture* 476 (July): 111–18.

Dong, Ha Thanh, Vuong Viet Nguyen, Hai Dinh Le, Pakkakul Sangsuriya, Sarocha Jitrakorn, Vanvimon Saksmerprome, Saengchan Senapin, and Channarong Rodkhum. 2015. "Naturally Concurrent Infections of Bacterial and Viral Pathogens in Disease Outbreaks in Cultured Nile Tilapia (Oreochromis Niloticus) Farms." *Aquaculture* 448 (November): 427–35.

El-Sayed, Abdel-fattah M., and Kevin Fitzsimmons. 2023. "From Africa to the World—The Journey of Nile Tilapia." *Reviews in Aquaculture* 15 (S1): 6–21.

Eyngor, Marina, Rachel Zamostiano, Japhette Esther Kembou Tsofack, Asaf Berkowitz, Hillel Bercovier, Simon Tinman, Menachem Lev, et al. 2014. "Identification of a Novel RNA Virus Lethal to Tilapia." *Journal of Clinical Microbiology* 52 (12): 4137–46.

2018a. *The State of World Fisheries and Aquaculture 2018: Meeting the Sustainable Development Goals*. {Food and Agriculture Organization of the United Nations}.

2018b. *2016 THE STATE OF WORLD FISHERIES AND AQUACULTURE: CONTRIBUTING TO FOOD SECURITY AND NUTRITION FOR ALL*. Food & Agriculture Org.

FAO. 2019. "The State of the World's Aquatic Genetic Resources for Food and Agriculture." ReliefWeb. 2019. https://reliefweb.int/report/world/state-world-s-aquatic-genetic-resources-food-and-agriculture.

FAO. 2020. "The State of World Fisheries and Aquaculture 2020." http://www.fao.org/3/ca9229en/ca9229en.pdf.https://doi.org/10.4060/CA9229EN.

FAO. 2022. *Blue Transformation – Roadmap 2022–2030: A Vision for FAO's Work on Aquatic Food Systems*. FAO. Accessed August 8, 2023. https://doi.org/10.4060/cc0461en.

Faria, Nuno Rodrigues, Ester C. Sabino, Marcio R. T. Nunes, Luiz Carlos Junior Alcantara, Nicholas J. Loman, and Oliver G. Pybus. 2016. "Mobile Real-Time Surveillance of Zika Virus in Brazil." *Genome Medicine* 8 (1): 97.

Farkas, Kata, Luke S. Hillary, Shelagh K. Malham, James E. McDonald, and David L. Jones. 2020. "Wastewater and Public Health: The Potential of Wastewater Surveillance for Monitoring COVID-19." *Current Opinion in Environmental Science & Health* 17 (October): 14–20.

Fathi, Mohamed, Cathryn Dickson, Malcolm Dickson, William Leschen, Johanna Baily, Fiona Muir, Kristina Ulrich, and Manfred Weidmann. 2017. "Identification of Tilapia Lake Virus in Egypt in Nile Tilapia Affected by 'Summer Mortality' Syndrome." *Aquaculture* 473 (April): 430–32.

Figueiredo, Henrique César Pereira, Guilherme Campos Tavares, Fernanda Alves Dorella, Júlio César Câmara Rosa, Sóstenes Apolo Correia Marcelino, and Felipe Pierezan. 2020. "First Report of Infectious Spleen and Kidney Necrosis

Virus in Nile Tilapia in Brazil." *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.10.08.331991.

Figueiredo, Henrique César Pereira, Guilherme Campos Tavares, Fernanda Alves Dorella, Júlio César Câmara Rosa, Sóstenes Apolo Correia Marcelino, Felipe Pierezan, and Felipe Luiz Pereira. 2022. "First Report of Infectious Spleen and Kidney Necrosis Virus in Nile Tilapia in Brazil." *Transboundary and Emerging Diseases* 69 (5): 3008–15.

Fu, Weixuan, Yong Li, Yuting Fu, Wenfeng Zhang, Panpan Luo, Qianqian Sun, Fangzhao Yu, et al. 2023. "The Inactivated ISKNV-I Vaccine Confers Highly Effective Cross-Protection against Epidemic RSIV-I and RSIV-II from Cultured Spotted Sea Bass Lateolabrax Maculatus." *Microbiology Spectrum* 11 (3): e0449522.

Fu, Xiaozhe, Ningqiu Li, Qiang Lin, Huizhi Guo, Lihui Liu, Zhibin Huang, and Shuqin Wu. 2015. "Early Protein ORF086 Is an Effective Vaccine Candidate for Infectious Spleen and Kidney Necrosis Virus in Mandarin Fish Siniperca Chuatsi." *Fish & Shellfish Immunology* 46 (2): 200–205.

Fu, Xiaozhe, Ningqiu Li, Lihui Liu, Qiang Lin, Fang Wang, Yingtiao Lai, Haiming Jiang, Houjun Pan, Cunbin Shi, and Shuqin Wu. 2011. "Genotype and Host Range Analysis of Infectious Spleen and Kidney Necrosis Virus (ISKNV)." *Virus Genes* 42 (1): 97–109.

Fusianto, Cahya, Paul M. Hick, and Joy A. Becker. 2019. "Stability of Infectious Spleen and Kidney Necrosis Virus and Susceptibility to Physical and Chemical Disinfectants." *Aquaculture* 506 (May): 104–11.

Fusianto, Cahya, Paul M. Hick, Murwantoko, Afri Herlambang, Richard J. Whittington, and Joy A. Becker. 2021. "Outbreak Investigation Attributes Infectious Spleen and Kidney Necrosis Virus as a Necessary Cause of a Mortality Epidemic in Farmed Grouper (Epinephelus Spp.) in Bali, Indonesia." *Aquaculture Reports* 20 (July): 100723.

Fusianto, Cahya K., Joy A. Becker, Kuttichantran Subramaniam, Richard J. Whittington, Samantha A. Koda, Thomas B. Waltzek, Undefined Murwantoko, and Paul M. Hick. 2023. "Genotypic Characterization of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Southeast Asian Aquaculture." *Transboundary and Emerging Diseases* 2023 (March). https://doi.org/10.1155/2023/6643006.

Gardy, Jennifer, Nicholas J. Loman, and Andrew Rambaut. 2015. "Real-Time Digital Pathogen Surveillance - the Time Is Now." *Genome Biology* 16 (1): 155.

Ghosh, Smita, Patrick K. Moonan, Lauren Cowan, Juliana Grant, Steve Kammerer, and Thomas R. Navin. 2012. "Tuberculosis Genotyping Information Management System: Enhancing Tuberculosis Surveillance in the United States." *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 12 (4): 782–88.

Go, J., T. B. Waltzek, K. Subramaniam, S. C. Yun, J. M. Groff, I. G. Anderson, R. Chong, et al. 2016. "Detection of Infectious Spleen and Kidney Necrosis Virus (ISKNV) and Turbot Reddish Body Iridovirus (TRBIV) from Archival Ornamental Fish Samples." *Diseases of Aquatic Organisms* 122 (2): 105–23.

Go, Jeffrey, and Richard Whittington. 2006. “Experimental Transmission and Virulence of a Megalocytivirus (Family Iridoviridae) of Dwarf Gourami (Colisa Lalia) from Asia in Murray Cod (Maccullochella Peelii Peelii) in Australia.” *Aquaculture* 258 (1): 140–49.

Grubaugh, Nathan D., Jason T. Ladner, Moritz U. G. Kraemer, Gytis Dudas, Amanda L. Tan, Karthik Gangavarapu, Michael R. Wiley, et al. 2017. “Genomic Epidemiology Reveals Multiple Introductions of Zika Virus into the United States.” *Nature* 546 (7658): 401–5.

Grubaugh, Nathan D., Jason T. Ladner, Philippe Lemey, Oliver G. Pybus, Andrew Rambaut, Edward C. Holmes, and Kristian G. Andersen. 2019. “Tracking Virus Outbreaks in the Twenty-First Century.” *Nature Microbiology* 4 (1): 10–19.

Haramoto, E., M. Kitajima, H. Katayama, T. Ito, and S. Ohgaki. 2009. “Development of Virus Concentration Methods for Detection of Koi Herpesvirus in Water.” *Journal of Fish Diseases* 32 (3): 297–300.

He, J. G., M. Deng, S. P. Weng, Z. Li, S. Y. Zhou, Q. X. Long, X. Z. Wang, and S. M. Chan. 2001. “Complete Genome Analysis of the Mandarin Fish Infectious Spleen and Kidney Necrosis Iridovirus.” *Virology* 291 (1): 126–39.

He, J. G., K. Zeng, S. P. Weng, and S-M Chan. 2002. “Experimental Transmission, Pathogenicity and Physical–Chemical Properties of Infectious Spleen and Kidney Necrosis Virus (ISKNV).” *Aquaculture* 204 (1): 11–24.

Howell, Megan. 2019. “Seven Deadly Fins: Study Sheds Light on Key Tilapia Viruses.” December 3, 2019. https://thefishsite.com/articles/seven-deadly-fins-study-sheds-light-on-key-tilapia-viruses.

Islam, Sk Injamamul, Sheikh Sunzid Ahmed, Nasim Habib, Md Akib Ferdous, Saloa Sanjida, and Moslema Jahan Mou. 2023. “High-Throughput Virtual Screening of Marine Algae Metabolites as High-Affinity Inhibitors of ISKNV Major Capsid Protein: An Analysis of in-Silico Models and DFT Calculation to Find Novel Drug Molecules for Fighting Infectious Spleen and Kidney Necrosis Virus (ISKNV).” *Heliyon* 9 (6): e16383.

Itokawa, Kentaro, Tsuyoshi Sekizuka, Masanori Hashino, Rina Tanaka, and Makoto Kuroda. 2020. “Disentangling Primer Interactions Improves SARS-CoV-2 Genome Sequencing by the ARTIC Network’s Multiplex PCR.” *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.03.10.985150.

Jansen, Mona Dverdal, Ha Thanh Dong, and Chadag Vishnumurthy Mohan. 2018. “Tilapia Lake Virus: A Threat to the Global Tilapia Industry?” *Reviews in Aquaculture* 46 (May): 120.

Jeong, Joon Bum, Ho Yeoul Kim, Lyu Jin Jun, Ji Hyo Lyu, Nam Gyu Park, Joong Kyun Kim, and Hyun Do Jeong. 2008. “Outbreaks and Risks of Infectious Spleen and Kidney Necrosis Virus Disease in Freshwater Ornamental Fishes.” *Diseases of Aquatic Organisms* 78 (3): 209–15.

Jung-Schroers, Verena, Mikolaj Adamek, Peter Wohlsein, Jan Wolter, Helmut Wedekind, and Dieter Steinhagen. 2016. “First Outbreak of an Infection with Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Ornamental Fish in Germany.” *Diseases of Aquatic Organisms* 119 (3): 239–44.

Kibenge, Frederick Sb. 2019. “Emerging Viruses in Aquaculture.” *Current Opinion in Virology* 34 (February): 97–103.

Kiselev, Daniel, Alina Matsvay, Ivan Abramov, Vladimir Dedkov, German Shipulin, and Kamil Khafizov. 2020. “Current Trends in Diagnostics of Viral Infections of Unknown Etiology.” *Viruses* 12 (2). https://doi.org/10.3390/v12020211.

Kiselinova, Maja, Alexander O. Pasternak, Ward De Spiegelaere, Dirk Vogelaers, Ben Berkhout, and Linos Vandekerckhove. 2014. “Comparison of Droplet Digital PCR and Seminested Real-Time PCR for Quantification of Cell-Associated HIV-1 RNA.” *PloS One* 9 (1): e85999.

Kittigul, L., B. Raengsakulrach, S. Siritantikorn, R. Kanyok, F. Utrarachkij, P. Diraphat, V. Thirawuth, et al. 2000. “Detection of Poliovirus, Hepatitis A Virus and Rotavirus from Sewage and Water Samples.” *The Southeast Asian Journal of Tropical Medicine and Public Health* 31 (1): 41–46.

Kurita, Jun, Kazuhiro Nakajima, Ikuo Hirono, and Takashi Aoki. 1998.“Polymerase Chain Reaction (PCR) Amplification of DNA of Red Sea Bream Iridovirus (RSIV).” *Fish Pathology* 33 (1): 17–23.

Kurita, Jun, and Kazuhiro Nakajima. 2012. “Megalocytiviruses.” *Viruses* 4 (4): 521–38.

Lauring, Adam S., and Emma B. Hodcroft. 2021. “Genetic Variants of SARS-CoV-2-What Do They Mean?” *JAMA: The Journal of the American Medical Association* 325 (6): 529–31.

Li, Chengjun, and Hualan Chen. 2014. “Enhancement of Influenza Virus Transmission by Gene Reassortment.” *Current Topics in Microbiology and Immunology* 385: 185–204.

Lin, Qiang, Xiaozhe Fu, Lihui Liu, Hongru Liang, Huizhi Guo, Shuwen Yin, Venkatesh Kumaresan, Zhibin Huang, and Ningqiu Li. 2017. “Application and Development of a TaqMan Real-Time PCR for Detecting Infectious Spleen and Kidney Necrosis Virus in Siniperca Chuatsi.” *Microbial Pathogenesis* 107 (June): 98–105.

Lin, Qiang, Xiaozhe Fu, Lihui Liu, Hongru Liang, Yinjie Niu, Yuanya Wen, Zhibin Huang, and Ningqiu Li. 2020. “Development and Application of a Sensitive Droplet Digital PCR (DdPCR) for the Detection of Infectious Spleen and Kidney Necrosis Virus.” *Aquaculture* 529 (December): 735697.

Liu, Yunxi, Joshua Kearney, Medhat Mahmoud, Bryce Kille, Fritz J. Sedlazeck, and Todd J. Treangen. 2022. “Rescuing Low Frequency Variants within Intra-Host Viral Populations Directly from Oxford Nanopore Sequencing Data.” *Nature Communications* 13 (1): 1321.

Lucca Maganha, Samara Rita de, Pedro Henrique Magalhães Cardoso, Simone de Carvalho Balian, Sabrina Ribeiro de Almeida-Queiroz, Andrezza Maria Fernandes, and Ricardo Luiz Moro de Sousa. 2018. “Molecular Detection and Phylogenetic Analysis of Megalocytivirus in Brazilian Ornamental Fish.” *Archives of Virology* 163 (8): 2225–31.

Lythgoe, Katrina A., Matthew Hall, Luca Ferretti, Mariateresa de Cesare, George MacIntyre-Cockett, Amy Trebes, Monique Andersson, et al. 2021. “SARS-CoV-2 within-Host Diversity and Transmission.” *Science*, March. https://doi.org/10.1126/science.abg0821.

Maar, Dianna, and Andrew M. Prantner. 2020. “Transitioning Your Assay from Quantitative PCR to Droplet Digital PCR.” 2020. https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_7320.pdf.

Machimbirike, Vimbai Irene, Mona Dverdal Jansen, Saengchan Senapin, Pongsak Khunrae, Triwit Rattanarojpong, and Ha Thanh Dong. 2019. “Viral Infections in Tilapines: More than Just Tilapia Lake Virus.” *Aquaculture* 503 (March): 508–18.

Mahardika, Ketut, Ahmad Muzaki, and Ketut Suwirya. 2009. “PATHOGENECITY OF GROUPER SLEEPY DISEASE IRIDOVIRUS (GSDIV: Megalocytivirus, FAMILY Iridoviridae) TO CORAL TROUT GROUPER Plectrophomus Leopardus.” *Indonesian Aquaculture Journal* 4 (2): 121–30.

Mohr, Peter G., Nicholas J. G. Moody, Lynette M. Williams, John Hoad, David M. Cummins, Kelly R. Davies, and Mark StJ Crane. 2015. “Molecular Confirmation of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Farmed and Imported Ornamental Fish in Australia.” *Diseases of Aquatic Organisms* 116 (2): 103–10.

Mondal, Haimanti, and John Thomas. 2022. “A Review on the Recent Advances and Application of Vaccines against Fish Pathogens in Aquaculture.” *Aquaculture International: Journal of the European Aquaculture Society* 30 (4): 1971–2000.

Moriah L. Szpara, Koenraad Van Doorslaer. 2019. “Mechanisms of DNA Virus Evolution.” *Reference Module in Life Sciences*. https://doi.org/10.1016/B978-0-12-809633-8.20993-X.

Mugimba, Kizito K., Denis K. Byarugaba, Stephen Mutoloki, Øystein Evensen, and Hetron M. Munang’andu. 2021. “Challenges and Solutions to Viral Diseases of Finfish in Marine Aquaculture.” *Pathogens* 10 (6). https://doi.org/10.3390/pathogens10060673.

Naveca, F. 2021. “SARS-CoV-2 Reinfection by the New Variant of Concern (VOC).” January 18, 2021. https://virological.org/t/sars-cov-2-reinfection-by-the-new-variant-of-concern-voc-p-1-in-amazonas-brazil/596.

Ni, Ying, Xudong Liu, Zemenu Mengistie Simeneh, Mengsu Yang, and Runsheng Li. 2023. “Benchmarking of Nanopore R10.4 and R9.4.1 Flow Cells in Single-Cell Whole-Genome Amplification and Whole-Genome Shotgun Sequencing.” *Computational and Structural Biotechnology Journal* 21 (March): 2352–64.

OIE. 2017. “Immediate Notification Report, Report Reference: , Ref OIE: 25278.” https://www.fao.org/fi/static-media/MeetingDocuments/TiLV/d10.pdf.

OIE. 2019. “World Organisation for Animal Health. Manual of Diagnostic Tests for Aquatic Animals.” Accessed October 7, 2023. https://www.woah.org/fileadmin/Home/eng/Health_standards/aahm/current/chapitre_rsbid.pdf.

Okai, Efua Konyim. 2021. “Ghana’s Fish Farmers Hit Hard by Covid-19.” February 5, 2021. https://thefishsite.com/articles/ghanas-fish-farmers-hit-hard-by-covid-19.

Pattanayak, Sabyasachi, Anirban Paul, and Pramoda Kumar Sahoo. 2020. “Detection and Genetic Analysis of Infectious Spleen and Kidney Necrosis

Virus (ISKNV) in 2 Ornamental Fish from Non-Clinical Cases: First Report from India." https://doi.org/10.1101/2020.08.12.247650.

Peck, Kayla M., and Adam S. Lauring. 2018. "Complexities of Viral Mutation Rates." *Journal of Virology* 92 (14). https://doi.org/10.1128/JVI.01031-17.

Quick, Josh, Kara Zivin, Jeff Turiczek, Jacob S. Sherkow, Matthew Lane, and Patrick Skerrett. 2020. "Real-Time Gene Sequencing Can Help Control Pandemics." 2020. https://www.statnews.com/2020/09/11/real-time-gene-sequencing-can-help-control-and-may-someday-prevent-pandemics/.

Quick, Joshua, Nathan D. Grubaugh, Steven T. Pullan, Ingra M. Claro, Andrew D. Smith, Karthik Gangavarapu, Glenn Oliveira, et al. 2017. "Multiplex PCR Method for MinION and Illumina Sequencing of Zika and Other Virus Genomes Directly from Clinical Samples." *Nature Protocols* 12 (6): 1261–76.

Quick, Joshua, Nicholas J. Loman, Sophie Duraffour, Jared T. Simpson, Ettore Severi, Lauren Cowley, Joseph Akoi Bore, et al. 2016. "Real-Time, Portable Genome Sequencing for Ebola Surveillance." *Nature* 530 (7589): 228–32.

Ragasa, Catherine, Harrison Charo-Karisa, Eugene Rurangwa, Nhuong Tran, and Kelvin Mashisia Shikuku. 2022. "Sustainable Aquaculture Development in Sub-Saharan Africa." *Nature Food* 3 (2): 92–94.

Ramírez-Paredes, José Gustavo, Richard K. Paley, William Hunt, Stephen W. Feist, David M. Stone, Terry Field, David J. Haydon, et al. 2021. "First Detection of Infectious Spleen and Kidney Necrosis Virus (ISKNV) Associated with Massive Mortalities in Farmed Tilapia in Africa." *BioRxiv*. https://doi.org/10.1101/680538.

Razak, Asrazitah Abd, Julian Ransangan, and Ahemad Sade. 2014. "First Report of Megalocytivirus (Iridoviridae) in Grouper Culture in Sabah, Malaysia" 3 (March): 896–909.

Resende, Paola Cristina, Fernando Couto Motta, Sunando Roy, Luciana Appolinario, Allison Fabri, Joilson Xavier, Kathryn Harris, et al. 2020. "SARS-CoV-2 Genomes Recovered by Long Amplicon Tiling Multiplex Approach Using Nanopore Sequencing and Applicable to Other Sequencing Platforms." *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.04.30.069039.

Shahin, Khalid, Kuttichantran Subramaniam, Alvin C. Camus, Zeinab Yazdi, Susan Yun, Samantha A. Koda, Thomas B. Waltzek, Felipe Pierezan, Ruixue Hu, and Esteban Soto. 2021. "Isolation, Identification and Characterization of a Novel Megalocytivirus from Cultured Tilapia (Oreochromis Spp.) from Southern California, USA." *Animals : An Open Access Journal from MDPI* 11 (12). https://doi.org/10.3390/ani11123524.

Shi, Cheng-Yin, Kun-Tong Jia, Bing Yang, and Jie Huang. 2010. "Complete Genome Sequence of a Megalocytivirus (Family Iridoviridae) Associated with Turbot Mortality in China." *Virology Journal* 7 (July): 159.

Shiu, Jing-Yi, Jiann-Ruey Hong, Chen-Chun Ku, and Chiu-Ming Wen. 2018. "Complete Genome Sequence and Phylogenetic Analysis of Megalocytivirus RSIV-Ku: A Natural Recombination Infectious Spleen and Kidney Necrosis Virus." *Archives of Virology* 163 (4): 1037–42.

Shlapobersky, Mark, Michael S. Sinyakov, Mark Katzenellenbogen, Ronit Sarid, Jeremy Don, and Ramy R. Avtalion. 2010. "Viral Encephalitis of Tilapia

Larvae: Primary Characterization of a Novel Herpes-like Virus." *Virology* 399 (2): 239–47.

Srivathsan, Amrita, Bilgenur Baloğlu, Wendy Wang, Wei X. Tan, Denis Bertrand, Amanda H. Q. Ng, Esther J. H. Boey, Jayce J. Y. Koh, Niranjan Nagarajan, and Rudolf Meier. 2018. "A MinION™-Based Pipeline for Fast and Cost-Effective DNA Barcoding." *Molecular Ecology Resources*, April. https://doi.org/10.1111/1755-0998.12890.

Subramaniam, K., M. Shariff, A. R. Omar, M. Hair-Bejo, and B. L. Ong. 2014. "Detection and Molecular Characterization of Infectious Spleen and Kidney Necrosis Virus from Major Ornamental Fish Breeding States in Peninsular Malaysia." *Journal of Fish Diseases* 37 (7): 609–18.

Subramaniam, Kuttichantran, Michael Gotesman, Charlie E. Smith, Natalie K. Steckler, Karen L. Kelley, Joseph M. Groff, and Thomas B. Waltzek. 2016. "Megalocytivirus Infection in Cultured Nile Tilapia Oreochromis Niloticus." *Diseases of Aquatic Organisms* 119 (3): 253–58.

Suebsing, R., P. J. Pradeep, S. Jitrakorn, S. Sirithammajak, J. Kampeera, W. A. Turner, V. Saksmerprome, B. Withyachumnarnkul, and W. Kiatpathomchai. 2016. "Detection of Natural Infection of Infectious Spleen and Kidney Necrosis Virus in Farmed Tilapia by Hydroxynapthol Blue-Loop-Mediated Isothermal Amplification Assay." *Journal of Applied Microbiology* 121 (1): 55–67.

Sukenda, Sukenda, Lila Gardenia, M. Zairin, Angela Lusiastuti, and Alimuddin Alimudin. 2020. "Identification of Giant Gourami Iridovirus (GGIV): A New Infectious Spleen and Kidney Necrosis Virus (ISKNV) from Natural Outbreak in Cultured Osphronemus Goramy." *Aquaculture International: Journal of the European Aquaculture Society*, February. https://doi.org/10.1007/s10499-020-00513-4.

Throngnumchai, Boonyalit, Sarocha Jitrakorn, Pakkakul Sangsuriya, Sasimanas Unajak, Pongsak Khunrae, Ha Thanh Dong, Vanvimon Saksmerprome, and Triwit Rattanarojpong. 2021. "Refolded Recombinant Major Capsid Protein (MCP) from Infectious Spleen and Kidney Necrosis Virus (ISKNV) Effectively Stimulates Serum Specific Antibody and Immune Related Genes Response in Nile Tilapia (Oreochromis Niloticus)." *Protein Expression and Purification* 184 (August): 105876.

Tsofack, Japhette Esther Kembou, Rachel Zamostiano, Salsabeel Watted, Asaf Berkowitz, Ezra Rosenbluth, Nischay Mishra, Thomas Briese, et al. 2017. "Detection of Tilapia Lake Virus in Clinical Samples by Culturing and Nested Reverse Transcription-PCR." *Journal of Clinical Microbiology* 55 (3): 759–67.

Tyson, John R., Phillip James, David Stoddart, Natalie Sparks, Arthur Wickenhagen, Grant Hall, Ji Hyun Choi, et al. 2020. "Improvements to the ARTIC Multiplex PCR Method for SARS-CoV-2 Genome Sequencing Using Nanopore." *BioRxiv : The Preprint Server for Biology*, September. https://doi.org/10.1101/2020.09.04.283077.

Verner-Jeffreys, D. W., T. J. Wallis, I. Cano Cejas, D. Ryder, D. J. Haydon, J. F. Domazoro, J. Dontwi, et al. 2018. "Streptococcus Agalactiae Multilocus Sequence Type 261 Is Associated with Mortalities in the Emerging Ghanaian Tilapia Industry." *Journal of Fish Diseases* 41 (1): 175–79.

Wang, Miao, and Maixin Lu. 2016. “Tilapia Polyculture: A Global Review.” *Aquaculture Research* 47 (8): 2363–74.

Wasswa, Fredrickson B., Kennedy Kassaza, Kirsten Nielsen, and Joel Bazira. 2022. “MinION Whole-Genome Sequencing in Resource-Limited Settings: Challenges and Opportunities.” *Current Clinical Microbiology Reports* 9 (4): 52–59.

Watanabe, R. A., J. L. Fryer, and J. S. Rohovec. 1988. “Molecular Filtration for Recovery of Waterborne Viruses of Fish.” *Applied and Environmental Microbiology* 54 (6): 1606–9.

Williams, Trevor, Valérie Barbosa-Solomieu, and V. Gregory Chinchar. 2005. “A Decade of Advances in Iridovirus Research.” *Advances in Virus Research* 65: 173–248.

Xu, Xiaopeng, Lichun Zhang, Shaoping Weng, Zhijian Huang, Jing Lu, Dongming Lan, Xuejun Zhong, Xiaoqiang Yu, Anlong Xu, and Jianguo He. 2008. “A Zebrafish (Danio Rerio) Model of Infectious Spleen and Kidney Necrosis Virus (ISKNV) Infection.” *Virology* 376 (1): 1–12.

Xu, Yifei, Kuiama Lewandowski, Sheila Lumley, Steven Pullan, Richard Vipond, Miles Carroll, Dona Foster, Philippa C. Matthews, Timothy Peto, and Derrick Crook. 2018. “Detection of Viral Pathogens With Multiplex Nanopore MinION Sequencing: Be Careful With Cross-Talk.” *Frontiers in Microbiology* 9 (September): 2225.

Yacout, Dalia M. M., Naglaa F. Soliman, and M. M. Yacout. 2016. “Comparative Life Cycle Assessment (LCA) of Tilapia in Two Production Systems: Semi-Intensive and Intensive.” *International Journal of Life Cycle Assessment* 21 (6): 806–19.

Yanong, Roy P. E., and Thomas B. Waltzek. 2010. “Megalocytivirus Infections in Fish, with Emphasis on Ornamental Species1.” *University of Florida, UF IFAS Extension*. http://agrilife.org/fisheries2/files/2013/09/Megalocytivirus-Infections-in-Fish-with-Emphasis-on-Ornamental-Species.pdf.

Zeng, Ruoyun, Weiqiang Pan, Yifan Lin, Jian He, Zhiyong Luo, Zhimin Li, Shaoping Weng, Jianguo He, and Changjun Guo. 2021. “Development of a Gene-Deleted Live Attenuated Candidate Vaccine against Fish Virus (ISKNV) with Low Pathogenicity and High Protection.” *IScience* 24 (7): 102750.

2. A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia *(Publication)*

Shayma Alathari[1,*], Dominique L. Chaput[1], Luis M. Bolaños[1], Andrew Joseph[2], Victoria L. N. Jackson [1], David Verner-Jeffreys[2,3], Richard Paley[2], Charles R. Tyler[1,3] and Ben Temperton[1,*]

1 Department of Biosciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK

2 Centre for Environment, Fisheries and Aquaculture Science (Cefas), The Nothe, Barrack Road, Weymouth DT4 8UB, UK

3 Sustainable Aquaculture Futures Centre, University of Exeter, Exeter EX4 4QD, UK

*Authors to whom correspondence should be addressed.

Viruses 2023, 15(4), 965; https://doi.org/10.3390/v15040965

Original submission received: 3 March 2023 / Revised: 6 April 2023 / Accepted: 10 April 2023 / **Published:** 14 April 2023 / Corrected: 29 June 2023

viruses

MDPI

Article

A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia

Shayma Alathari [1,*], Dominique L. Chaput [1], Luis M. Bolaños [1], Andrew Joseph [2], Victoria L. N. Jackson [1], David Verner-Jeffreys [2,3], Richard Paley [2], Charles R. Tyler [1,3] and Ben Temperton [1,*]

1 Department of Biosciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK
2 Centre for Environment, Fisheries and Aquaculture Science (Cefas), The Nothe, Barrack Road, Weymouth DT4 8UB, UK
3 Sustainable Aquaculture Futures Centre, University of Exeter, Exeter EX4 4QD, UK
* Correspondence: sa655@exeter.ac.uk (S.A.); b.temperton@exeter.ac.uk (B.T.)

Abstract: Tilapia farming is one of the most important sectors in aquaculture worldwide and of major importance to global food security. Infectious spleen and kidney necrosis virus (ISKNV) has been identified as an agent of high morbidity and mortality, threatening tilapia aquaculture. ISKNV was detected in Lake Volta, Ghana, in September 2018 and spread rapidly, with mortality rates between 60 and 90% and losses of more than 10 tonnes of fish per day. Understanding the spread and evolution of viral pathogens is important for control strategies. Here, we developed a tiled-PCR sequencing approach for the whole-genome sequencing of ISKNV, using long read sequencing to enable field-based, real-time genomic surveillance. This work represents the first use of tiled-PCR for whole genome recovery of viruses in aquaculture, with the longest genome target (>110 kb dsDNA) to date. Our protocol was applied to field samples collected from the ISKNV outbreaks from four intensive tilapia cage culture systems across Lake Volta, between October 2018 and May 2022. Despite the low mutation rate of dsDNA viruses, 20 single nucleotide polymorphisms accumulated during the sampling period. Droplet digital PCR identified a minimum requirement of template in a sample to recover 50% of an ISKNV genome at 275 femtograms (2410 viral templates per 5 µL sequencing reaction). Overall, tiled-PCR sequencing of ISKNV provides an informative tool to assist in disease control in aquaculture.

Keywords: *Oreochromis niloticus*; ISKNV; Artic-Network; aquaculture; long-read sequencing

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

Academic Editor: Tae-Jin Choi

Received: 3 March 2023
Revised: 6 April 2023
Accepted: 10 April 2023
Published: 14 April 2023
Corrected: 29 June 2023

1. Introduction

Nile tilapia (*Oreochromis niloticus*) is a key fish species for freshwater aquaculture, with a global production estimated at 4,525,400 tonnes [1], providing food, employment, as well as domestic and export earnings to large populations worldwide [2,3]. Tilapia production has almost doubled over the past decade [1], due to their relative ease of farming, marketability and stable prices [4]. Tilapia aquaculture provides an important source of nutrition, especially for populations that are dependent on a narrow range of staple foods. In Ghana, nearly 70,000 metric tonnes of tilapia were produced in 2018 [5], rising rapidly from only 954 tonnes in 2005 [6]. Most production in Ghana is conducted under high density stocking in floating cage systems, and is centred around Lake Volta (Figure 1), with hatcheries predominantly located besides the River Volta, below the dam to the lake [6]. Intensification of production in aquaculture is associated with risks of disease emergence and spread, as high stocking density and the number of reported viral outbreaks has increased steadily over the last few decades, resulting in catastrophic losses to fish farmers globally [7,8]. Although, the major disease agents are predominantly bacterial infections, such as Streptococcal infections [9], there is an increasing global burden of

emerging viral infections, such as tilapia lake virus (TiLV), which has been a causative agent of high cumulative mortalities estimated at (80–90%) in farmed tilapia in Israel, Ecuador and Colombia [10].

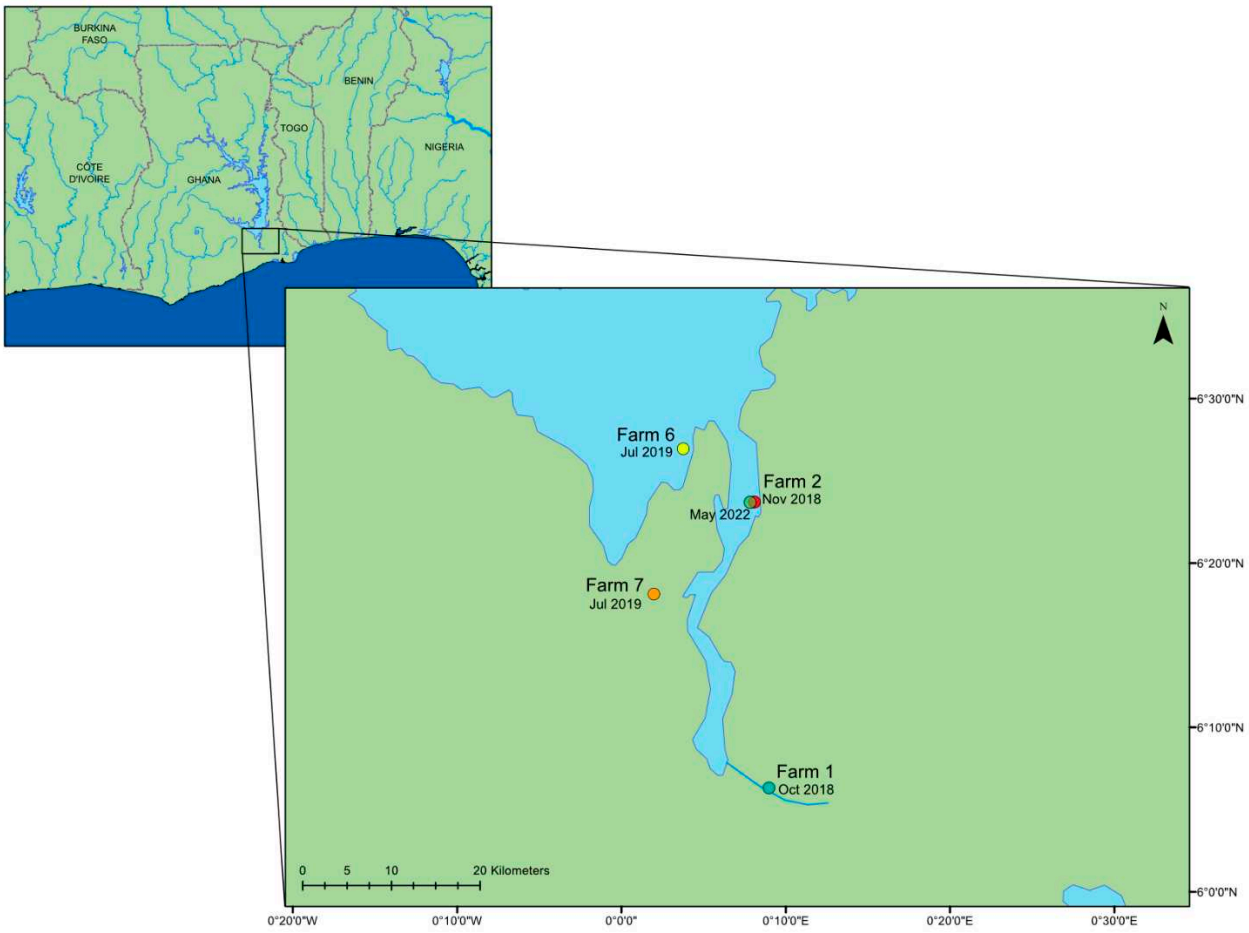


Figure 1. A map of the lower region of Lake Volta in Ghana, West Africa, showing the date and location of the farms where the outbreaks of mortality occurred; locations retrieved from [11]. This map was constructed using ArcGIS (GIS software). Version 10.0. Redlands, CA, USA: Environmental Systems Research Institute, Inc., 2010.

Infectious spleen and kidney necrosis virus (ISKNV) is a Megalocytivirus, and one of five genera within the Iridoviridae family of large, enveloped, double stranded DNA viruses [12]. ISKNV virions are icosahedral, around 150 nm in diameter and contain a single linear dsDNA molecule 111,362 bp in length, whose structure is highly methylated at cytosines in the CpG and circularly permuted during infection [13,14]. The host range of ISKNV was previously considered to be narrow: extended surveys did not detect ISKNV in 18 fish species, including tilapia [15]. However, ISKNV has been observed in mandarin fish (*Siniperca chuatsi*) [14] and large-mouth bass (*Micropterus salmoides*) [16]. In 2012, an ISKNV outbreak occurred in tilapia in the United States with a 50–75% mortality rate over a two-month period [17]. In late 2018, unusual patterns of very high mortality in the Asutsuare region of Ghana (Figure 1, Farm 1) were reported in intensive tilapia cage culture systems across Lake Volta, with ISKNV confirmed as the likely causative agent through PCR and DNA sequencing of the major capsid protein. Samples collected from the same farms had all tested negative for the virus the previous year. A week following the first report, a second farm located in the Akuse region reported similar mortalities. By the end of 2018, despite the attempts to reduce losses by increasing the production of fingerlings, or treatment with antibiotics, most tilapia farmers in Lake Volta were not able to contain these

mortalities. By mid-October, the Dodi region and the Asikuma region reported 10 tonnes of fish losses per day [11].

To mitigate the effects and spread of viral diseases in aquaculture, it is critical to achieve rapid detection of the causative agent, understand their epidemiology and disseminate the information efficiently to raise awareness [18]. Analysis of outbreaks in viral aquatic diseases in particular requires methods that offer a high level of strain discrimination [19]. Understanding the phylogeography of a viral outbreak provides vital information for containment, source identification and prevention of future outbreaks, yet current practices for epidemiological tracing focus on partial fragments of the MCP gene, do not inform changes occurring in other structural proteins or identify mutation sites on other relevant proteins that may alter vaccine development regions and/or changes in virulence. Whole genome sequencing (WGS) has revolutionised our ability to track infectious disease outbreaks by providing greater resolution of emerging diversity, allowing rapid and accurate identification of virulence factors of pathogens [20,21]. However, historically the long lead times and requirement for expensive sequencing infrastructure have limited application of WGS to understanding disease phylogeography.

The Artic-Network pipeline is an end-to-end system for generating real-time WGS epidemiological information, coupling a tiled-PCR approach to portable sequencing devices from Oxford Nanopore, enabling rapid deployment in resource-limited settings. This approach was successfully deployed to determine the phylogeography of Ebola and Zika outbreaks, as well as global surveillance of emerging variants of SARS-CoV-2 [22–26]. To date, the Artic-Network pipeline has been used for short, rapidly evolving RNA viruses (19 kbp for Ebola and 30 kbp for SARS-CoV-2). Using a similar approach to understand the phylogeography of large dsDNA viral genomes, which require many more tiled PCR products and evolve more slowly, limiting the emergence of novel variants, had not been tested. Here, we developed a protocol for WGS sequencing for real-time surveillance of ISKNV, optimising primers to recover ~96% of the ISKNV genome using the Artic-Network tiled PCR approach. When applied to samples from infected tilapia from Lake Volta, we were able to detect the accumulation of mutations across the ISKNV genome, during the sampling period. Successful field testing in Ghana showed that our method could be deployed for real-time surveillance as a field diagnostics tool. We confirm that the Artic-Network protocol can be adapted for long dsDNA viruses to provide useful phylogeographic information for managing disease outbreaks in aquaculture and beyond. A minimal viral load to recover >50% of the ISKNV genome (at least 50% of the nucleotides from ISKNV genome represented at >20-fold coverage in the sequence data) using tiled-PCR was established using droplet digital PCR (ddPCR) [27], to guide future sequencing efforts.

2. Materials and Methods

2.1. Samples and DNA Extraction

A total of 36 tissue samples from the spleen, liver and brain were collected from Nile tilapia (*Oreochromis niloticus)* from seven different fish farms in Ghana (Figure 1), during outbreaks of ISKNV. A total of 31 samples were collected by Cefas, from 10 October 2018 to 11 July 2019. An additional five samples were collected from a more recent outbreak in May 2022 (Table 1). Samples were stored in RNAlater® and shipped to Cefas Weymouth Laboratory for processing. Samples collected from farms labelled 1 and 2 were washed twice in 750 µL of sterile 1× PBS to remove the RNAlater® and homogenised using Matrix A and the FastPrep-24™ apparatus. Total nucleic acid was extracted using nanomagnetic beads (Genesig Easy DNA/RNA Extraction Kit, Primerdesign, Southampton, UK). For farms labelled 3 to 7, RNAlater® was removed and the tissue samples weighed. Depending on the weight of the tissue available, tissue samples were diluted in RLT buffer (Qiagen, Manchester, UK) at either 1:10 *w/v* or a 1:5 *w/v*, pooled and homogenised per fish using Matrix A and the FastPrep-24™ apparatus to homogenise the tissues (MP Biomedicals, Eschwege, Germany). Following homogenisation, samples were diluted further with RLT buffer for a 1:60 *w/v* homogenate and clarified by centrifugation for 10 min at 3000× *g*. Total

nucleic acid was extracted from 300 µL of the clarified sample using the EZ1&2 RNA Tissue Mini Kit without DNase (Qiagen, Manchester, UK) and eluted in 60 µL of RNAse-free water. DNA extraction for samples collected in May 2022 was performed using the DNeasy Blood and Tissue kit (Qiagen, Manchester, UK). A starting sample of 10 mg of pooled organs (kidney and spleen) were collected and dried for 5 min prior to DNA extraction using the manufacturer's protocol. Eluted nucleic acid was stored at 4 °C for one week until processing.

Table 1. Dates and regions for collection of the 36 samples from four different farms in Lake Volta, Ghana. Data for Farms 3–5 have been described previously [11].

Farm	Number of Samples	Date	Region
1	5	18 October 2018	Dodi
2	11	28 November 2018	Asikuma
2	5	20 May 2022	Asikuma
6	10	10 July 2019	Dasasi
7	5	11 July 2019	Akosombo

2.2. Design of Primers

Primers to produce 2 kbp amplicons with an overlap of 50 nt were generated with PrimalScheme (v 1.3.2) [23], using ISKNV reference sequence (accession Number: AF371960.1). A total of 62 primer pairs spanned the full ISKNV genome, and the version (v1) was designated to this set of primers (Supplementary Table S1). As an initial development and testing of our methodology for genome recovery we used viral samples collected from the 2019 outbreak in Ghana (Farm 7) and propagated in Bluegill fry BF-2 cell lines (American Type Culture Collection, ATCC CCL 91) at Cefas [11]. Template DNA was recovered (total nucleic acid kit (ThermoFisher, Heysham, UK), extracted in a Maxwell® RSC Instrument (Promega, Southampton, UK)). Each primer pair was individually tested against the template DNA, following the nCoV-2019 sequencing protocol v3: https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye?step=6 (accessed on 29 August 2020). PCR was performed with Q5 Hotstart High-Fidelity Polymerase (NEB) as follows: 98 °C for an initial heat activation for 30 s, 15 s at 98 °C for denaturation, followed by a 65 °C for annealing and extension step for 5 min for 30 cycles, and amplicons were visualised by gel electrophoresis. Four out of the 62 primer pairs failed to produce a product of appropriate size and were replaced by newer primers (v2) generated from a sequence alignment produced in Geneious Prime® 2021.1.1 from the following sequences: Accession numbers (NC_003494, MT128666, MW 273354, MW273353). Finally, the (v3) primer set contained the v1 primers with additional alternate primers for drop-out regions. All primers can be found in Supplementary Table S1.

2.3. Primer Pool Preparation

Two primer pools were generated (Pool A and Pool B), containing odd and even numbered genomic regions, respectively, at 15 nM concentration per primer. Template DNA concentration was increased from 2.5 ng to 7.5 ng (freshly diluted viral DNA in nuclease-free water (NFW)). Two pools were prepared with alternating primer sets, as previously described [23].

2.4. Failed Regions Recovered Using Neighbouring Pairs

To determine whether variation from the reference sequence was responsible for failure of four primer pairs, we generated larger amplicon products to span these regions, using neighbouring primers to generate a 6 kb product which spanned the drop-out regions (Supplementary Figure S1).

2.5. Library Preparation, Sequencing Protocols and Bioinformatic Processing

Using ISKNV cell line extracts, 2 kb amplicons were generated and pooled for sequencing using a FLO-MIN 106 (R9.4.1) (Oxford Nanopore Technologies, Oxfrord, UK) MinION flow cell. Library preparation was conducted using the Ligation Sequencing kit 1D (SQK-LSK109) (ONT) and Native Barcoding system (EXP-NBD104) (ONT), according to the manufacturer's instructions, and following the Native barcoding amplicon protocol (version NBA_9093_v109_revD_12Nov2019). Amplicons were quantified using the Qubit dsDNA broad range kit (Invitrogen, Waltham, MA, USA), and pools A & B for each sample were combined and assigned a single barcode per sample. The equimolar amounts of each barcoded sample were pooled and taken forward for the adaptor ligation step using a total volume of 60 µL of DNA. An amount of 5 µL of Adaptor Mix II (AMII), 25 µL of Ligation Buffer (LNB) and 10 µL of T4 DNA Ligase were all added to the barcoded DNA. The reaction was incubated for 10 min at room temperature, and a 0.5× AMPure XP bead clean-up was performed, followed by 2 × 250 µL of SFB (ONT) washes. The pellet was resuspended in 15 µL of Elution Buffer (EB) for 10 min at 37 °C. An amount of 15 µL of the elute was retained and ~1 µg of adaptor ligated DNA was taken forward for priming and loading onto the flow cell.

Sequencing was run for 23 h. High accuracy base calling was carried out using the Oxford Nanopore Guppy tool (v. 4.0.15). Adapter trimming was performed, and samples were demultiplexed using guppy_barcoder. Reads below 1800 bp in length and above 2200 bp were removed. Reads were mapped to the reference genome from the NCBI (NC_003494) [14] using minimap2 (v.2.17, parameters: -x map-ont) [28]. Genome coverage is visualised in Tablet [29].

2.6. Construction of the Full ISKNV Genome Infecting Tilapia in Ghana

A complete reference ISKNV genome from the Ghana outbreak was reconstructed in a three-step protocol. First, consensus genomes were constructed separately using 2 kb amplicons and 6 kb amplicons. The 6 kb amplicons were individually amplified as they failed to amplify with a multiplex PCR. The consensus sequences of 6 kb and 2 kb were aligned using LASTZ v1.02, with default parameters in Geneious Prime® 2021.2.2, revealing a gap spanning the region between primers 46 and 48. A separate amplicon library generated only using these primers was sequenced as above, and the sequences were aligned to the 2 kb/6 kb genome to close the gap. To recover the ends of the ISKNV genome, an amplicon library was generated from the last primer (62 f) and the first primer (1 r), sequenced as above and manually aligned to the constructed genome. All constructed consensus genomes used had a minimum of 20× coverage of the genome. The ISKNV consensus was annotated using Prokka version 1.13 (Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. Bioinformatics 2014, 30, 2068–2069.) [30]. Prokka was run with following parameters: –addgenes –compliant –kingdom Viruses. Predicted single nucleotide polymorphisms (SNPs) were assigned to the corresponding genes in Geneious.

2.7. Droplet Digital PCR to Determine Minimal Input for Genome Recovery of ISKNV Using the Tiled PCR Protocol

We first established the number of ISKNV templates required in a sample to recover at least 50% of the ISKNV genome at 20-fold coverage for robust error correction, using tiled PCR. Triplicates of 10-fold serial dilution from 6 ng to 6×10^{-6} ng of ISKNV from cell line extracts were used as a standard curve. Quantification of template strands in each dilution was performed in accordance with the manufacturer's instructions (Bio-Rad, Watford, UK). The reactions included 10 µL of 2× ddPCR™ Evagreen (Bio-Rad, Hercules, CA, USA), 1 µL of each forward primer (5′ CGCCTTTAACGTGGGATATATTG 3′) and reverse primer (5′ CGAGGCCACATCCAACATC 3′) (200 nM) [31], and 8 µL of DNase/RNase-free H_2O and 1 µL of DNA template. PCR amplification was performed with an initial step of 95 °C for 5 min, followed by 40 cycles of 95 °C for 30 s, 54.6 °C for 60 s and 1 cycle of 4 °C for 5 min, 1 cycle of 90 °C for 10 min, followed by 12 °C of 10 min. Microdroplets from each well were

read using a QX200 Droplet Reader (Bio-Rad, WAtford, UK). The copy number of each well was evaluated by QuantaSoft™ version 1.2 (Bio-Rad, Watford, UK).

Serial dilutions from the above were sequenced on a single MinION flow cell following library preparation using a Ligation Sequencing kit 1D (SQK-LSK109) (ONT) and Native Barcoding system (EXP-NBD196) (ONT), according to the manufacturer's instructions, and following the Native barcoding amplicon protocol: version NBA_9093_v109_revD_12Nov2019. The percentage of genome covered was estimated by at least 20-fold coverage (for consensus sequence polishing). A linear regression model (*genome recovery @ >20× coverage ~→log*$_{10}$ *(number of template strands per µL)*) was used to determine the number of viral particles to achieve at least 50% recovery of the genome. ddPCR was also employed to detect the number of ISKNV templates present in samples collected from Farms 3–5, as these samples failed to amplify by tiled PCR. A positive control of 20 ng/µL ISKNV and similar conditions were followed, as mentioned above.

2.8. Epidemiology and Phylogeographic Analysis of ISKNV

To investigate the origin and diversity of ISKNV in Ghana, we performed whole-genome alignment of 40 genomes of samples collected from Lake Volta and different ISKNV strains previously sequenced (Supplementary Table S2). Consensus genomes were aligned using Geneious Prime® 2021.1.1. Specifically, sequences were aligned using MAFFT [32], and a phylogeny was reconstructed using IQ-Tree [33,34]. The consensus sequences generated from each sample collected from Lake Volta were aligned to previously sequenced ISKNV genomes available from the GenBank NCBI. GenBank accession numbers and host species are listed in Supplementary Table S2.

2.9. Tiled PCR for MinION Sequencing of ISKNV Directly from Samples Collected from Lake Volta Outbreak

Samples from the Lake Volta ISKNV outbreak were processed using the Ligation Sequencing kit 1D (SQK-LSK109) (ONT) and Native Barcoding system (EXP-NBD104) (ONT), according to the manufacturer's instructions, and following the Native barcoding amplicon protocol, as described in detail above. Nucleic acid extracts of ISKNV, along with a negative control, were pooled for sequencing, and MinION FLO-MIN 106 (R9.4.1) and flongle (FLO-FLG001) runs were performed. Sequencing was performed for 48 h for the MinION and ~24 h for the flongle. Super high accuracy (SUP) base calling was carried out after sequencing using the Oxford Nanopore Guppy tool (v. 6.0.1). Read demultiplexing was enabled by requiring barcodes for both ends, and reads below 1800 bp in length and above 2200 bp were removed. The Artic network pipeline was used to generate the consensus sequences for each genome. The workflow can be found in Supplementary Figure S2. Augur bioinformatics toolkit (version 3.0.6) [35] (github.com/Nextstrain/augur) was used to process the genomes. Consensus genomes were aligned using MAFFT [32], and a phylogeny was reconstructed using IQ-Tree [33]. The tree was further processed using augur translate and augur clade to assign clades to nodes and to integrate phylogenetic analysis with metadata. Augur output was exported and visualised in auspice (github.com/Nextstrain/auspice) [34].

3. Results

3.1. Tiled PCR Recovers near Complete ISKNV Genome from Cell-Line Extracts

The full ISKNV genome was generated with MinION sequencing of ISKNV harvested from cell lines. Tiled-PCR products using the v1 primer scheme generated 192,317 reads, with a median read length of 1942 bp, and yielded ~75% genome recovery of the ISKNV genome, when aligned to the reference genome. All but four primer pairs were successful in generating 2 kb amplicons. Following this, primers generated from a sequence alignment (v2) of ISKNV ancestral genomes (listed in Supplementary Table S2), successfully amplified dropped regions when tested individually, and were used to replace the four failing primers,

creating a newer primer version (v3). This version was used for all subsequent Lake Volta ISKNV samples.

To reconstruct the full ISKNV genome, 6 kb amplicons spanning the full genome were recovered, and the percentage of the genome with at least 20× coverage was 83.76%. These amplicons were combined with 2 kb amplicons to recover the full genome, end regions and primer pair 47 (a 6 kbp region that did not amplify within the pool, but did amplify separately). This reconstruction generated a near complete ISKNV genome spanning 99.79% of the ISKNV reference (NC_003494) with 99.82% average nucleotide identity and 19 ambiguous bases. A total of 137 SNPs were identified when compared with the ISKNV reference genome (NC_003494)—58 of these mutations were non-synonymous. Mutations were located in the putative ankyrin repeat protein (NP_612299.1), NTPase (NP_612285.1), DNA-directed RNA polymerase II (NP_612256.1) and thymidine kinase (NP_612254.1).

3.2. ddPCR Determined Minimal Input for Genome Recovery of ISKNV Using the Tiled PCR Protocol

To evaluate the optimal concentration of ISKNV needed for genome recovery using the tiled PCR method, we measured the number of ISKNV viral templates from 6 ng to 6×10^{-6} ng. A minimum of 10 template molecules of ISKNV (to 6×10^{-5} ng) were needed to recover any of the genome with the required per-nucleotide coverage of >20-fold for accurate error correction. Genome recovery increased logarithmically from 10 template strands to 10,000 template strands, where ~75% of the genome was recovered (Figure 2). The minimum requirement to recover 50% of an ISKNV genome was 275 fg (~2410 viral templates) in 5 µL of input DNA for each sequencing reaction. Figures generated by the QuantaSoft™ version 2.1 (Bio-Rad, Hercules, CA, USA) are found in Supplementary Figure S4.

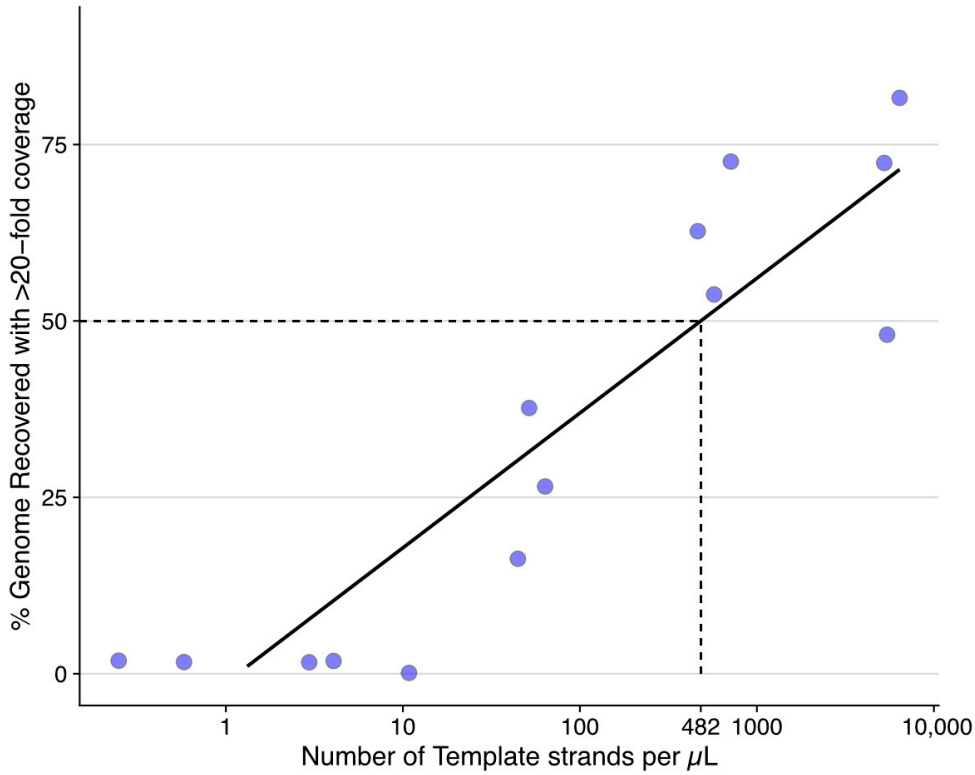


Figure 2. Successful recovery of >50% of the ISKNV genome required 482 template strands per µL (2410 viral templates per 5 µL sequencing reaction), with a minimum of 0.2 copies per µL to recover >0% of the genome with at least 20-fold coverage for error correction. Number of viral templates was measured using ddPCR from a serially diluted ISKNV template, which was subsequently sequenced and processed as described in the text.

ddPCR was also applied for the field samples (farm 3–5) that failed to amplify using the tiled PCR protocol. Detecting the number of viral templates in samples collected from Farm 3 and 4 showed very low ISKNV concentration (<1 viral template), while samples collected from Farm 5 had a high concentration of ISKNV, with up to ~5000 viral templates per ng. According to Ramierez et al., July 2019, samples from Farms 4 and 5 were recovered from recent mortality events but had no remaining observable clinical disease [11].

3.3. Epidemiology and Phylogeographic Analysis of ISKNV Is Not Solely Related to Host Species

Whole genomes from previously published reference strains from different hosts were aligned with samples collected from Ghana, aligning with MAFFT v7.450 [32], in Geneious Prime (Figure 3). ISKNV within samples collected from Ghana belonged to a separate lineage compared to samples collected from other ISKNV outbreaks. The Brazilian strain ON212400.1, although also infecting Nile tilapia, seemed distantly related to samples collected from Ghana. ISKNV from tilapia samples in Ghana were most closely related to those from an outbreak in Albino sharks (MW273353), in the United States, and were least related to samples collected from mandarin fish (*Siniperca chuatsi*) and barramundi (*Lates calcarifer*). Host species are listed in Supplementary Table S2.

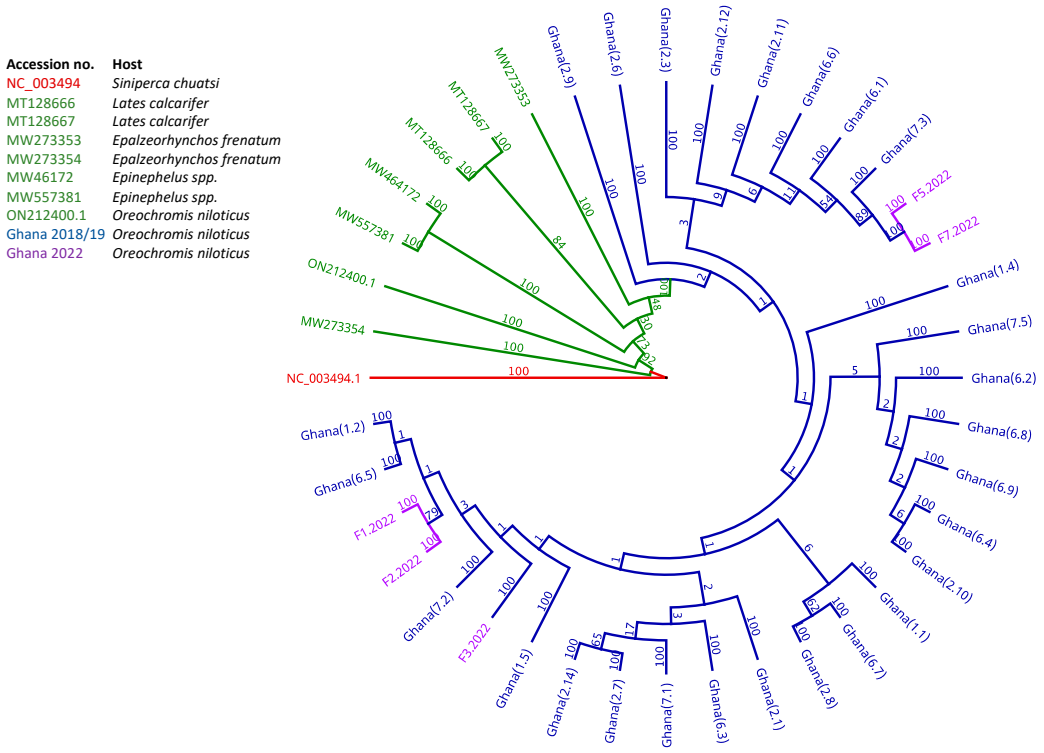


Figure 3. Phylogeny of whole ISKNV genomes of 36 samples collected from Lake Volta (2019 in blue and 2022 in purple) with whole ISKNV genomes reported in the GenBank, in green (listed in Supplementary Table S2), using MAFFT [32] with the bootstrapped branch support. The tree was rooted to the ISKNV reference genome (NC_003494), shown in red. Numbers in brackets after the Ghana samples from this study are in the format <*farm identifier*>.<*sample identifier*>.

3.4. Phylogenetic Analysis Indicates Multiple Introductions of ISKNV in Fish Samples Collected from Lake Volta

Phylogenetic analysis of ISKNV within the Ghana outbreak of 2018–2022 was performed using Augur and visualised in Auspice (Figure 4). Initial outbreaks in Lake Volta clustered into four distinct clades, and each clade had a mix of samples from different farms. The three most closely related to the reference strain were identical and from three different farms, indicating possible multiple introduction events. The highest genome recovery was obtained from sample 6.2 at ~96%, with samples 2.2 and 2.11 having the lowest coverage,

at 44% and 35% recovery, respectively (Supplementary Table S3), and the median genome recovery for all samples was 87.83% (85.61–88.63%, 95% CI, 1000 bootstraps). The consensus sequence of all the ISKNV samples obtained in this study displayed similar dropout regions in several locations of their genomes, with poorly recovered regions including: (1) a repeat region located at 23,273 bp to 23,768 bp; (2) between ORF014R and ORF018L; and (3) in the putative DNA polymerase (ORF025).

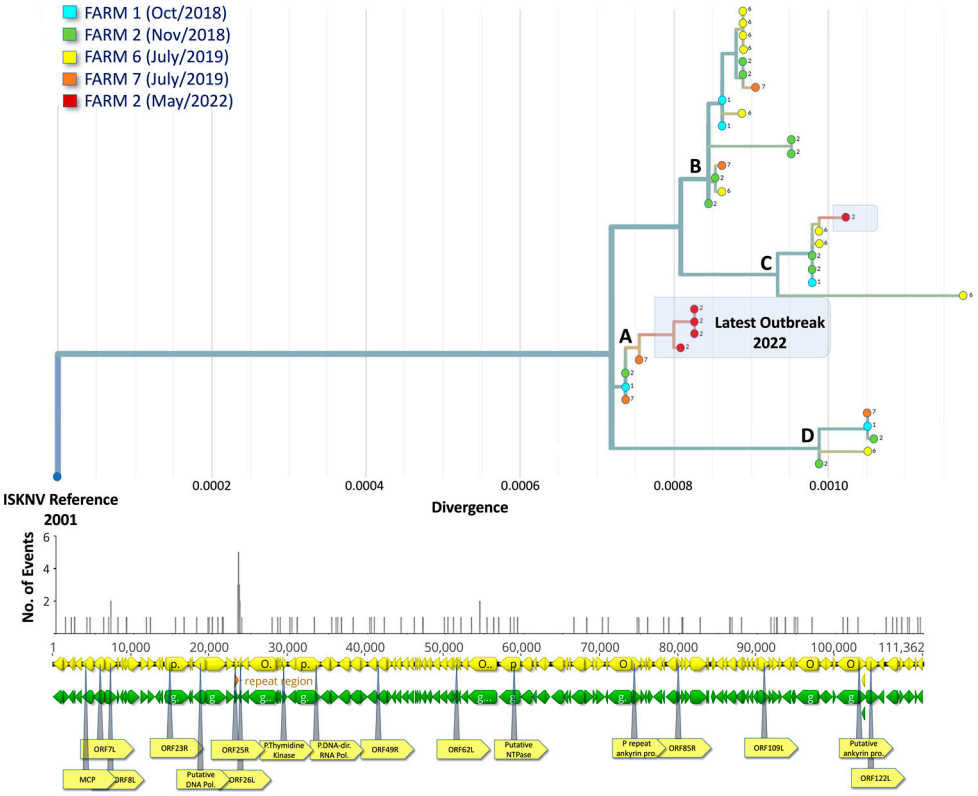


Figure 4. Phylogenetic placement of ISKNV genomes from Ghana and their associated farms. The horizontal axis indicates divergence relevant to the root of the tree. Clades are labelled A–D. The colour of the tips represents the date of sample collection; the number and location of mutational events are shown in the diversity panel below. Sequences from the latest outbreak in 2022 are highlighted.

A total of 137 polymorphisms were observed when comparing samples from the first outbreak in Ghana, in 2018, to the ancestral strain (based on SNP-calling against reference genome NC_003494.1). Of those SNPs, 20 showed variations among samples taken in 2018 and 2022 (excluding the dropout regions). Four of the five samples taken in May 2022 from the Asikuma region (Farm 2) clustered with those taken at the same location in 2018, but they have diverged independently due to a non-synonymous SNP (T3934C) within the MCP. These were highlighted in Figure 4 as "Latest outbreak". An additional mutation (C4328T) in the MCP was also unique to all Ghana samples compared to other outbreaks. A second SNP in the virulence gene ORF022L was also unique to the Ghana samples.

4. Discussion

ISKNV has caused major losses in aquaculture, with infections reported for more than 52 marine and freshwater species, and is continuously expanding to different continents [35]. To understand infectious disease dynamics in aquaculture, we have described the development and implementation of a new workflow to track viral outbreaks in fish using whole genome sequencing.

We report here the first tiled PCR that successfully generated near complete genomes of the large nucleocytoplasmic DNA virus ISKNV and its use to assess the epidemiology of

an ongoing epidemic of ISKNV in Nile tilapia in Ghana. A full-length genome was initially obtained from 2 kb and 6 kb amplicons of ISKNV from cell culture isolates from early in the epidemic. Fifty-eight non-synonymous SNPs were identified relative to reference genome sequence (NC_003494). We observed a mutation in ISKNV thymidine kinase (TK), which has previously been correlated to increased neurovirulence and mortality of the host in another dsDNA virus—Herpes Simplex Virus (HSV-1) [36], and natural mutation in the TK gene of these viruses have been associated with an increase in drug resistance [37]. Thus, observed variations may in part explain the rapid spread of ISKNV early on in the outbreak, in conjunction with the naivety of the regional tilapia Akosombo strain to the disease. When compared to ancestral samples collected from different hosts, two mutations in the ankyrin repeat protein were only seen in the strains infecting tilapia fish in Ghana. This protein has previously been shown to play a role in modulating host range and cellular immune signalling [38]. A single mutation located in the ORF022L may have increased the virulence of ISKNV in Lake Volta. Zeng et al. have defined this part of the genome as a possible virulence gene and selected it as a target gene when constructing a gene deletion vaccine for ISKNV [39].

Sequencing viral material directly from 36 samples provided insight into the relatedness of viruses collected from four different fish farms, by examining their evolution in relation to geographical spread. ISKNV samples that only showed a positive result by nested PCR failed to amplify using our tiled PCR, despite some samples showing very high ISKNV concentrations when tested by ddPCR. This could be due to fragmented DNA of ISKNV samples collected from these farms with less than 2 kb fragments, or the presence of residual fragments of non-replicating ISKNV, as these farms witnessed a past mortality event and had no clinical signs of infection during the time of sampling.

We recovered high-quality (>72% complete) genomes of ISKNV from 31 out of 36 samples collected during the ISKNV outbreak in Ghana fish farms. Phylogenetic analyses showed patterns of similar haplotypes circulating both within and between farms, indicating a shared source of infection, possibly through epidemiological links such as movement of fish or equipment, including infected live fry and fingerlings for stocking purposes, water, wild and escaped cultured tilapia as vector reservoir, and potentially other vector species. The three most closely related samples to the reference strain were identical and from three different farms, confirming multiple introductions of ISKNV and/or rapid transmission across the farms. Samples taken from farm 2 were most closely related to farm 6, despite samples being collected after seven months. Moreover, there is evidence of the rapid mutation of ISKNV in Lake Volta following the first outbreak in 2018, in comparison to the previous evolutionary rate since the first documented ISKNV outbreak in 2001. The original probable index case for the introduction of this virus into the naïve population of tilapia in Lake Volta, Ghana, from Ramirez-Paredez et al. [11], was not accessible and not sampled, nor was the virus sequenced. Given the timing of subsequent disease events on various farms on the lake, most subsequent detections, isolations, and sequence data ([11] and this publication) are likely secondary re-introductions/movements.

ISKNV genomes from Ghana appear to include two polymorphisms in the major capsid protein, a standard target for single-gene phylogenies of this disease. Within the latest samples taken in May 2022, a new non-synonymous mutation in this region was observed and has not been identified previously in any ISKNV genomes to date, suggesting continued evolution of the outbreak and requiring further study.

Twenty polymorphisms were observed across the sampling period, within the recovered regions of the genomes. This number is likely to be an underestimate because of the several dropout regions across the sequenced genome and/or error correction of the DNA genome between sampling periods. Amplicon drop-offs are common, and usually affected by viral load, sample quality, and constant viral evolution, resulting in mutations on primer binding sites [40]. Therefore, despite examining the outbreak occurring across a short period of time, and ISKNV being a dsDNA virus, divergence was seen in samples collected from the farms under investigation, shedding some light on evolutionary origins

in the phylogenetic analysis and confirming the utility of PCR-tiled approaches for viral phylogeography in large dsDNA viruses.

We did not attempt to evaluate intra-host variation of ISKNV in this study, and it is possible that the consensus sequences generated for each sample represent a flattening of true biological variability within samples. However, unlike RNA viruses, there are few reports of quasi-species within dsDNA viruses. The capacity to maintain a large viral genome without extinction through accumulated mutation is correlated to polymerase fidelity. Bacteriophages T2 and T4 are dsDNA viruses with similar genome size to ISKNV and have mutation rates of ~10^{-8} substitutions per nucleotide per replication cycle [41], approximately four orders of magnitude lower than the estimated mutation rate required to sustain a quasi-species population [42]. Therefore, it is likely that loss of intra-host variants within the consensus sequence of each sample would be minimal. In addition, the error-rate of individual Oxford Nanopore reads limits the capacity to discriminate between sequencing error and biological variability. Even with advanced methods for identifying intra-host SNVs (iSNVs), the false discovery rate of iSNVs using Oxford Nanopore data alone was ~55% in a rapidly evolving RNA virus, and is expected to be higher in viruses that evolve more slowly, such as ISKNV. In studies where quantification of iSNVs is required, replicated Illumina sequencing libraries per sample, combined with Nanopore data, is recommended for accurate quantification of iSNVs [43–47]. The increased costs and loss of field-based sequencing of this approach would need to be weighed against the likelihood and importance of detecting intra-host variability. Although not yet a day-to-day diagnostic tool for fish diseases, the method described here significantly reduces the cost of whole genome sequencing of important pathogens and makes it feasible in the field. Such information supports control strategies including the modelling of epidemiological links and potential vaccination or resistance breeding.

5. Conclusions

This work represents promising results with the potential to reveal a real-time view into the evolution and spread of ISKNV and other viral pathogens in aquaculture. This work here provides a platform from which it is feasible to replicate the Artic-Network "lab-in-a-suitcase" approach to disease tracking and management in aquaculture in remote and resource-limited locations. With appropriate training and guidance, this workflow represents a suitable framework for local authorities in lower- and middle-income countries to contain and track different viral diseases in their localities.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/v15040965/s1, Figure S1: A schematic diagram for the 6kb amplicon product targeting the primer 47 region. Figure S2: An illustration of processing tissue samples collected from infected fish. Figure S3: Gel-like image and electropherogram of the 6kb ISKNV amplicons targeting the full ISKNV genome; Figure S4. The ddPCR output data showing the number of templates generated for a serial dilution in triplicates of ISKNV. Table S1: ISKNV PCR primer scheme used by the ZiBRA project generated by the Primal Scheme software Table S2: A list of ISKNV genomes reported in the GenBank and their hosts. Table S3: Sequencing results for each sequenced sample, collected during the ISKNV outbreak from Lake Volta/Ghana.

Author Contributions: Conceptualization, B.T. and C.R.T.; methodology, S.A. and D.L.C.; software, B.T.; validation, B.T., R.P. and C.R.T.; formal analysis, S.A.; investigation, S.A.; resources, B.T., C.R.T. and A.J.; data curation, S.A. and L.M.B.; writing—original draft preparation, S.A.; writing—review and editing, B.T., C.R.T.; L.M.B. and R.P.; visualization, S.A. and V.L.N.J.; supervision, B.T. and C.R.T.; project administration, B.T., C.R.T., R.P. and D.V.-J.; funding acquisition, R.P., D.V.-J., B.T. and C.R.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by joint University of Exeter and Cefas doctoral training programme. The work was supported by funding from the UK Department of Environment, Food and Rural Affairs (Defra) under contract FX001 and Cefas Seedcorn SP003Q.

Institutional Review Board Statement: Not applicable as no experiments were performed on fish. Samples were taken jointly by staff of the CCEAAD (Cefas UK) and Ghanaian Ministry of Fisheries, Fish Health Unit as part of their statutory disease diagnosis functions. Fish were euthanised using approved schedule 1 methods.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are deposited in NCBI BioProject ID: PRJNA935699. Accession number for the ISKNV strain infecting tilapia in Lake Volta in July 2019: OQ513807.

Acknowledgments: Support from Cefas is gratefully acknowledged. The farmers involved in the investigation are thanked for their support and provision of information and samples. We thank Joshua Quick and David Studholme for their insightful comments. We thank Joanna Warwick-Dugdale, Sariqa Wagley, Karen Moore and Aaron Jeffries for their support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. FAO. 2020. Available online: http://www.fao.org/documents/card/en/c/ca9229en (accessed on 10 February 2023).
2. Machimbirike, V.I.; Jansen, M.D.; Senapin, S.; Khunrae, P.; Rattanarojpong, T.; Dong, H.T. Viral infections in tilapines: More than just tilapia lake virus. *Aquaculture* **2019**, *503*, 508–518. [CrossRef]
3. Eyngor, M.; Zamostiano, R.; Tsofack, J.E.K.; Berkowitz, A.; Bercovier, H.; Tinman, S.; Lev, M.; Hurvitz, A.; Galeotti, M.; Bacharach, E.; et al. Identification of a novel RNA virus lethal to tilapia. *J. Clin. Microbiol.* **2014**, *52*, 4137–4146. [CrossRef] [PubMed]
4. Wang, M.; Lu, M. Tilapia polyculture: A global review. *Aquac. Res.* **2016**, *47*, 2363–2374. [CrossRef]
5. Sasu, D.D. Ghana: Annual Volume of Tilapia Production. Statista. 2021. Available online: https://www.statista.com/statistics/1278956/production-volume-of-tilapia-in-ghana/ (accessed on 3 March 2023).
6. Verner-Jeffreys, D.W.; Wallis, T.J.; Cejas, I.C.; Ryder, D.; Haydon, D.J.; Domazoro, J.F.; Dontwi, J.; Field, T.R.; Adjei-Boteng, D.; Wood, G.; et al. *Streptococcus agalactiae* Multilocus sequence type 261 is associated with mortalities in the emerging Ghanaian tilapia industry. *J. Fish Dis.* **2018**, *41*, 175–179. [CrossRef] [PubMed]
7. McMurtrie, J.; Alathari, S.; Chaput, D.L.; Bass, D.; Ghambi, C.; Nagoli, J.; Delamare-Debouteville, J.; Mohan, C.V.; Cable, J.; Temperton, B.; et al. Relationships between pond water and tilapia skin microbiomes in aquaculture ponds in Malawi. *Aquaculture.* **2022**, *558*, 73836710. [CrossRef]
8. Sivasankar, P. A Review on DNA Viral Diseases of Fish. 2017. Available online: https://www.fisheriesjournal.com/archives/2017/vol5issue4/PartF/5-4-44-714.pdf (accessed on 11 February 2023).
9. Dong, H.T.; Nguyen, V.V.; Le, H.D.; Sangsuriya, P.; Jitrakorn, S.; Saksmerprome, V.; Senapin, S.; Rodkhum, C. Naturally concurrent infections of bacterial and viral pathogens in disease outbreaks in cultured Nile tilapia (*Oreochromis niloticus*) farms. *Aquaculture* **2015**, *448*, 427–435. [CrossRef]
10. Dong, H.; Siriroob, S.; Meemetta, W.; Santimanawong, W.; Gangnonngiw, W.; Pirarat, N.; Khunrae, P.; Rattanarojpong, T.; Vanichviriyakit, R.; Senapin, S. Emergence of tilapia lake virus in Thailand and an alternative semi-nested RT-PCR for detection. *Aquaculture* **2017**, *476*, 111–118. [CrossRef]
11. Ramírez-Paredez, J.G.; Paley, R.K.; Hunt, W.; Feist, S.W.; Stone, D.M.; Field, T.R.; Haydon, D.J.; Ziddah, P.A.; Nkansa, M.; Guilder, J.; et al. First detection of Infectious Spleen and kidney Necrosis Virus (ISKNV) associated with massive mortalities in farmed tilapia in Africa. *Transbound. Emerg. Dis.* **2021**, *68*, 680538. [CrossRef]
12. Mohr, P.; Moody, N.; Williams, L.; Hoad, J.; Cummins, D.; Davies, K.; Crane, M.S. Molecular confirmation of infectious spleen and kidney necrosis virus (ISKNV) in farmed and imported ornamental fish in Australia. *Dis. Aquat. Organ.* **2015**, *116*, 103–110. [CrossRef]
13. Fu, X.; Li, N.; Liu, L.; Lin, Q.; Wang, F.; Lai, Y.; Jiang, H.; Pan, H.; Shi, C.; Wu, S. Genotype and host range analysis of infectious spleen and kidney necrosis virus (ISKNV). *Virus Genes* **2011**, *42*, 97–109. [CrossRef]
14. He, J.G.; Denga, M.; Weng, S.P.; Lia, Z.; Zhou, S.Y.; Long, Q.X.; Wang, X.Z.; Chanb, S.M. Complete genome analysis of the mandarin fish infectious spleen and kidney necrosis iridovirus. *Virology* **2001**, *291*, 126–139. [CrossRef] [PubMed]
15. He, J.G.; Zeng, K.; Weng, S.P.; Chan, S.M. Experimental transmission, pathogenicity and physical–chemical properties of infectious spleen and kidney necrosis virus (ISKNV). *Aquaculture* **2002**, *204*, 11–24. [CrossRef]
16. Jung-Schroers, V.; Adamek, M.; Wohlsein, P.; Wolter, J.; Wedekind, H.; Steinhagen, D. First outbreak of an infection with infectious spleen and kidney necrosis virus (ISKNV) in ornamental fish in Germany. *Dis. Aquat. Organ.* **2016**, *119*, 239–244. [CrossRef] [PubMed]
17. Howell, M. Seven Deadly Fins: Study Sheds Light on Key Tilapia Viruses [Internet]. 2019. Available online: https://thefishsite.com/articles/seven-deadly-fins-study-sheds-light-on-key-tilapia-viruses (accessed on 5 December 2019).
18. Assefa, A.; Abunna, F. Maintenance of Fish Health in Aquaculture: Review of Epidemiological Approaches for Prevention and Control of Infectious Disease of Fish. *Vet. Med. Int.* **2018**, *2018*, 5432497. [CrossRef]
19. Avarre, J.C. Editorial: Molecular Tracing of Aquatic Viruses: Where Epidemiology Needs to Meet Genomics. *Front. Microbiol.* **2017**, *8*, 1498. [CrossRef]

20. Gilchrist, C.A.; Turner, S.D.; Riley, M.F.; Petri, W.A., Jr.; Hewlett, E.L. Whole-genome sequencing in outbreak analysis. *Clin. Microbiol. Rev.* **2015**, *28*, 541–563. [CrossRef]
21. Quick, J.; Zivin, K.; Turiczek, J.; Sherkow, J.S.; Lane, M.; Skerrett, P. Real-Time GENE sequencing Can Help Control Pandemics. 2020. Available online: https://www.statnews.com/2020/09/11/real-time-gene-sequencing-can-help-control-and-may-someday-prevent-pandemics/ (accessed on 1 February 2023).
22. Quick, J.; Loman, N.J.; Duraffour, S.; Simpson, J.T.; Severi, E.; Cowley, L.; Bore, J.A.; Koundouno, R.; Dudas, G.; Mikhail, A.; et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* **2016**, *530*, 228–232. [CrossRef]
23. Quick, J.; Grubaugh, N.D.; Pullan, S.T.; Claro, I.M.; Smith, A.D.; Gangavarapu, K.; Oliveira, G.; Robles-Sikisaka, R.; Rogers, T.F.; Beutler, A.N.; et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **2017**, *12*, 1261–1276. [CrossRef]
24. Tyson, J.R.; James, P.; Stoddart, D.; Sparks, N.; Wickenhagen, A.; Hall, G.; Choi, J.H.; Lapointe, H.; Kamelian, K.; Smith, A.D.; et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *bioRxiv* **2020**. [CrossRef]
25. Resende, P.C.; Motta, F.C.; Roy, S.; Appolinario, L.; Fabri, A.; Xavier, J.; Harris, K.; Matos, A.R.; Caetano, B.; Orgeswalska, M.; et al. SARS-CoV-2 Genomes Recovered by Long Amplicon Tiling Multiplex Approach Using Nanopore Sequencing and Applicable to Other Sequencing Platforms. Cold Spring Harbor Laboratory. 2020. Available online: https://www.biorxiv.org/content/10.1101/2020.04.30.069039v1 (accessed on 23 January 2021).
26. Du Plessis, L.; McCrone, J.T.; Zarebski, A.E.; Hill, V.; Ruis, C.; Gutierrez, B.; Raghwani, J.; Ashworth, J.; Colquhoun, R.; Connor, T.R.; et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **2021**, *371*, 708–712. [CrossRef]
27. Jiang, N.; Shen, J.; Zhou, Y.; Liu, W.; Meng, Y.; Li, Y.; Xue, M.; Xu, C.; Fan, Y. Development of a droplet digital PCR method for the sensitive detection and quantification of largemouth bass ranavirus. *J. Fish Dis.* **2023**, *46*, 91–98. [CrossRef]
28. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **2018**, *34*, 3094–3100. [CrossRef]
29. Milne, I.; Stephen, G.; Bayer, M.; Cock, P.; Pritchard, L.; Cardle, L.; Shaw, P.D.; Marshall, D. Using Tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **2013**, *14*, 193–202. [CrossRef] [PubMed]
30. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]
31. Lin, Q.; Fu, X.; Liu, L.; Liang, H.; Guo, H.; Yin, S.; Kumaresan, V.; Huang, Z.; Li, N. Application and development of a TaqMan real-time PCR for detecting infectious spleen and kidney necrosis virus in *Siniperca chuatsi*. *Microb. Pathog.* **2017**, *107*, 98–105. [CrossRef] [PubMed]
32. Katoh, K.; Misawa, K.; Kuma, K.I.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [CrossRef] [PubMed]
33. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]
34. Hadfield, J.; Megill, C.; Bell, S.M.; Huddleston, J.; Potter, B.; Callender, C.; Sagulenko, P.; Bedford, T.; Neher, R.A. Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **2018**, *34*, 4121–4123. [CrossRef]
35. Wang, Y.Q.; Lü, L.; Weng, S.P.; Huang, J.N.; Chan, S.M.; He, J.G. Molecular epidemiology and phylogenetic analysis of a marine fish infectious spleen and kidney necrosis virus-like (ISKNV-like) virus. *Arch. Virol.* **2007**, *152*, 763–773. [CrossRef]
36. Andrei, G.; Balzarini, J.; Fiten, P.; De Clercq, E.; Opdenakker, G.; Snoeck, R. Characterization of herpes simplex virus type 1 thymidine kinase mutants selected under a single round of high-dose brivudin. *J. Virol.* **2005**, *79*, 5863–5869. [CrossRef]
37. Xie, Y.; Wu, L.; Wang, M.; Cheng, A.; Yang, Q.; Wu, Y.; Jia, R.; Zhu, D.; Zhao, X.; Chen, S.; et al. Alpha-Herpesvirus Thymidine Kinase Genes Mediate Viral Virulence and Are Potential Therapeutic Targets. *Front. Microbiol.* **2019**, *10*, 941. [CrossRef] [PubMed]
38. Odon, V.; Georgana, I.; Holley, J.; Morata, J.; Maluquer de Motes, C. Novel Class of Viral Ankyrin Proteins Targeting the Host E3 Ubiquitin Ligase Cullin-2. *J. Virol.* **2018**, *92*, e01374-18. [CrossRef] [PubMed]
39. Zeng, R.; Pan, W.; Lin, Y.; He, J.; Luo, Z.; Li, Z.; Weng, S.; He, J.; Guo, C. Development of a gene-deleted live attenuated candidate vaccine against fish virus (ISKNV) with low pathogenicity and high protection. *iScience* **2021**, *24*, 102750. [CrossRef] [PubMed]
40. Lambisia, A.W.; Mohammed, K.S.; Makori, T.O.; Ndwiga, L.; Mburu, M.W.; Morobe, J.M.; Moraa, E.O.; Musyoki, J.; Murunga, N.; Mwangi, J.N.; et al. Optimization of the SARS-CoV-2 ARTIC Network V4 Primers and Whole Genome Sequencing Protocol. *Front. Med.* **2022**, *9*, 836728. [CrossRef] [PubMed]
41. Drake, J.W. The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. *Ann. N. Y. Acad. Sci.* **1999**, *870*, 100–107. [CrossRef] [PubMed]
42. Eigen, M.; Schuster, P. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften* **1977**, *64*, 541–565. [CrossRef]
43. Grubaugh, N.D.; Gangavarapu, K.; Quick, J.; Matteson, N.L.; De Jesus, J.G.; Main, B.J.; Tan, A.L.; Paul, L.M.; Brackney, D.E.; Grewal, S.; et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* **2019**, *20*, 8. [CrossRef]
44. Kerddee, P.; Dinh-Hung, N.; Dong, H.T.; Hirono, I.; Soontara, C.; Areechon, N.; Srisapo, P.; Kayansamruaj, P. Molecular evidence for homologous strains of infection spleen and kidney necrosis virus(ISKNV) genotype I infecting inland freshwater cultured Asian sea bass (*Lates calcarifer*) in Thailand. *Arch. Virol.* **2021**, *166*, 3061–3074. [CrossRef]

45. Koda, S.A.; Subramaniam, K.; Pouder, B.; Yanong Roy, P.; Frasca, J.r.S.; Popov, V.L.; Waltzek, T.B. Complete genome sequences of infectious spleen and kidney necrosis virus isolated from farmed albino rainbow sharks *Epalzeorhynchos frenatum* in the United States. *Virus Genes* **2021**, *57*, 448–452. [CrossRef]
46. Fusianto, C.; Hick, P.M.; Murwantoko, H.A.; Herlambang, A.; Whittington, R.J.; Becker, A.J. Outbreak investigation attributes Infectious spleen and kidney necrosis virus as a necessary cause of a mortality epidemic in farmed grouper (*Epinephelus* spp.) in Bali, Indonesia. *Aquac. Rep.* **2021**, *20*, 100723. [CrossRef]
47. Figueiredo, H.C.P.; Tavares, G.C.; Dorella, F.A.; Rosa, J.C.C.; Marcelino, S.A.C.; Pierezan, F.; Pereira, F.L. First report of infectious spleen and kidney necrosis virus in Nile tilapia in Brazil. *Transbound. Emerg. Dis.* **2022**, *69*, 3008–3015. [CrossRef] [PubMed]

2. 5. Supplementary Figures and Tables (*Published*)

A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia

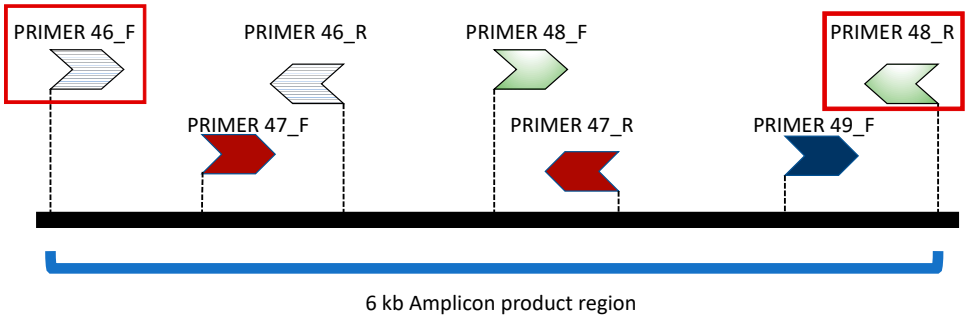


Figure S1. A schematic diagram for the 6kb amplicon product targeting the primer 47 region. Product selected using the 46f and 48r primers (circled in red) from the generated primer sets for ISKNV.

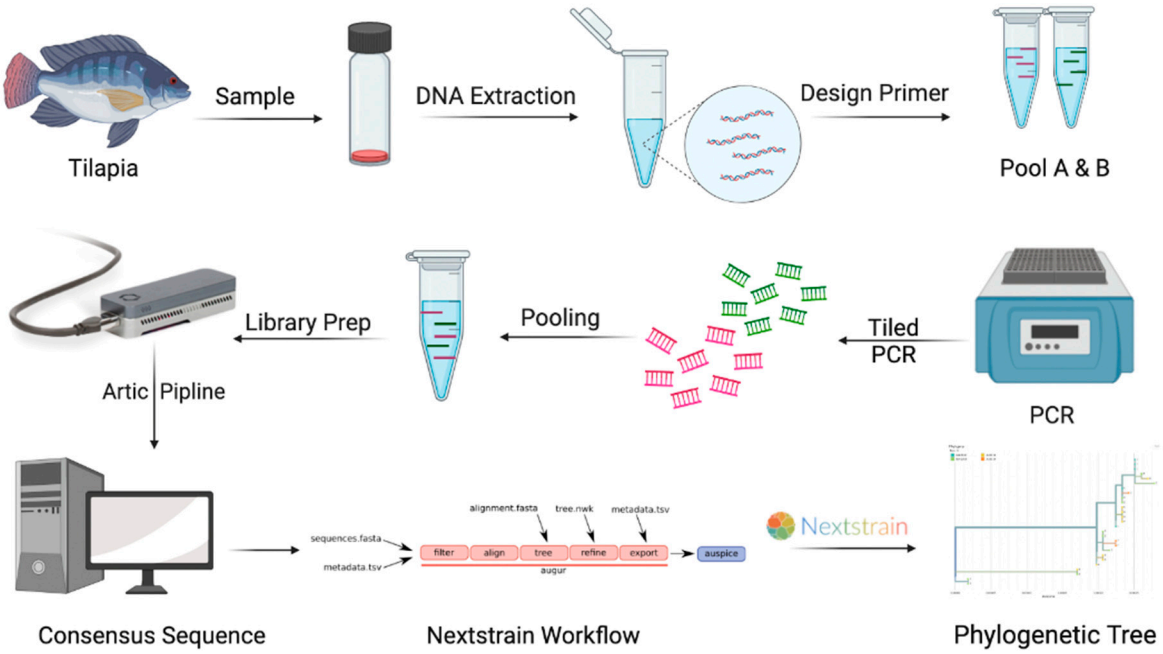
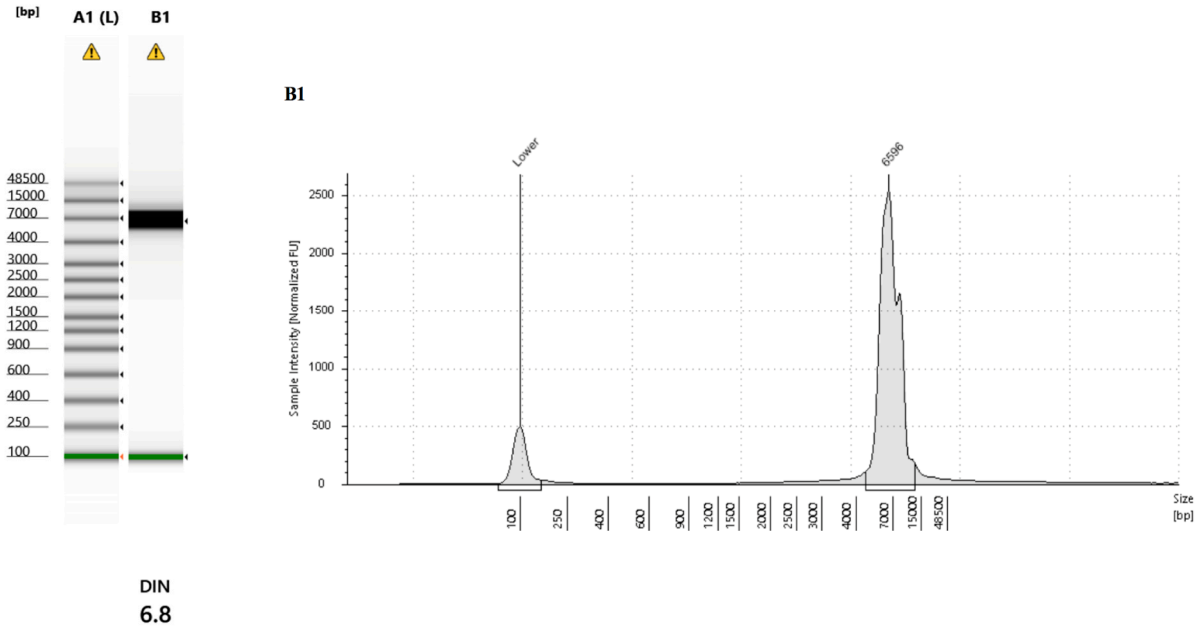


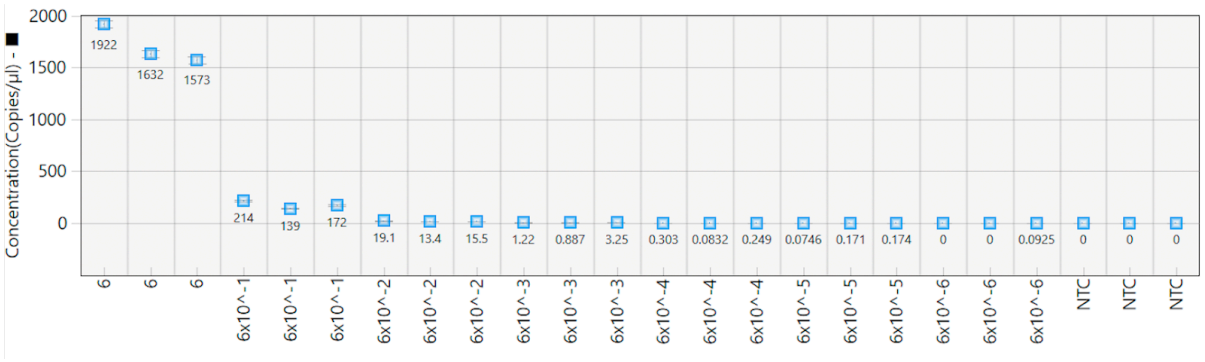
Figure S2. An illustration of processing tissue samples collected from infected fish, across the ISKNV outbreak in Lake Volta/ Ghana (2018-2019).



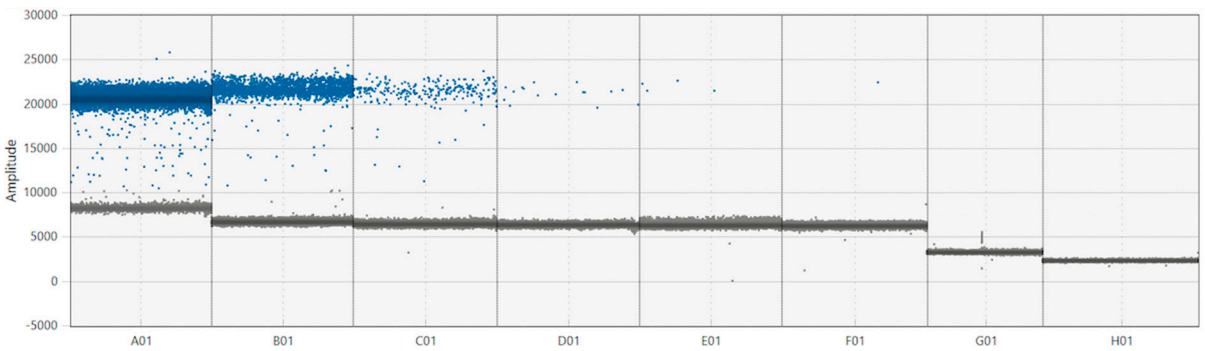
(a) **(b)**

Figure S3. Gel-like image and electropherogram of the 6kb ISKNV amplicons targeting the full ISKNV genome. (**a**). Amplicons of interest were observed as a thick band at the expected location. (**b).** The x-axis on the electropherogram represents amplicon size (bp), while the y-axis represents the measurement response of fluorescence units (FUs). Highly intact DNA was shown as a narrow peak above the highest marker peak of 6,956 bp.

(A)



(b)



(c)

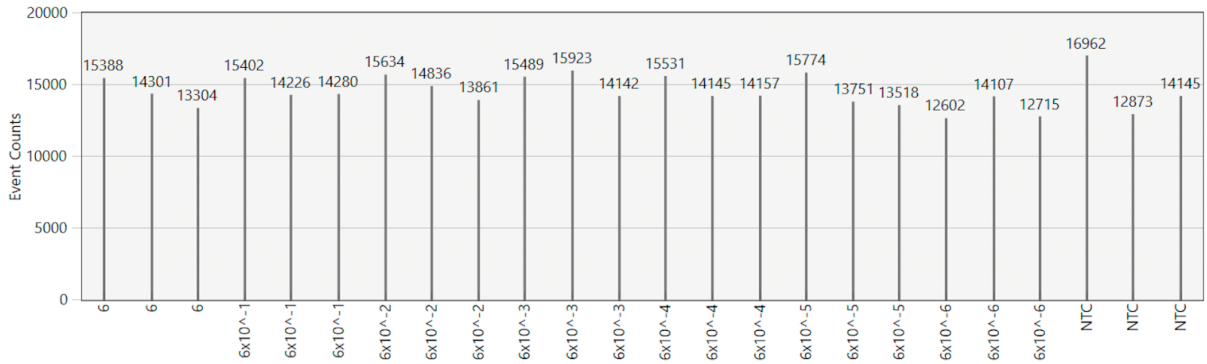


Figure S4. The ddPCR output data showing the number of templates generated for a serial dilution in triplicates of ISKNV. **a.** One-dimensional scatterplot of event number (droplets) vs. fluorescence amplitude for one set of serial dilutions, showing an ideal assay with a clear separation of positive (blue) and negative (grey) droplets **b.** Copies of ISKNV template/μL for all dilutions **c.** Event count of the total number of droplets generated.

Table S1

Different ISKNV PCR primer schemes used by the ZiBRA project generated by the Primal Scheme software [23] using: a. the reference genome (NC_003494). b. the reference genome (NC_003494), with four primer pairs replacing the dropping regions. c. the genome alignment created from the reference genome. d. the reference (NC_003494), targeting 6kb amplicons.

a. V1.

Name	Pool	Sequence	Length
ShaymaAlath_1_LEFT	ISKNV_1	AGTGTGCAGAGCATCCATGTTG	22
ShaymaAlath_1_RIGHT	ISKNV_1	ACAGTGGTTGTCCGTACCAGAA	22
ShaymaAlath_2_LEFT	ISKNV_2	AGCACATCACATATTGTAAAGGCCA	25
ShaymaAlath_2_RIGHT	ISKNV_2	CCATGGGTTCAACCAACTACGG	22
ShaymaAlath_3_LEFT	ISKNV_1	AAACTTTTGGGCCACCGTGTAG	22
ShaymaAlath_3_RIGHT	ISKNV_1	CGTCAAGCCCATGATACGCTAC	22
ShaymaAlath_4_LEFT	ISKNV_2	TACCGCTTTCACTGTGCAGGTA	22
ShaymaAlath_4_RIGHT	ISKNV_2	CCACACGTCACATAGTTCTGCC	22
ShaymaAlath_5_LEFT	ISKNV_1	GACACTGTGTTTATCTGTCGTGGA	24
ShaymaAlath_5_RIGHT	ISKNV_1	GGGTGGTGTTGCCCTAATCAAG	22
ShaymaAlath_6_LEFT	ISKNV_2	AGCTTGTCGATGTGCTGGTAAC	22
ShaymaAlath_6_RIGHT	ISKNV_2	CCCAACCTGTGCACCAAGTATC	22
ShaymaAlath_7_LEFT	ISKNV_1	ATGTCAACAGTCATAACGCCCG	22
ShaymaAlath_7_RIGHT	ISKNV_1	TTGTCAAACACCAACTTGGCCA	22
ShaymaAlath_8_LEFT	ISKNV_2	GTCGGATGCCACAGAGAAGTCT	22

ShaymaAlath_8_RIGHT	ISKNV_2	GAAACGCAGGTCACCCACTAAG	22
ShaymaAlath_9_LEFT	ISKNV_1	GGACATGTGCGCATCTAACGAC	22
ShaymaAlath_9_RIGHT	ISKNV_1	TGCAAGACACCAATCTCGATGC	22
ShaymaAlath_10_LEFT	ISKNV_2	CCGAGCATCATCATATCCAAGAACA	25
ShaymaAlath_10_RIGHT	ISKNV_2	GTTGCTGTATCCGAACACCTGG	22
ShaymaAlath_11_LEFT	ISKNV_1	GGTGATTGGCGTCACTGTATGG	22
ShaymaAlath_11_RIGHT	ISKNV_1	ATGTACCACCTCGCCATGTACA	22
ShaymaAlath_12_LEFT	ISKNV_2	AATTGACAACCAGACGACCACC	22
ShaymaAlath_12_RIGHT	ISKNV_2	TTTGCATTCTCTCTTGGGTGGC	22
ShaymaAlath_13_LEFT	ISKNV_1	CACCGTAGCAACCACTACAGTG	22
ShaymaAlath_13_RIGHT	ISKNV_1	TGTGTGTTATTAGAAATCTTCAGTCATTGT	30
ShaymaAlath_14_LEFT	ISKNV_2	AAACAAACTTCTTTGAACGCCGT	23
ShaymaAlath_14_RIGHT	ISKNV_2	TCGCCACCGACTATCTGTAGTG	22
ShaymaAlath_15_LEFT	ISKNV_1	GTTGCTGTATGGTAGCCACTGC	22
ShaymaAlath_15_RIGHT	ISKNV_1	TGTCGTGGTATCCCTTCAGCAT	22
ShaymaAlath_16_LEFT	ISKNV_2	ATCGCGGGCACTTTCCATTAAC	22
ShaymaAlath_16_RIGHT	ISKNV_2	GGATATCGGCCGGTTTGTGTTT	22
ShaymaAlath_17_LEFT	ISKNV_1	TCTGACGGCAACATAAATGGCC	22
ShaymaAlath_17_RIGHT	ISKNV_1	TTTCCATGCAAGGCGACATTGA	22
ShaymaAlath_18_LEFT	ISKNV_2	TCGTACAGGCACATCTTCCTCC	22
ShaymaAlath_18_RIGHT	ISKNV_2	TTGGTGATGGCATTGACAGAGC	22
ShaymaAlath_19_LEFT	ISKNV_1	GCAATCTGTTCAAGCAGTGGGT	22
ShaymaAlath_19_RIGHT	ISKNV_1	TTCCCCAATTTTTATGCCCCCG	22
ShaymaAlath_20_LEFT	ISKNV_2	AAAGCATCTGGTGGCCAACAAG	22
ShaymaAlath_20_RIGHT	ISKNV_2	CGTGTTTGTCATAGGCACCCTC	22
ShaymaAlath_21_LEFT	ISKNV_1	GCGCATTGTCACACAGCACATA	22
ShaymaAlath_21_RIGHT	ISKNV_1	TATCCTGTAGACAAGGACGCGG	22
ShaymaAlath_22_LEFT	ISKNV_2	ACGTGTCATGTCTATAAGCATGCG	24
ShaymaAlath_22_RIGHT	ISKNV_2	GCGCATAGCCACAGATACTGTC	22
ShaymaAlath_23_LEFT	ISKNV_1	GTACAATCAGCCGTGTGACAGC	22
ShaymaAlath_23_RIGHT	ISKNV_1	TGTCTATGTGCACGATGGGTCT	22
ShaymaAlath_24_LEFT	ISKNV_2	TTACACGTGGGTCTAGGGACAC	22

ShaymaAlath_24_RIGHT	ISKNV_2	TATTGCCAAAACCACGGACGAG	22
ShaymaAlath_25_LEFT	ISKNV_1	GACATTTGTGGTGCACGCAGAA	22
ShaymaAlath_25_RIGHT	ISKNV_1	CGTGCTATGTATACGCGCATGT	22
ShaymaAlath_26_LEFT	ISKNV_2	GTTCCAGAACAAGACACACGGT	22
ShaymaAlath_26_RIGHT	ISKNV_2	ACAACACCACTTGCTGTGTACG	22
ShaymaAlath_27_LEFT	ISKNV_1	CTGTTCTGGAGACGAGGCTACT	22
ShaymaAlath_27_RIGHT	ISKNV_1	GACGCTGACCTGAGTGCTATTG	22
ShaymaAlath_28_LEFT	ISKNV_2	AGGAACGGCATTTTAAATTGGGAAG	25
ShaymaAlath_28_RIGHT	ISKNV_2	CAGCTGCGCAACAATAGGTACA	22
ShaymaAlath_29_LEFT	ISKNV_1	CAACCTGGGCTGCTCACATATG	22
ShaymaAlath_29_RIGHT	ISKNV_1	ATGATGACAACTCTTGCGCTGG	22
ShaymaAlath_30_LEFT	ISKNV_2	CACTGTTGTTGTGCAGTAGTCAC	23
ShaymaAlath_30_RIGHT	ISKNV_2	CGATCGCTATTATGCACCCCAC	22
ShaymaAlath_31_LEFT	ISKNV_1	TACGCCTCCAGAACATCGTCAA	22
ShaymaAlath_31_RIGHT	ISKNV_1	CGGCAGGTTACATACACACCAC	22
ShaymaAlath_32_LEFT	ISKNV_2	ATCCTCAATGGGCAGCTTGGTA	22
ShaymaAlath_32_RIGHT	ISKNV_2	GCCGGTGGGATATTATGGCATG	22
ShaymaAlath_33_LEFT	ISKNV_1	CATTTGTCCATGTCCACGCACC	22
ShaymaAlath_33_RIGHT	ISKNV_1	TCAGGGTGCAAAGAAAGTGCTG	22
ShaymaAlath_34_LEFT	ISKNV_2	CATCCGGTGGCAATATGAGGTT	22
ShaymaAlath_34_RIGHT	ISKNV_2	ACATACGGCTTCAATCGCACTG	22
ShaymaAlath_35_LEFT	ISKNV_1	GTAGTCTGCCTTGTACATGCCG	22
ShaymaAlath_35_RIGHT	ISKNV_1	GTACAGGACAGCATTGGGAACG	22
ShaymaAlath_36_LEFT	ISKNV_2	CCAAGTTGGGTTATTGTAACCGTCA	25
ShaymaAlath_36_RIGHT	ISKNV_2	TGACGACAAGCTATTGGTGCAC	22
ShaymaAlath_37_LEFT	ISKNV_1	AACTGTTGTAGCTCGTTGCCTC	22
ShaymaAlath_37_RIGHT	ISKNV_1	CCATGCTTGTATCTCATCGGCC	22
ShaymaAlath_38_LEFT	ISKNV_2	AAGATGCTGTACTTTGTGGCGC	22
ShaymaAlath_38_RIGHT	ISKNV_2	ATGTGCAGCGACATCTCAATGG	22
ShaymaAlath_39_LEFT	ISKNV_1	CTGTCTGTATGTCACGAAGGGC	22
ShaymaAlath_39_RIGHT	ISKNV_1	TCCTGAAGTTCAAGCATTCGGC	22
ShaymaAlath_40_LEFT	ISKNV_2	CCAAAGTGGCGTGTGATGTCAT	22

ShaymaAlath_40_RIGHT	ISKNV_2	AGCTACCCAATGTCGTACGTCA	22
ShaymaAlath_41_LEFT	ISKNV_1	ACACGGCTTGACATACTGTTCG	22
ShaymaAlath_41_RIGHT	ISKNV_1	CCGATACCCCAAACATTACGGC	22
ShaymaAlath_42_LEFT	ISKNV_2	GGTATGGCAAGGTCACGTCATC	22
ShaymaAlath_42_RIGHT	ISKNV_2	TGGTGAAGAAGGGCCCTATGTT	22
ShaymaAlath_43_LEFT	ISKNV_1	CTTTACGCCCACATTCTCGGAG	22
ShaymaAlath_43_RIGHT	ISKNV_1	CGTATGCGTGTGTTCCAGACAA	22
ShaymaAlath_44_LEFT	ISKNV_2	CGAGACCATCACATTTGTCGGT	22
ShaymaAlath_44_RIGHT	ISKNV_2	TCGGTTCACCACGTTGAAATGG	22
ShaymaAlath_45_LEFT	ISKNV_1	AAATGCCCATATGCGCCGTTTC	22
ShaymaAlath_45_RIGHT	ISKNV_1	ACAATCTAGCTCCAGGTGCTGT	22
ShaymaAlath_46_LEFT	ISKNV_2	ATAGTGGGATCTGTGGCACCTG	22
ShaymaAlath_46_RIGHT	ISKNV_2	TCCTGGGAAAAGAGTGTCAGGG	22
ShaymaAlath_47_LEFT	ISKNV_1	CAGTACCCGCACATACTTGAGC	22
ShaymaAlath_47_RIGHT	ISKNV_1	GGCGGTCACATACAACCTTCAG	22
ShaymaAlath_48_LEFT	ISKNV_2	TCGTCAGAGTTGGGGTCGTTTA	22
ShaymaAlath_48_RIGHT	ISKNV_2	ATTATGCATTGTGCCGTGCTCA	22
ShaymaAlath_49_LEFT	ISKNV_1	GACAGCATATGCACCGATGTCG	22
ShaymaAlath_49_RIGHT	ISKNV_1	GAAACTACGTGACCAGACGCTG	22
ShaymaAlath_50_LEFT	ISKNV_2	GAGCTGTCTACATTGCGCACAA	22
ShaymaAlath_50_RIGHT	ISKNV_2	TTGAGCATGCGTATGTGGTGTC	22
ShaymaAlath_51_LEFT	ISKNV_1	AGCCGTTGGAGATCATTGTTCT	22
ShaymaAlath_51_RIGHT	ISKNV_1	GTTTCCTTCGGCCATCTCCTTG	22
ShaymaAlath_52_LEFT	ISKNV_2	TTCTTGTGTGAGGACCCCAAGA	22
ShaymaAlath_52_RIGHT	ISKNV_2	TTGGTCTCTGTGGTCATGGGTT	22
ShaymaAlath_53_LEFT	ISKNV_1	TGTGTGGTACAATAAACAGTACAAAATACA	30
ShaymaAlath_53_RIGHT	ISKNV_1	GTTCAAGGCGTACATGACAGCA	22
ShaymaAlath_54_LEFT	ISKNV_2	CATGACGTCAATTAGGTGGCCG	22
ShaymaAlath_54_RIGHT	ISKNV_2	ATGGTCGCATGCGTTACAAGAG	22
ShaymaAlath_55_LEFT	ISKNV_1	TGGCTGTTGTTGTATCATCAACTGT	25
ShaymaAlath_55_RIGHT	ISKNV_1	GTATGTCGGCATTGTCTGTGCA	22
ShaymaAlath_56_LEFT	ISKNV_2	GACACACGACACACCTGACAAC	22

ShaymaAlath_56_RIGHT	ISKNV_2	TGGGTAGTTGGTTCCCATTCGT	22
ShaymaAlath_57_LEFT	ISKNV_1	GGGCATGCTGTCCAACAACATA	22
ShaymaAlath_57_RIGHT	ISKNV_1	TGCCTGTACTCACGCCATATCA	22
ShaymaAlath_58_LEFT	ISKNV_2	GGGAGGGCTTAACGGAGATGTT	22
ShaymaAlath_58_RIGHT	ISKNV_2	GCCGACTGAGCCAATGTGATAG	22
ShaymaAlath_59_LEFT	ISKNV_1	GAGATTGGAGATGTACTGGCCG	22
ShaymaAlath_59_RIGHT	ISKNV_1	CCAGGAGAACACAAAGGATGGC	22
ShaymaAlath_60_LEFT	ISKNV_2	TTGCCTCGAGCTGGTTGACAAA	22
ShaymaAlath_60_RIGHT	ISKNV_2	CTCCATGGTGTCTGTTGATGCC	22
ShaymaAlath_61_LEFT	ISKNV_1	CATGCTGGTGTCGTAGCGTATG	22
ShaymaAlath_61_RIGHT	ISKNV_1	CGTGTGATAATGTCGGCGTCAA	22
ShaymaAlath_62_LEFT	ISKNV_2	GGACACAATGACACGACAGGTT	22
ShaymaAlath_62_RIGHT	ISKNV_2	GCTGTGATGACAAGAGACCTGC	22

b. **V2.**

ShaymaAlath_1_LEFT	ShaymaAlath_1	AGTGTGCAGAGCATCCATGTTG	22
ShaymaAlath_1_RIGHT	ShaymaAlath_1	ACAGTGGTTGTCCGTACCAGAA	22
ShaymaAlath_2_LEFT	ShaymaAlath_2	AGCACATCACATATTGTAAAGGCCA	25
ShaymaAlath_2_RIGHT	ShaymaAlath_2	CCATGGGTTCAACCAACTACGG	22
ShaymaAlath_3_LEFT	ShaymaAlath_1	AAACTTTTGGGCCACCGTGTAG	22
ShaymaAlath_3_RIGHT	ShaymaAlath_1	CGTCAAGCCCATGATACGCTAC	22
ShaymaAlath_4_LEFT	ShaymaAlath_2	TACCGCTTTCACTGTGCAGGTA	22
ShaymaAlath_4_RIGHT	ShaymaAlath_2	CCACACGTCACATAGTTCTGCC	22
ShaymaAlath_5_LEFT	ShaymaAlath_1	GACACTGTGTTTATCTGTCGTGGA	24
ShaymaAlath_5_RIGHT	ShaymaAlath_1	GGGTGGTGTTGCCCTAATCAAG	22
ShaymaAlath_6_LEFT	ShaymaAlath_2	AGCTTGTCGATGTGCTGGTAAC	22
ShaymaAlath_6_RIGHT	ShaymaAlath_2	CCCAACCTGTGCACCAAGTATC	22
ShaymaAlath_7_LEFT	ShaymaAlath_1	ATGTCAACAGTCATAACGCCCG	22
ShaymaAlath_7_RIGHT	ShaymaAlath_1	TTGTCAAACACCAACTTGGCCA	22
ShaymaAlath_8_LEFT	ShaymaAlath_2	GTCGGATGCCACAGAGAAGTCT	22
ShaymaAlath_8_RIGHT	ShaymaAlath_2	GAAACGCAGGTCACCCACTAAG	22
ShaymaAlath_9_LEFT	ShaymaAlath_1	GGACATGTGCGCATCTAACGAC	22

ShaymaAlath_9_RIGHT	ShaymaAlath_1	TGCAAGACACCAATCTCGATGC	22
ShaymaAlath_10_LEFT	ShaymaAlath_2	CCGAGCATCATCATATCCAAGAACA	25
ShaymaAlath_10_RIGHT	ShaymaAlath_2	GTTGCTGTATCCGAACACCTGG	22
ShaymaAlath_11_LEFT	ShaymaAlath_1	GGTGATTGGCGTCACTGTATGG	22
ShaymaAlath_11_RIGHT	ShaymaAlath_1	ATGTACCACCTCGCCATGTACA	22
ShaymaAlath_12_LEFT	ShaymaAlath_2	AATTGACAACCAGACGACCACC	22
ShaymaAlath_12_RIGHT	ShaymaAlath_2	TTTGCATTCTCTCTTGGGTGGC	22
ShaymaAlath_13_LEFT	ShaymaAlath_1	CACCGTAGCAACCACTACAGTG	22
ShaymaAlath_13_RIGHT	ShaymaAlath_1	TGTGTGTTATTAGAAATCTTCAGTCATTGT	30
ShaymaAlath_14_LEFT	ShaymaAlath_2	AAACAAACTTCTTTGAACGCCGT	23
ShaymaAlath_14_RIGHT	ShaymaAlath_2	TCGCCACCGACTATCTGTAGTG	22
ShaymaAlath_15_LEFT	ShaymaAlath_1	GTTGCTGTATGGTAGCCACTGC	22
ShaymaAlath_15_RIGHT	ShaymaAlath_1	TGTCGTGGTATCCCTTCAGCAT	22
ShaymaAlath_16_LEFT	ShaymaAlath_2	ATCGCGGGCACTTTCCATTAAC	22
ShaymaAlath_16_RIGHT	ShaymaAlath_2	GGATATCGGCCGGTTTGTGTTT	22
ShaymaAlath_17_LEFT	ShaymaAlath_1	TCTGACGGCAACATAAATGGCC	22
ShaymaAlath_17_RIGHT	ShaymaAlath_1	TTTCCATGCAAGGCGACATTGA	22
ShaymaAlath_18_LEFT	ShaymaAlath_2	TCGTACAGGCACATCTTCCTCC	22
ShaymaAlath_18_RIGHT	ShaymaAlath_2	TTGGTGATGGCATTGACAGAGC	22
ShaymaAlath_19_LEFT	ShaymaAlath_1	GCAATCTGTTCAAGCAGTGGGT	22
ShaymaAlath_19_RIGHT	ShaymaAlath_1	TTCCCCAATTTTTATGCCCCCG	22
ShaymaAlath_20_LEFT	ShaymaAlath_2	AAAGCATCTGGTGGCCAACAAG	22
ShaymaAlath_20_RIGHT	ShaymaAlath_2	CGTGTTTGTCATAGGCACCCTC	22
ShaymaAlath_21_LEFT	ShaymaAlath_1	GCGCATTGTCACACAGCACATA	22
ShaymaAlath_21_RIGHT	ShaymaAlath_1	TATCCTGTAGACAAGGACGCGG	22
ShaymaAlath_22_LEFT	ShaymaAlath_2	ACGTGTCATGTCTATAAGCATGCG	24
ShaymaAlath_22_RIGHT	ShaymaAlath_2	GCGCATAGCCACAGATACTGTC	22
ISKNV_Alignment_23_LEFT	ShaymaAlath_1	CTGGTCAACACATCGTCCACAT	22
ISKNV_Align-ment_23_RIGHT	ShaymaAlath_1	GGGACATGGGCATCGATGTAAA	22
ShaymaAlath_24_LEFT	ShaymaAlath_2	TTACACGTGGGTCTAGGGACAC	22
ShaymaAlath_24_RIGHT	ShaymaAlath_2	TATTGCCAAAACCACGGACGAG	22
ShaymaAlath_25_LEFT	ShaymaAlath_1	GACATTTGTGGTGCACGCAGAA	22

ShaymaAlath_25_RIGHT	ShaymaAlath_1	CGTGCTATGTATACGCGCATGT	22
ShaymaAlath_26_LEFT	ShaymaAlath_2	GTTCCAGAACAAGACACACGGT	22
ShaymaAlath_26_RIGHT	ShaymaAlath_2	ACAACACCACTTGCTGTGTACG	22
ShaymaAlath_27_LEFT	ShaymaAlath_1	CTGTTCTGGAGACGAGGCTACT	22
ShaymaAlath_27_RIGHT	ShaymaAlath_1	GACGCTGACCTGAGTGCTATTG	22
ShaymaAlath_28_LEFT	ShaymaAlath_2	AGGAACGGCATTTTAAATTGGGAAG	25
ShaymaAlath_28_RIGHT	ShaymaAlath_1	CAGCTGCGCAACAATAGGTACA	22
ShaymaAlath_29_LEFT	ShaymaAlath_2	CAACCTGGGCTGCTCACATATG	22
ShaymaAlath_29_RIGHT	ShaymaAlath_2	ATGATGACAACTCTTGCGCTGG	22
ShaymaAlath_30_LEFT	ShaymaAlath_1	CACTGTTGTTGTGCAGTAGTCAC	23
ShaymaAlath_30_RIGHT	ShaymaAlath_1	CGATCGCTATTATGCACCCCAC	22
ShaymaAlath_31_LEFT	ShaymaAlath_2	TACGCCTCCAGAACATCGTCAA	22
ShaymaAlath_31_RIGHT	ShaymaAlath_2	CGGCAGGTTACATACACACCAC	22
ShaymaAlath_32_LEFT	ShaymaAlath_1	ATCCTCAATGGGCAGCTTGGTA	22
ShaymaAlath_32_RIGHT	ShaymaAlath_1	GCCGGTGGGATATTATGGCATG	22
ShaymaAlath_33_LEFT	ShaymaAlath_2	CATTTGTCCATGTCCACGCACC	22
ShaymaAlath_33_RIGHT	ShaymaAlath_2	TCAGGGTGCAAAGAAAGTGCTG	22
ShaymaAlath_34_LEFT	ShaymaAlath_1	CATCCGGTGGCAATATGAGGTT	22
ShaymaAlath_34_RIGHT	ShaymaAlath_1	ACATACGGCTTCAATCGCACTG	22
ShaymaAlath_35_LEFT	ShaymaAlath_2	GTAGTCTGCCTTGTACATGCCG	22
ShaymaAlath_35_RIGHT	ShaymaAlath_2	GTACAGGACAGCATTGGGAACG	22
ShaymaAlath_36_LEFT	ShaymaAlath_1	CCAAGTTGGGTTATTGTAACCGTCA	25
ShaymaAlath_36_RIGHT	ShaymaAlath_1	TGACGACAAGCTATTGGTGCAC	22
ShaymaAlath_37_LEFT	ShaymaAlath_2	AACTGTTGTAGCTCGTTGCCTC	22
ShaymaAlath_37_RIGHT	ShaymaAlath_2	CCATGCTTGTATCTCATCGGCC	22
ShaymaAlath_38_LEFT	ShaymaAlath_1	AAGATGCTGTACTTTGTGGCGC	22
ShaymaAlath_38_RIGHT	ShaymaAlath_1	ATGTGCAGCGACATCTCAATGG	22
ISKNV_Alignment_39_LEFT	ShaymaAlath_2	CTGTCTGTATGTCACGAAGGGC	22
ISKNV_Align-ment_39_RIGHT	ShaymaAlath_2	TAGCGTGTCCTGAAGTTCAAGC	22
ShaymaAlath_40_LEFT	ShaymaAlath_1	CCAAAGTGGCGTGTGATGTCAT	22
ShaymaAlath_40_RIGHT	ShaymaAlath_1	AGCTACCCAATGTCGTACGTCA	22
ShaymaAlath_41_LEFT	ShaymaAlath_2	ACACGGCTTGACATACTGTTCG	22

ShaymaAlath_41_RIGHT	ShaymaAlath_2	CCGATACCCCAAACATTACGGC	22
ShaymaAlath_42_LEFT	ShaymaAlath_1	GGTATGGCAAGGTCACGTCATC	22
ShaymaAlath_42_RIGHT	ShaymaAlath_1	TGGTGAAGAAGGGCCCTATGTT	22
ShaymaAlath_43_LEFT	ShaymaAlath_2	CTTTACGCCCACATTCTCGGAG	22
ShaymaAlath_43_RIGHT	ShaymaAlath_2	CGTATGCGTGTGTTCCAGACAA	22
ShaymaAlath_44_LEFT	ShaymaAlath_1	CGAGACCATCACATTTGTCGGT	22
ShaymaAlath_44_RIGHT	ShaymaAlath_1	TCGGTTCACCACGTTGAAATGG	22
ShaymaAlath_45_LEFT	ShaymaAlath_2	AAATGCCCATATGCGCCGTTTC	22
ShaymaAlath_45_RIGHT	ShaymaAlath_2	ACAATCTAGCTCCAGGTGCTGT	22
ShaymaAlath_46_LEFT	ShaymaAlath_1	ATAGTGGGATCTGTGGCACCTG	22
ShaymaAlath_46_RIGHT	ShaymaAlath_1	TCCTGGGAAAAGAGTGTCAGGG	22
ISKNV_Alignment_47_LEFT	ShaymaAlath_2	ACAAGACTCGCAGTGTGTTTGA	22
ISKNV_Align-ment_47_RIGHT	ShaymaAlath_2	GTAGCATCGTGTCGCGCATAAA	22
ShaymaAlath_48_LEFT	ShaymaAlath_1	TCGTCAGAGTTGGGGTCGTTTA	22
ShaymaAlath_48_RIGHT	ShaymaAlath_1	ATTATGCATTGTGCCGTGCTCA	22
ShaymaAlath_49_LEFT	ShaymaAlath_2	GACAGCATATGCACCGATGTCG	22
ShaymaAlath_49_RIGHT	ShaymaAlath_2	GAAACTACGTGACCAGACGCTG	22
ISKNV_Alignment_50_LEFT	ShaymaAlath_1	CCATGTGCTTTTTGGCCACATC	22
ISKNV_Align-ment_50_RIGHT	ShaymaAlath_1	CGGTTGGGGCATAATACGGAAT	22
ShaymaAlath_51_LEFT	ShaymaAlath_2	AGCCGTTGGAGATCATTGTTCT	22
ShaymaAlath_51_RIGHT	ShaymaAlath_2	GTTTCCTTCGGCCATCTCCTTG	22
ShaymaAlath_52_LEFT	ShaymaAlath_1	TTCTTGTGTGAGGACCCCAAGA	22
ShaymaAlath_52_RIGHT	ShaymaAlath_1	TTGGTCTCTGTGGTCATGGGTT	22
ShaymaAlath_53_LEFT	ShaymaAlath_2	TGTGTGGTACAATAAACAG-TACAAAATACA	30
ShaymaAlath_53_RIGHT	ShaymaAlath_2	GTTCAAGGCGTACATGACAGCA	22
ShaymaAlath_54_LEFT	ShaymaAlath_1	CATGACGTCAATTAGGTGGCCG	22
ShaymaAlath_54_RIGHT	ShaymaAlath_1	ATGGTCGCATGCGTTACAAGAG	22
ShaymaAlath_55_LEFT	ShaymaAlath_2	TGGCTGTTGTTGTATCATCAACTGT	25
ShaymaAlath_55_RIGHT	ShaymaAlath_2	GTATGTCGGCATTGTCTGTGCA	22
ShaymaAlath_56_LEFT	ShaymaAlath_1	GACACACGACACACCTGACAAC	22
ShaymaAlath_56_RIGHT	ShaymaAlath_1	TGGGTAGTTGGTTCCCATTCGT	22

ShaymaAlath_57_LEFT	ShaymaAlath_2	GGGCATGCTGTCCAACAACATA	22
ShaymaAlath_57_RIGHT	ShaymaAlath_2	TGCCTGTACTCACGCCATATCA	22
ShaymaAlath_58_LEFT	ShaymaAlath_1	GGGAGGGCTTAACGGAGATGTT	22
ShaymaAlath_58_RIGHT	ShaymaAlath_1	GCCGACTGAGCCAATGTGATAG	22
ShaymaAlath_59_LEFT	ShaymaAlath_2	GAGATTGGAGATGTACTGGCCG	22
ShaymaAlath_59_RIGHT	ShaymaAlath_2	CCAGGAGAACACAAAGGATGGC	22
ShaymaAlath_60_LEFT	ShaymaAlath_1	TTGCCTCGAGCTGGTTGACAAA	22
ShaymaAlath_60_RIGHT	ShaymaAlath_1	CTCCATGGTGTCTGTTGATGCC	22
ShaymaAlath_61_LEFT	ShaymaAlath_2	CATGCTGGTGTCGTAGCGTATG	22
ShaymaAlath_61_RIGHT	ShaymaAlath_2	CGTGTGATAATGTCGGCGTCAA	22
ShaymaAlath_62_LEFT	ShaymaAlath_1	GGACACAATGACACGACAGGTT	22
ShaymaAlath_62_RIGHT	ShaymaAlath_1	GCTGTGATGACAAGAGACCTGC	22

c. **V3**

scheme_1_LEFT	scheme_1	GGCGCCTGTAATATAGCCATGT	22
scheme_1_RIGHT	scheme_1	ACCAACGTTGTCTTGGCGAATA	22
scheme_2_LEFT	scheme_2	TGAAAAGGCAGAGCACATCACA	22
scheme_2_RIGHT	scheme_2	TGGACATGGGCAATATCAACCC	22
scheme_3_LEFT	scheme_1	GGGGCAATCCATAGCTTACAGG	22
scheme_3_RIGHT	scheme_1	AAATTCACTGACGTCACCCTGG	22
scheme_4_LEFT	scheme_2	GCCAGGTTCAGTTTGTAGCGTA	22
scheme_4_RIGHT	scheme_2	GTTTGCCACACACGTGTACTTG	22
scheme_5_LEFT	scheme_1	GTAGGTCACATGAAAGGGCCAG	22
scheme_5_RIGHT	scheme_1	AGCAGAATGGATGGTTGCATGT	22
scheme_6_LEFT	scheme_2	TTCGATGTCTGTCGGCTTGTTG	22
scheme_6_RIGHT	scheme_2	GTTAAGGCGACCGAGCTTTACC	22
scheme_7_LEFT	scheme_1	CAAATAGCATGTTGCAGCACCG	22
scheme_7_RIGHT	scheme_1	CCGGCACTCTTTCCAGTATCTG	22
scheme_8_LEFT	scheme_2	AGACGTGCTAAAGCGACATGAC	22
scheme_8_RIGHT	scheme_2	GGTGCATTCGACATACACCACT	22
scheme_9_LEFT	scheme_1	GCTGGCAGTGCACTTGAAGATA	22
scheme_9_RIGHT	scheme_1	CTTAAGGAGATGAGCGGCCTTC	22

scheme_10_LEFT	scheme_2	CGACATTAGGGCAGTTCTGGAC	22
scheme_10_RIGHT	scheme_2	TCCACACTGTTGGCCCTATACT	22
scheme_11_LEFT	scheme_1	CCGGGGTGCTGTATGACATATG	22
scheme_11_RIGHT	scheme_1	ATTCGATTGGTAGCACTGGAGC	22
scheme_12_LEFT	scheme_2	CGGAATGGACGACTACGATGAC	22
scheme_12_RIGHT	scheme_2	TTCATTTTGGATGGCAGGTGGT	22
scheme_13_LEFT	scheme_1	TGCAAAAAGAAGACCACCACCA	22
scheme_13_RIGHT	scheme_1	TGCATGTCAGGCCTGTGGTA	20
scheme_14_LEFT	scheme_2	CGCCAAGCAGCATGGAATTTAC	22
scheme_14_RIGHT	scheme_2	TCAGCAAAGCCACATTTGAGGA	22
scheme_15_LEFT	scheme_1	CCCATTTTTAACTACGCGCCAC	22
scheme_15_RIGHT	scheme_1	CACTCAGTGGGGGTTACGATTC	22
scheme_16_LEFT	scheme_2	GCGGGCACTTTCCATTAACATC	22
scheme_16_RIGHT	scheme_2	ATAAAGAGGCTGAACAGGACCG	22
scheme_17_LEFT	scheme_1	GTACATGCCAAACACAAACCGG	22
scheme_17_RIGHT	scheme_1	GACAAAGAAGAGGGCGATTCGT	22
scheme_18_LEFT	scheme_2	CATAAGGCAGGGTCATCATGGG	22
scheme_18_RIGHT	scheme_2	CGGTGTGTATGTGTTCTTGCTG	22
scheme_19_LEFT	scheme_1	ACTGTCAAGCGTGTGATGGAAA	22
scheme_19_RIGHT	scheme_1	CGCTCATAGTGGTCATGCTCTC	22
scheme_20_LEFT	scheme_2	TAAAGCATCTGGTGGCCAACAA	22
scheme_20_RIGHT	scheme_2	TCATAGGCACCCTCCATGTCAT	22
scheme_21_LEFT	scheme_1	AAAAGGACCTCGAAGGCAAACT	22
scheme_21_RIGHT	scheme_1	AGACATGACACGTGACATGAGT	22
scheme_22_LEFT	scheme_2	ACCGCTATATGTCTGCCACAAC	22
scheme_22_RIGHT	scheme_2	ACAACATGTAGGCCAGCTGAAG	22
scheme_23_LEFT	scheme_1	CTGGTCAACACATCGTCCACAT	22
scheme_23_RIGHT	scheme_1	GGGACATGGGCATCGATGTAAA	22
scheme_24_LEFT	scheme_2	ACCGCTCACGCATAAGCTTAAT	22
scheme_24_RIGHT	scheme_2	ACAACTTCGGCTATCCCTCTCA	22
scheme_25_LEFT	scheme_1	AATGGAGTCACAGCTTCTTGCC	22
scheme_25_RIGHT	scheme_1	TTCCAACTACACTGTGCTGTGC	22

scheme_26_LEFT	scheme_2	CACGTCGAGCCGTCTTATGAAT	22
scheme_26_RIGHT	scheme_2	TTGTCGATGCTCTCCTTGACAC	22
scheme_27_LEFT	scheme_1	ACGTGACACTGATGGAGAGGAT	22
scheme_27_RIGHT	scheme_1	GACCAGTCGACATATGTGCCTC	22
scheme_28_LEFT	scheme_2	ATTCGCCAATACGTGATCTGGG	22
scheme_28_RIGHT	scheme_2	CACGAATCGCAAAGCACACAAA	22
scheme_29_LEFT	scheme_1	ATACAACAGGTCGTCAATGGGC	22
scheme_29_RIGHT	scheme_1	AGCACCAACTCGTACAACTGTC	22
scheme_30_LEFT	scheme_2	TGGTAACATCCCAGTTGTGCTG	22
scheme_30_RIGHT	scheme_2	CTAAAGGTCAGTGACGTGGAGC	22
scheme_31_LEFT	scheme_1	TCCCCAAAGTCCCTGATGGTAA	22
scheme_31_RIGHT	scheme_1	CAAGACGAGCCAACCTTCAGAC	22
scheme_32_LEFT	scheme_2	GCTTGACACGTCTCTCATAGGC	22
scheme_32_RIGHT	scheme_2	CAACGCGCACACAATGTCATAC	22
scheme_33_LEFT	scheme_1	TAGGTTACGCACGGTGTTTGAA	22
scheme_33_RIGHT	scheme_1	ATAGTGGGGTGCGTGTAGTGTA	22
scheme_34_LEFT	scheme_2	GATGCCTTGCGTGTACTTGTGA	22
scheme_34_RIGHT	scheme_2	GGCATGTACAAGGCAGACTACA	22
scheme_35_LEFT	scheme_1	TACAATCTCCGCAGCCAACAC	21
scheme_35_RIGHT	scheme_1	CGATGACATCTCTGCAAAACGC	22
scheme_36_LEFT	scheme_2	CAATGTACGTGCAAGGACTCCA	22
scheme_36_RIGHT	scheme_2	AGGCAACGAGCTACAACAGTTT	22
scheme_37_LEFT	scheme_1	GCAACACGTTCAGTAGCCTGTA	22
scheme_37_RIGHT	scheme_1	TGTGCACCATGGGATTGTAGTG	22
scheme_38_LEFT	scheme_2	TCTGTGTGCCGTCATCTTTGAG	22
scheme_38_RIGHT	scheme_2	TCTTCGAGGTATCCGGGATCTG	22
scheme_39_LEFT	scheme_1	CTGTCTGTATGTCACGAAGGGC	22
scheme_39_RIGHT	scheme_1	TAGCGTGTCCTGAAGTTCAAGC	22
scheme_40_LEFT	scheme_2	GAATGCTTGCCGAACGGATGTA	22
scheme_40_RIGHT	scheme_2	CCGCCGTACCCAGTGTATATTG	22
scheme_41_LEFT	scheme_1	TATGCTGTGCCTCAAAGGTGTC	22
scheme_41_RIGHT	scheme_1	TTGTAGATGACGTGACCTTGCC	22

scheme_42_LEFT	scheme_2	GCATGTTGTCAGGGTACTTGGT	22
scheme_42_RIGHT	scheme_2	GAGAATGTGGGCGTAAAGGTGT	22
scheme_43_LEFT	scheme_1	CCATGTCCCACCGTACATGTAC	22
scheme_43_RIGHT	scheme_1	ACCGACAAATGTGATGGTCTCG	22
scheme_44_LEFT	scheme_2	ATGTGGTTTCTGACACTCAGGC	22
scheme_44_RIGHT	scheme_2	ACTGTGCACATCTCACGTACAC	22
scheme_45_LEFT	scheme_1	CCCACGAATGTACATGAGGCAT	22
scheme_45_RIGHT	scheme_1	CAATCTAGCTCCAGGTGCTGTC	22
scheme_46_LEFT	scheme_2	TGTCTTCAAGCCGTCTTTGTGT	22
scheme_46_RIGHT	scheme_2	CAAGTATGTGCGGGTACTGTCC	22
scheme_47_LEFT	scheme_1	ACAAGACTCGCAGTGTGTTTGA	22
scheme_47_RIGHT	scheme_1	GTAGCATCGTGTCGCGCATAAA	22
scheme_48_LEFT	scheme_2	GACCTTGTCATACAGCGGATCG	22
scheme_48_RIGHT	scheme_2	TTTGGTCGCGTGGACATTATGT	22
scheme_49_LEFT	scheme_1	CGCACATCGTCAGTATTGTCCA	22
scheme_49_RIGHT	scheme_1	GAGATGGACGACACCATGGAAG	22
scheme_50_LEFT	scheme_2	CCATGTGCTTTTTGGCCACATC	22
scheme_50_RIGHT	scheme_2	CGGTTGGGGCATAATACGGAAT	22
scheme_51_LEFT	scheme_1	GCACTGCTGCTACTGAAGAAGG	22
scheme_51_RIGHT	scheme_1	GGGCGTTGATGTCGTAGTTGTA	22
scheme_52_LEFT	scheme_2	GGACAAGGATGCCATTGTGACT	22
scheme_52_RIGHT	scheme_2	AAACCTGCCGTCATACCATAGC	22
scheme_53_LEFT	scheme_1	GCTGCTAATACTGTTGCAACGT	22
scheme_53_RIGHT	scheme_1	TTATGCGTACCACCGTCATGAC	22
scheme_54_LEFT	scheme_2	AAGAGCTTGCAATGAGCCAAGA	22
scheme_54_RIGHT	scheme_2	CCTGGCCAGCATTATGTACAGT	22
scheme_55_LEFT	scheme_1	TTGTTCAGCACCTGGTGTATGG	22
scheme_55_RIGHT	scheme_1	CATGGTTCTGTCTCCGGTGTTT	22
scheme_56_LEFT	scheme_2	CACTGTCACGTGTTTGCTGAAC	22
scheme_56_RIGHT	scheme_2	TGACATTTGTCTGCTTGTGGTC	22
scheme_57_LEFT	scheme_1	GGTGATCTCCTCCTCGCTATCA	22
scheme_57_RIGHT	scheme_1	GGACAAGTACACGCTGCTAGAC	22

scheme_58_LEFT	scheme_2	TTGCATTCAGGGTGTCGCTTAA	22
scheme_58_RIGHT	scheme_2	CCTGTGTAGGAGCTCCAGTACA	22
scheme_59_LEFT	scheme_1	CAACTACTTGAGTTCCCCAGGC	22
scheme_59_RIGHT	scheme_1	GGCATCATGTTCACCTGTTGTG	22
scheme_60_LEFT	scheme_2	CCGTGTCGTCACTACAGTTTGT	22
scheme_60_RIGHT	scheme_2	GCAGTCATTTCATGCGAAGACG	22
scheme_61_LEFT	scheme_1	TGTTGGCCATGACATCGTACTG	22
scheme_61_RIGHT	scheme_1	CTTCGCATCAAACACCCTGAGA	22
scheme_62_LEFT	scheme_2	CGGGGTGTGTTGCATTTGTATG	22
scheme_62_RIGHT	scheme_2	ACACATTTATGGGCATGCGACT	22

d. **V1.6.**

Name	Pool	Sequence	Length
ISKNV6kb_1_LEFT	ISKNV_1	AGTGTGCAGAGCATCCATGTTG	22
ISKNV6kb_1_RIGHT	ISKNV_1	CGTCAAGCCCATGATACGCTAC	22
ISKNV6kb_2_LEFT	ISKNV_2	TACCGCTTTCACTGTGCAGGTA	22
ISKNV6kb_2_RIGHT	ISKNV_2	CCCAACCTGTGCACCAAGTATC	22
ISKNV6kb_3_LEFT	ISKNV_1	ATGTCAACAGTCATAACGCCCG	22
ISKNV6kb_3_RIGHT	ISKNV_1	TGCAAGACACCAATCTCGATGC	22
ISKNV6kb_4_LEFT	ISKNV_2	CCGAGCATCATCATATCCAAGAACA	25
ISKNV6kb_4_RIGHT	ISKNV_2	TTTGCATTCTCTCTTGGGTGGC	22
ISKNV6kb_5_LEFT	ISKNV_1	CACCGTAGCAACCACTACAGTG	22
ISKNV6kb_5_RIGHT	ISKNV_1	TGTCGTGGTATCCCTTCAGCAT	22
ISKNV6kb_6_LEFT	ISKNV_2	ATCGCGGGCACTTTCCATTAAC	22
ISKNV6kb_6_RIGHT	ISKNV_2	TTGGTGATGGCATTGACAGAGC	22
ISKNV6kb_7_LEFT	ISKNV_1	GCAATCTGTTCAAGCAGTGGGT	22
ISKNV6kb_7_RIGHT	ISKNV_1	TATCCTGTAGACAAGGACGCGG	22
ISKNV6kb_8_LEFT	ISKNV_2	ACGTGTCATGTCTATAAGCATGCG	24
ISKNV6kb_8_RIGHT	ISKNV_2	TATTGCCAAAACCACGGACGAG	22
ISKNV6kb_9_LEFT	ISKNV_1	GACATTTGTGGTGCACGCAGAA	22
ISKNV6kb_9_RIGHT	ISKNV_1	GACGCTGACCTGAGTGCTATTG	22

ISKNV6kb_10_LEFT	ISKNV_2	AGGAACGGCATTTTAAATTGGGAAG	25
ISKNV6kb_10_RIGHT	ISKNV_2	CGATCGCTATTATGCACCCCAC	22
ISKNV6kb_11_LEFT	ISKNV_1	TACGCCTCCAGAACATCGTCAA	22
ISKNV6kb_11_RIGHT	ISKNV_1	TCAGGGTGCAAAGAAAGTGCTG	22
ISKNV6kb_12_LEFT	ISKNV_2	CATCCGGTGGCAATATGAGGTT	22
ISKNV6kb_12_RIGHT	ISKNV_2	TGACGACAAGCTATTGGTGCAC	22
ISKNV6kb_13_LEFT	ISKNV_1	AACTGTTGTAGCTCGTTGCCTC	22
ISKNV6kb_13_RIGHT	ISKNV_1	TCCTGAAGTTCAAGCATTCGGC	22
ISKNV6kb_14_LEFT	ISKNV_2	CCAAAGTGGCGTGTGATGTCAT	22
ISKNV6kb_14_RIGHT	ISKNV_2	TGGTGAAGAAGGGCCCTATGTT	22
ISKNV6kb_15_LEFT	ISKNV_1	CTTTACGCCCACATTCTCGGAG	22
ISKNV6kb_15_RIGHT	ISKNV_1	ACAATCTAGCTCCAGGTGCTGT	22
ISKNV6kb_16_LEFT	ISKNV_2	ATAGTGGGATCTGTGGCACCTG	22
ISKNV6kb_16_RIGHT	ISKNV_2	ATTATGCATTGTGCCGTGCTCA	22
ISKNV6kb_17_LEFT	ISKNV_1	GACAGCATATGCACCGATGTCG	22
ISKNV6kb_17_RIGHT	ISKNV_1	GTTTCCTTCGGCCATCTCCTTG	22
ISKNV6kb_18_LEFT	ISKNV_2	TTCTTGTGTGAGGACCCCAAGA	22
ISKNV6kb_18_RIGHT	ISKNV_2	ATGGTCGCATGCGTTACAAGAG	22
ISKNV6kb_19_LEFT	ISKNV_1	TGGCTGTTGTTGTATCATCAACTGT	25
ISKNV6kb_19_RIGHT	ISKNV_1	TGCCTGTACTCACGCCATATCA	22
ISKNV6kb_20_LEFT	ISKNV_2	GGGAGGGCTTAACGGAGATGTT	22
ISKNV6kb_20_RIGHT	ISKNV_2	CTCCATGGTGTCTGTTGATGCC	22
ISKNV6kb_21_LEFT	ISKNV_2	TTGCCTCGAGCTGGTTGACAAA	22
ISKNV6kb_21_RIGHT	ISKNV_2	GCTGTGATGACAAGAGACCTGC	22

Table S2. A list of ISKNV genomes reported in the GenBank and their hosts.

Accession no.	Host	Country	Date	Reference
NC_003494	*Siniperca chuatsi*	China	2001	He et al. 2001 [14]
MT128666	*Lates calcarifer*	Thailand	2018	Kerddee et al. 2021 [44]
MT128667	*Lates calcarifer*	Thailand	2018	Kerddee et al. 2021 [44]
MW273353	*Epalzeorhynchos frenatum*	USA	2018-2019	Koda et al. 2021 [45]
MW273354	*Epalzeorhynchos frenatum*	USA	2018-2019	Koda et al. 2021 [45]

MW46172	*Epinephelus spp.*	Indonesia	2016	Fusianto et al. 2021 [46]
MW557381	*Epinephelus spp.*	Indonesia	2016	Fusianto et al. 2021 [46]
ON212400.1	*Oreochromis nilocticus*	Brazil	2020	Figueiredo et al 2022 [47]

Table S3. Sequencing results for each sequenced sample, collected during the ISKNV outbreak from Lake Volta/ Ghana**.** We show location and date of sampling, the number of sequenced reads, and the percentage of the coverage of MinION reads.

Sample	Farm	Date	No. of reads	% Coverage x20
Farm1.1	1	18.10.2018	23310	89.72
Farm1.2	1	18.10.2018	114824	46.9117
Farm1.3	1	18.10.2018	76204	92.95
Farm1.4	1	18.10.2018	144891	84.99
Farm1.5	1	18.10.2018	146888	86.78
Farm 2.1	2	28.11.2019	24763	84.99
Farm 2.2	2	28.11.2018	11382	44
Farm 2.3	2	28.11.2018	241475	75.38
Farm 2.6	2	28.11.2018	174175	49.96
Farm 2.7	2	28.11.2018	67367	86.54
Farm 2.8	2	28.11.2018	127231	89.59
Farm 2.9	2	28.11.2018	66533	74.63
Farm 2.10	2	28.11.2018	47804	91.39
Farm 2.11	2	28.11.2018	5473	35.22
Farm 2.12	2	28.11.2018	18600	51.18
Farm 2.13	2	28.11.2018	52392	86.22
Farm 2.14	2	28.11.2018	86402	85.04
Farm 6.1	6	10.07.2019	101587	91.72
Farm 6.2	6	10.07.2019	100300	95.91
Farm 6.3	6	10.07.2019	41219	88.53
Farm 6.4	6	10.07.2019	73099	94.42
Farm6.5	6	10.07.2019	538869	94.50
Farm6.6	6	10.07.2019	35121	84.78
Farm6.7	6	10.07.2019	271267	92.71
Farm6.8	6	10.07.2019	435008	89.85
Farm6.9	6	10.07.2019	185147	72.23

Farm6.10	6	10.07.2019	60290	91.39
Farm7.4	7	10.07.2019	44084	89.4
Farm7.1	7	11.07.2019	22914	87.83
Farm7.2	7	11.07.2019	148211	88.23
Farm7.3	7	11.07.2019	3113	44.72
Farm7.5	7	11.07.2019	338190	82.27
Farm 2.1/22	2	20.05.2022	114551	91.16
Farm 2.2/22	2	20.05.2022	177855	89.75
Farm 2.3/22	2	20.05.2022	98879	88.29
Farm 2.4/22	2	20.05.2022	94272	63.48
Farm2.5/22	2	20.05.2022	16934	88.73
Negative cont.	-	20.05.2022	8	0

3. In field use of water samples for genomic surveillance of ISKNV infecting tilapia fish in Lake Volta, Ghana

Shayma Alathari[1,*], Andrew Joseph[2], Luis M. Bolaños[1], David Studholme[1], Aaron Jeffries[1], Patrick Appenteng[4], Kwaku Duodu[4], Eric Sawyerr[4], Richard Paley[2], Charles R. Tyler[1, 3], and Ben Temperton[1,*]

[1] Biosciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK.
[2]Centre for Environment, Fisheries and Aquaculture Science (Cefas), The Nothe, Barrack Road, Weymouth DT4 8UB, UK.
[3]Sustainable Aquaculture Futures Centre, University of Exeter, Exeter EX4 4QD, UK.
[4]Fisheries Commission, Ministry of Fisheries and Aquaculture Development, Accra, Ghana.

*Corresponding Authors:
Shayma Alathari, Ben Temperton

Department of Biosciences, University of Exeter, Stocker Road, Exeter EX4 4QD, UK.
Email address: sa655@exeter.ac.uk, b.temperton@exeter.ac.uk

Submitted and currently under review in PeerJ.
Submitted to BioRxiv preprint: https://doi.org/10.1101/2023.12.02.569710.

Abstract

Viral outbreaks are a constant threat to aquaculture, limiting production for better global food security. A lack of diagnostic testing and monitoring in resource-limited areas hinders the capacity to respond rapidly to disease outbreaks and to prevent viral pathogens becoming endemic in fisheries productive waters. Recent developments in diagnostic testing for emerging viruses, however, offers a solution for rapid *in situ* monitoring of viral outbreaks. Genomic epidemiology has furthermore proven highly effective in detecting viral mutations involved in pathogenesis and assisting in resolving chains of transmission.

Here, we demonstrate the application of an in-field epidemiological tool kit to track viral outbreaks in aquaculture on farms with reduced access to diagnostic labs, and with non-destructive sampling. Inspired by the “lab in a suitcase” approach used for genomic surveillance of human viral pathogens and wastewater monitoring of COVID19, we evaluated the feasibility of real-time genome sequencing surveillance of the fish pathogen, Infectious spleen and kidney necrosis virus (ISKNV) in Lake Volta. Viral fractions from water samples collected from cages holding Nile tilapia (*Oreochromis niloticus*) with suspected ongoing ISKNV infections were concentrated and used as a template for whole genome sequencing, using a previously developed tiled PCR method for ISKNV. Mutations in ISKNV in samples collected from the water surrounding the cages matched those collected from infected caged fish, illustrating that water samples can be used for detecting predominant ISKNV variants in an ongoing outbreak. This approach allows for the detection of ISKNV and tracking of the dynamics of variant frequencies, and may thus assist in guiding control measures for the rapid isolation and quarantine of infected farms and facilities.

3.1 Introduction

Today, 811 million people globally suffer from hunger and 3 billion cannot afford a healthy diet. The United Nations has listed Zero Hunger as one of the global sustainable development goals and to end extreme poverty by 2030 (Boykin et al. 2018). As populations continue to grow, aquaculture is expected to play an increasingly important role in improving food security, and most notably in Low- and Middle-Income Countries (Cai 2022). New strategies have been developed such as “Blue Transformation" to enhance the role of aquaculture in food production, by providing the legal, policy and technical frameworks required to sustain growth and innovation systems to do so (FAO 2022). Despite significant increases in aquaculture output in the last few decades, all forms of aquaculture are limited by infectious diseases (FAO 2019). A study by You and Hedgcock suggested that the boom-and-bust production dynamics in aquaculture – periods of rapid growth spanning several years, followed by collapse - cause significant losses of production, with disease identified as a major cause of collapse, as well as economic factors, water quality and inbreeding that reduces fitness and increases susceptibility to disease (You and Hedgcock 2017). Fish disease is usually triggered by poor water and poor farm management and inadequate biosecurity practices (Ragasa et al. 2022). Implementation of biosecurity measures in resource-limited countries is, in part, challenging due to a lack of suitable real-time and/or effective diagnostics.

In Ghana, ISKNV, a Megalocytivirus, has become endemic in tilapia in Lake Volta, following a series of outbreaks in 2018 and this has significantly affected local farmers and their livelihoods (Ramírez-Paredes et al. 2021). According to these farmers, attempts to minimise the effects of the impact of outbreaks through heat shocking fish, to reduce the effectiveness of the virus, or increasing fingerling production, have not helped to improve total production. Genome sequencing provides an unparalleled ability to track infectious disease outbreaks, from the initial detection to understanding factors that contribute to the geographical spread. Indeed, it is emerging as a critical tool in real-time responses to these outbreaks, by providing insights into how viruses transmit, spread and evolve (Quick et al. 2017; Gardy, Loman, and Rambaut 2015). Accurate reconstruction of strain-resolved genomes is useful to monitor the

outbreak of viruses, to track their evolutionary history and develop effective vaccines and drugs, as well as detect the emergence of novel variants that may impact the course of an epidemic (Luo, Kang, and Schönhuth 2022; Child et al. 2023).

In aquaculture, monitoring large numbers of infections through tissue sampling poses challenges in large-scale outbreaks, particularly in resource-limited settings, as it is time consuming and requires well practised personnel. In human health, analyses of wastewater samples have been used to understand mutations and infection dynamics, as well as an early indicator of infection (Dharmadhikari et al. 2022). This method was used to monitor the ongoing evolution of SARS-CoV-2 during the pandemic, and the water-based epidemiological programmes has provided insights into its prevalence and diversity in different communities and detecting the emergence and spread of variants (Brunner et al. 2023). In the context of fish pathogens, water-based epidemiology provides a non-invasive routine method to early detection of viruses in asymptomatic fish and ongoing infections, reducing the sacrifice of fish for testing.

In this study, we tested the utility of an in-field water sampling method for whole genome sequencing of ISKNV, using a tiled PCR method that we developed previously (Alathari et al. 2023), as a potential alternative to destructive tissue sampling for genomic surveillance of a disease outbreak in Lake Volta, Ghana. We show water samples collected in the immediate vicinity of the cage fish showed similar variants to infected tissue samples in tilapia at that site, providing confidence in-field water sampling method for genomic surveillance.

3.2 Materials & Methods

3.2.1 Samples

In an ongoing outbreak of ISKNV, water and tilapia tissue samples were collected from six geographically distinct Nile tilapia farms (*Oreochromis niloticus)* situated on Lake Volta, Ghana, in January 2023, (Figure 1 & Table 1). Water samples (250-500 mL) were collected from high density cage-based farms on the lake and processed by sequential filtration through a 0.45 µm pore (PES filters), 0.22 µm pore (Merck, Millipore (Durapore PVDF Membrane)), and finally concentrating viral particles on 0.1 µm pore filters (Merck, Millipore (Durapore PVDF Membrane)), housed within Luer-lock syringe-compatible casings. An Erwin® quick-grip minibar clamp (6") was used to facilitate the pumping of the water, with a custom 3D-printed adaptor for the syringe (Supplementary Figure 1).

Viruses on 0.1 µm filters were preserved *in* situ by addition of RNALater®, filling the filter housing, and the inlet and outlet of the filters were sealed with Parafilm®. Filters were transferred to the University of Exeter for further processing. For matching tissue samples, a total of 12 fish were selected from each of the six farms, typically four fish from each of three cages across various fish life stages. Fish were humanely euthanized with a lethal overdose of tricaine methanesulfonate 1,000 mg/g (Pharmaq, Hampshire, UK), and the spleen, liver and kidney were collected on site. Tissue samples were either processed in the field, or were preserved in RNALater®, and taken for further processing at the University of Exeter. Fish size, life stages, and any observed clinical signs are detailed in Supplementary Table 1. For the samples from farm (F), one cage (number three) had been heat-shocked by the farmers as part of their routine treatment before sampling (timeframe unknown).

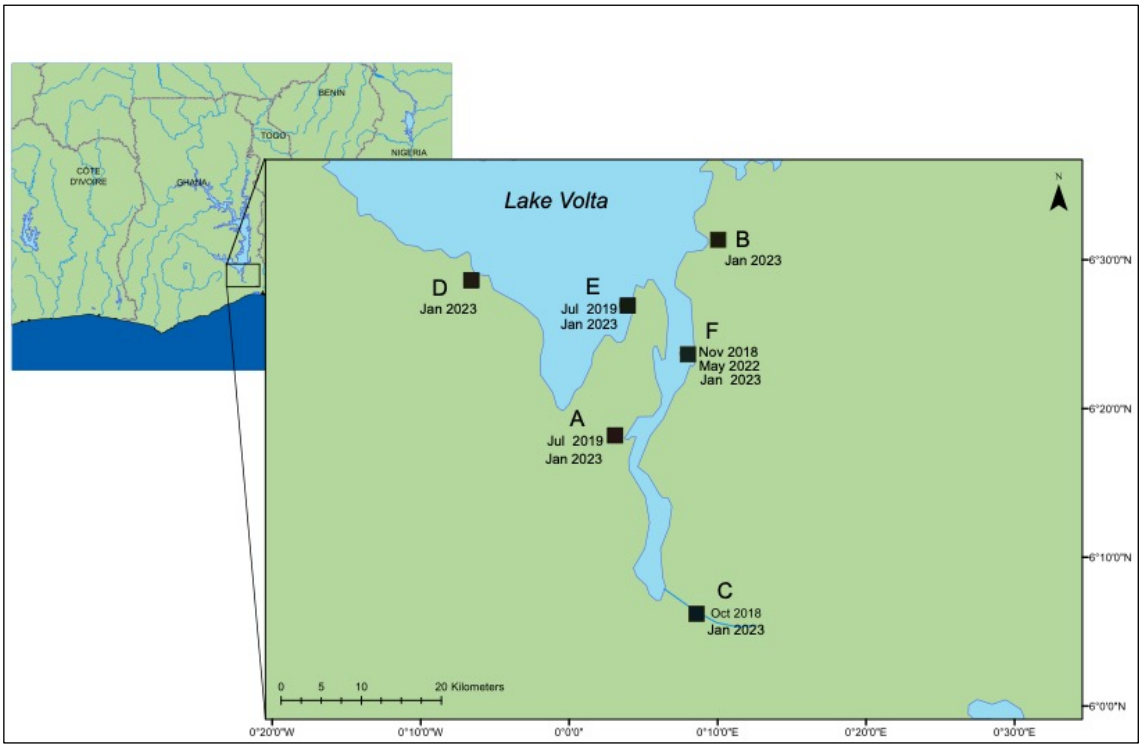


Figure 1. A Map of the lower region of Lake Volta; showing sampled farms between 2018-2023, and the date of sampling. Sample ID, type, and date of collection are listed in Supplementary Table 1 & 2.

Table 1. Labelling system for fish farms on Lake Volta, and a comparison with labels in previous study (Ramírez-Paredes et al. 2021).

Farm name (current study)	***Region***	***Farm name (Previous study)***
A	Akosombo	Farm 3 (near farm 7)
B	Dodi	New
C	Akuse	Farm 1
D	Akaten	New
E	Dasasi	Farm 6
F	Asikuma	Farm 2

3.2.2 DNA Extraction

DNA extraction from viral filters was undertaken using the Total nucleic acid Extraction Kit (MasterPure complete DNA/RNA purification kit, Epicenter). Using a Luer-lock syringe (Figure 2) excess liquid was flushed from the filter's housing prior to adding the extraction buffer. Extraction buffer was prepared by adding 2

µL from the supplied Proteinase K to 1 mL of the either X1 T+C lysis solution or Red Lysis buffer, resulting in 100 µg mL-1 Proteinase K concentration. A total of 1 mL of the extraction buffer was gently pushed from the outlet to the inlet of the filter using a 3 mL syringe. A further 3 mL syringe was connected to the filter inlet and the assembly was placed into a rotating incubator for 15 minutes at 65°C in a hybridization oven (Steward and Culley 2010; Mueller, Culley, and Steward 2014).The assembly was removed and allowed to cool briefly at room temperature. The extract was pulled into the aspiration syringe and transferred into a 2 mL microcentrifuge tube and chilled on ice for 3 minutes. One-half volume of MPC protein precipitation reagent was added and vortexed for 10 seconds. The debris was pelleted by centrifugation at 20,000 ×g for 15 minutes at 4°C, and the supernatant was transferred to a sterile 2 mL microcentrifuge tube, adding 1 µL of polyacryl carrier to the sample. An equal volume 100% isopropanol was added and mixed by inverting the tube. The sample was centrifuged at 20,000 ×g, for 45 min. The supernatant was then discarded, retaining the pellet, which was washed twice with 1 mL of 70% ethanol and centrifuged for 1 min. The pellet was air-dried, then dissolved in a 35 µL elution buffer (EB, NEB) heated to 50°C. An additional water sample from farm (F) was eluted in nuclease free water (NFW, Ambion).

DNA extraction from tissue samples was performed using the DNeasy Blood and Tissue kit (Qiagen, Manchester, UK), with a starting material of ~10 mg of tissue from pooled organs (liver, kidney and spleen), which were dried for 5 min prior to DNA extraction using the manufacturer's protocol. The nucleic acid, eluted in Elution Buffer, was stored at 4°C until processing. Quantification of DNA for water samples was performed using the high sensitivity reagents for the Qubit Fluorometer, with broad range reagents used for the tissue samples. Tissue samples were given an alphanumeric name in the format <farm>.<cage>.<fish>.

A positive control for water samples was used to test the efficiency of the DNA extraction method. This was done using the ISKNV viral particles collected from the 2019 outbreak from Lake Volta (lot: PM 38259) and passaged on BF-2 and or GF cell lines at Cefas. Infected cell lines were stored at -20°C and thawed at room temperature before filtration. Cell debris was removed by centrifugation at

900 ×g for 20 min, and the clarified supernatant was retained. Isolated virus from clarified harvested cell culture supernatant was filtered and DNA extracted, as mentioned above for field samples.

3.2.3 Droplet digital PCR for viral quantification:

To quantify the number of template strands of ISKNV in water samples, a droplet-digital PCR (ddPCR) amplification test was performed, using an Evagreen assay, described in (Alathari et al. 2023), in accordance with the manufacturer's instructions (Bio-Rad, USA). The positive control mentioned above was used as a positive control for viral quantification and detection using the ddPCR. The concentration of DNA input and results are shown in Supplementary Table 2.

For tissue samples, a probe-based ddPCR assay, using primers and probes by (Lin et al. 2017), were used following the manufacturer's instructions (Bio-Rad, USA), generating a 22 µL reaction. This was achieved following the same method described for the Evagreen assay, except the total concentration of the forward and reverse primer was 900 nM, and a concentration of 200 nM for the probe. The DNA volume template added was different according to sample concentration (Supplementary Table 2).

3.2.4 Tiled PCR

Extracted DNA from filtered water and tissue samples was quantified using a Qubit fluorometer, and a tiled PCR approach was performed to generate 2kb amplicons for sequencing. For water samples a total of 5µL of each DNA template was added to the reaction, and 1 µL of DNA was added for the tissue samples (concentrations are listed in Supplementary Table 2). The amount of DNA template in the water sample from farm (D) was too high and failed to amplify, therefore the amount was reduced to 2.5 µL. For tissue samples 0.1 µL of extracted DNA was taken forward for the tiled PCR (Supplementary Table 2). Two primer pools were prepared with alternating primer sets, described in (Alathari et al. 2023), and Q5 Hotstart High-Fidelity Polymerase (NEB) was used for amplification. Amplicons were quantified using the Qubit dsDNA BR kit (Invitrogen), and the two pools (A & B) of amplicons were combined.

3.2.5 Library preparation an sequencing

Long read Sequencing:

a. Water samples:

Amplicons generated from water samples from each farm and the prepared mock sample were taken forward for sequencing. Library preparation was performed using the Ligation Sequencing kit 1D (SQK-LSK109) (ONT) and Native Barcoding system (EXP-NBD104) (ONT), according to the manufacturer's instructions, and following the Native barcoding amplicon protocol: version *NBA_9093_v109_revD_12Nov2019.* Equimolar amounts of each barcoded sample were pooled and taken forward for the adaptor ligation step using a total volume of 60 µL of DNA, 5 µL of Adaptor Mix II (AMII), and 25 µL of Ligation Buffer (LNB) and 10 µL of T4 DNA Ligase were all added to the barcoded DNA. The reaction was incubated for 10 min at room temperature, and a 0.5× AMPure XP bead clean-up was performed, followed by 2 × 250 µL of SFB (ONT) washes. The pellet was then resuspended in 15 µL of Elution Buffer (EB) for 10 min at 37°C. 15 µL of the elute was retained and ~1 µg of adaptor ligated DNA was taken forward for priming and loading onto a FLO-MIN 106 (R9.4.1) flow cell.

A MinION run was performed for ~70 hours, and the flow cell was refuelled with FB after 25 hrs from the start of the sequencing run. All generated sequences were basecalled using the Oxford Nanopore Guppy tool, version v.6.0.4 with super high accuracy, and demultiplexed using guppy_barcoder. The Artic-Network pipeline and its accompanying tools were used to generate the consensus genomes. The pipeline uses viral nanopore sequencing data produced from tiling amplicon schemes, which are aligned to a reference genome to generate a consensus sequence. Initially, amplicons were filtered at 1800-2200 bp read lengths. Files generated by the Primal Scheme software were used to map the amplicons to the ISKNV reference genome. Nanopolish was used to produce a consensus sequence and identify genuine variants, and the percentage of genome recovery with at least 20× coverage was calculated (Alathari et al. 2023). The Artic pipeline was run with the parameter '--normalise 200' to subsample coverage >200×, and all other parameters were set to default: (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html). All

sequences were visualised and polymorphisms were evaluated in Geneious Prime 2022.1.1.

b. Tissue samples matching water samples

ONT updated their flow cells during this study, therefore a second library was prepared using the new R10.4 flow cell, to evaluate impact on variant calling. One tissue sample was selected from the same water sampled cages. One filter sample from farm F and one positive control filter sample (both previously sequenced), were sequenced alongside the matching tissue samples from the same cage, as a positive control, and were barcoded using the Native barcoding kit SQK-NBD114-24. Real-time basecalling was performed on MinKNOW version 23.04.5 with super high accuracy, to produce pod5 files, and demultiplexed with a requirement for barcodes on both ends and a minimum average q-score of 10. The total run was for ~22 hrs. Pod5 files were converted to fast5 files and downstream analysis was performed in a similar way to all previous samples except using Medaka (v.1.4.3) was used instead of nanopolish for variant calling, due to incompatibility between nanopolish and R10 data. Reads were processed using the Artic MinION method of the Artic bioinformatics pipeline: (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html)

c. All tissue samples

All amplicons generated from tissue samples that produced a visible band on gel electrophoresis following the tiled PCR, and where quantification indicated a concentration more than 10 ng/µL, were taken forward for sequencing. Samples that showed less than 400 viral templates/µL in a ddPCR assay were not taken forward for sequencing (Alathari et al. 2023). A total of 259 ng of DNA was loaded to a FLO-MIN 106 (R9.4.1) flow cell with following library preparation using the Ligation Sequencing kit 1D (SQK-LSK109) (ONT) and Barcoding system (EXP-NBD104) (ONT), according to the manufacturer's instructions: version *NBA_9093_v109_revD_12Nov2019*. The total run was for 72 hrs. A total of 5.24 million reads were generated, and the reads were processed as described above.

Short read sequencing

In contrast to tissue samples, where a fish is assumed to be infected by a single variant of ISKNV, water samples capture the population of variants circulating within a population. In such samples, consensus basecalling to remove read error from ONT reads is unable to discriminate between natural variation and sequencing error. Therefore, water samples from three farms (C, D, F) were selected to be sequenced using short read sequencing to identify the variants circulating the floating cages in the lake and determine if more than one variant was present. DNA was extracted as previously described and a tiled PCR was performed using the v2 primers (Alathari et al. 2023), to generate 2 kb amplicons spanning the full genome, followed by 0.6× bead clean-up with AMPure XP beads. Library preparation was performed with the DNA NEB PCR-free kit, followed by sequencing using the Illumina NovaSeq 6000 using a SP 300 flowcell. Short read sequences were trimmed using Artic guppylex, and mapped against the ISKNV reference genome from the NCBI (NC_003494) with minimap2 (Li 2018) to generate a bam file, which was visualised in Geneious (v. 2022.1.1). Reads were visualised and polymorphisms were identified in Geneious and IGV (v. 2.16.2).

3.2.6 Phylogeographic analysis

A phylogeographic tree was constructed comprising 52 whole genome sequences from fish samples collected between 2018- 2023, from (Alathari et al. 2023), and this study (Supplementary Table 3). Consensus genomes were aligned using the *augur* toolkit version 3.0.6 (github.com/Nextstrain/augur) in Nextstrain, where sequences were aligned using MAFFT (Katoh et al. 2002). The phylogenetic tree was reconstructed using IQ-Tree, with the GTR model, following all other default parameters and arguments in Nextstrain (Nguyen et al. 2015). The tree was further processed using augur translate and augur clade to assign clades to nodes and to integrate phylogenetic analysis with the metadata, where finally augur output was exported and visualised in auspice (github.com/Nextstrain/auspice) (Hadfield et al. 2018). All the consensus sequences generated from each sample were aligned to the ISKNV reference genome, accession no. (NC_003494). For a summary of methods see Figure 2.

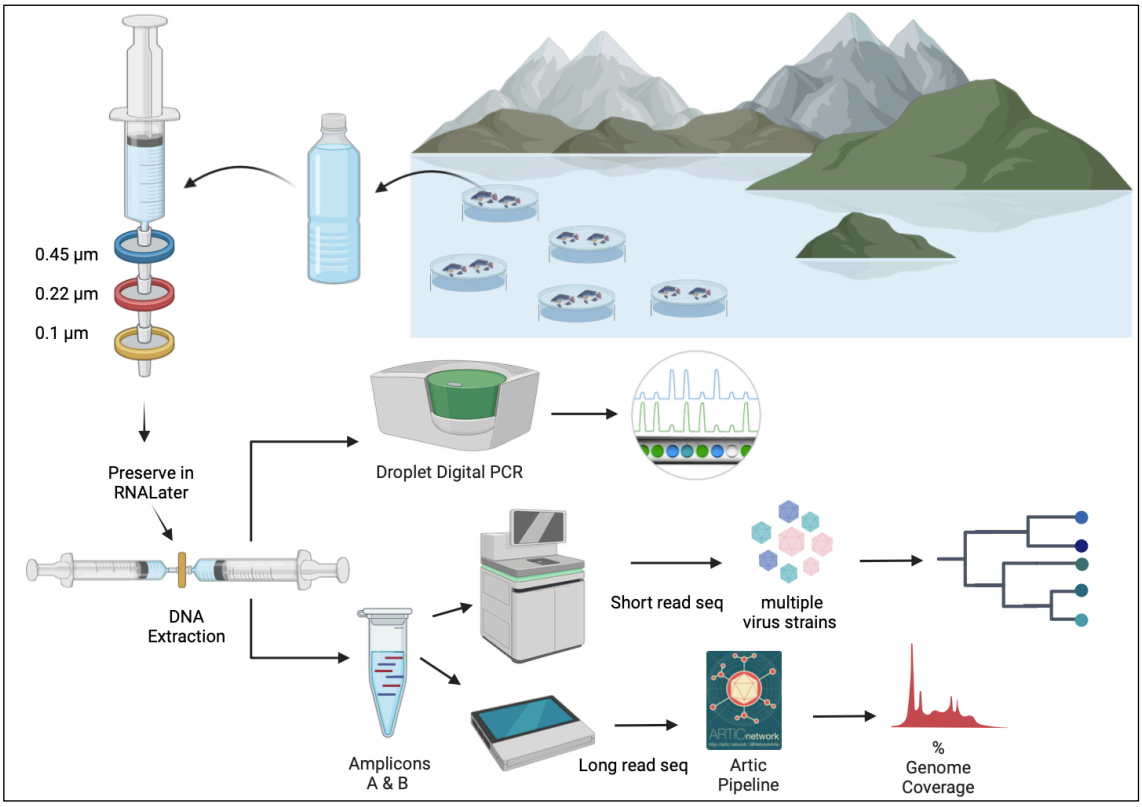


Figure 2. An overview of processing of water samples from around the tilapia cages on Lake Volta. The figure illustrates the concentrating of ISKNV onto filters, through to DNA extraction, quantification, and sequencing for variant detection. Figure was generated with BioRender (https://biorender.com/).

3.3 Results

3.3.1 ISKNV detection and quantification in tissue and water samples

DNA extraction was performed for tissue samples collected from six different locations across Lake Volta, with matching water samples taken at five locations. Quantification of DNA for all samples was performed using the Qubit Fluorometer and are provided in Supplementary Table 2.

ddPCR was used to detect and quantify the number of template strands of ISKNV in the extracted DNA from each tissue and water sample. Tissue samples were dominated by non-ISKNV DNA (most likely host DNA). There was an average of 317.45 ng/µL of DNA for all tissue samples, however ddPCR revealed low ISKNV viral template copies in most tissue samples; 71% of the tissue samples had fewer than 100 copies/µL, and in 14 out of the 74 tissue samples no ISKNV was

detected, mainly in fish sampled from farms (B) and (C). For water samples, the highest DNA concentration seen, at 21.4 ng/µL, was collected from farm (C).

The number of ISKNV templates in samples collected from water and tissue samples varied considerably across the different farm sites (Figure 3). At farm (C), tissue samples contained on average only 5 copies/µL, while the matching water sample had 174 copies/µL (Supplementary Figure 2). The average concentration of ISKNV templates found in tissue samples collected from farm (D) in contrast was much higher at 70.6 copies/µL except for one sample (D.3.3) and with very high viral templates, at 382,700 copies/µL, from one fish fingerling. The water samples collected from this cage site also had a high concentration of ISKNV at 361.2 copies/µL. The highest concentration of ISKNV in water samples was seen at farm (F), at 7,560 ISKNV copies /µL, followed by farms (D) & (C), respectively. Contrasting with these farms, (B) and (E) had very low concentrations of ISKNV in the water (~ 1 copy/µL). Despite the low water concentration of ISKNV at farm E tissues samples had a high ISKNV copy number, with at least 200 copies/µL. In six tissue samples collected from farm (F), the ddPCR failed to provide an accurate count. This was due to saturation of positive droplets at high concentration of DNA template, and this persisted despite further testing with a 20-fold dilution. Negative samples showed no viral template, while the mock filter sample (using viral particles harvested from cell culture) contained 1,584 ISKNV copies/µL. Heat-shocked fish samples from one cage in farm (F) showed no difference in the concentration of ISKNV compared with untreated (non-heat shocked) fish.

Spatial distribution of ISKNV detected across Lake Volta, showed the two farms (B, E) with very low concentrations of ISKNV in the water were both floating cages located far away from other farm cages, and were furthest from the shore (approximately 12 km). The highest titre of ISKNV, were seen in water samples collected from farm (F), and the highest concentration of ISKNV in tilapia were in juveniles and fingerlings. Moreover, fish in this farm showed the most obvious clinical signs and were experiencing ongoing mortality (Supplementary Table 1). In general, all life stages were positive for ISKNV, but the lowest concentrations were seen in adult fish (Supplementary Table 4).

A tiled PCR was performed on each sample, followed by a gel electrophoresis for each pool. All water samples yielded bands at 2kb, indicative of amplification of ISKNV. Bands for farms (B) & (E) were faint, supporting low template concentrations as measured by ddPCR (Supplementary Figure 3). Farm (E) showed multiple bands, with the strongest bands at 1kb. Despite some samples showing faint bands, all the tiled PCR products with any bands at 2kb were taken forward for sequencing.

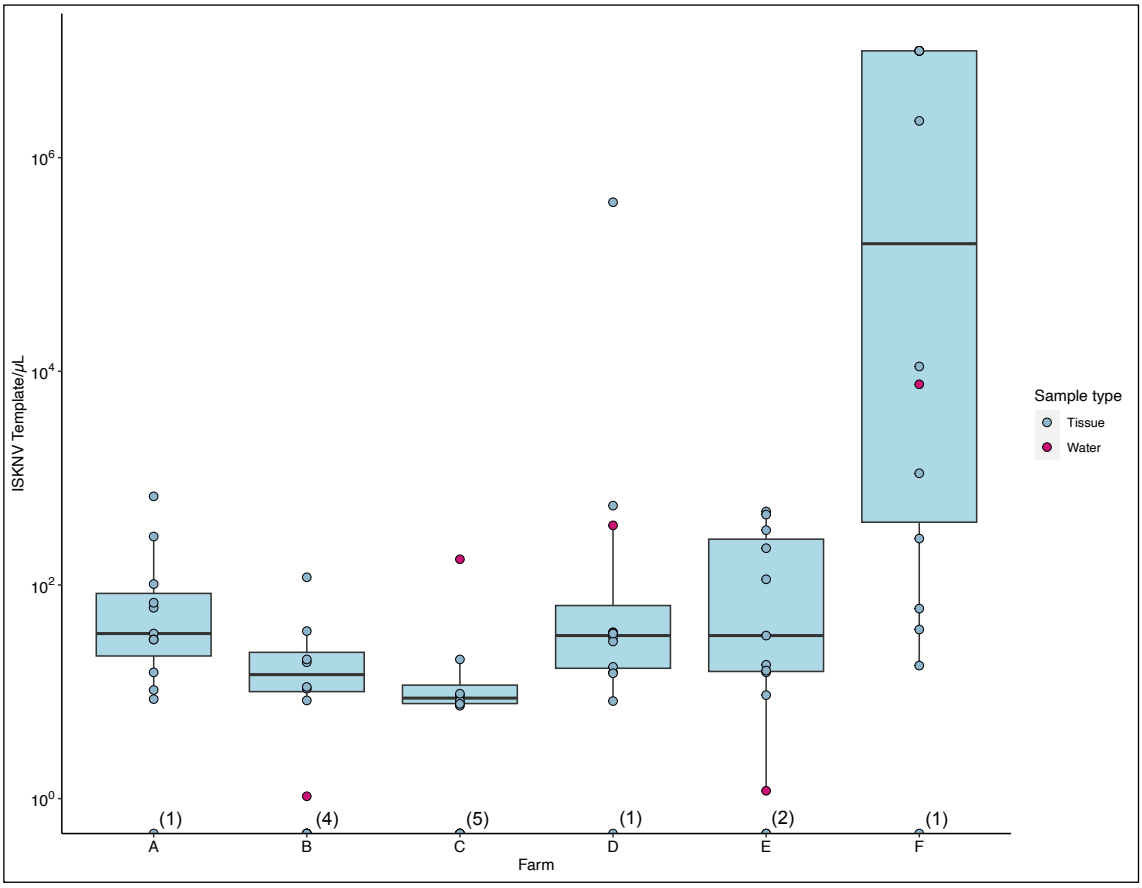


Figure 3: The number of viral templates of ISKNV in tissue and water samples collected from the ISKNV outbreak of 2023 in Lake Volta, Ghana; Distribution of ISKNV template strands in tissue samples (blue) and water samples (red). The number of samples with no ISKNV detected are given parentheses on the x axis.

3.3.2 Sequencing and phylogeographic analysis for all samples collected from Ghana- Changes to MinION chemistry do not affect our tiled PCR method

A total number of 4.93 M reads were produced from the five water samples, and a total of 5.23 M reads were generated from the five matched tissue samples.

The final sequencing run for ISKNV collected from tissue samples was 1.19 M reads. The median length of all samples is reported in Supplementary Table 3.

When compared to the ISKNV reference genome, the greatest proportion of the whole genome recovered was 98.18% in a tissue sample of a fingerling from cage 4 at farm (F). The highest genome recovery for water samples was 97.49%, collected from the same cage at farm F. Additionally, one sample (from fingerling tissue) from farm (D) had high genome recovery of 97.51%, matching water samples that showed high concentration of ISKNV by ddPCR, and sequencing resulted in genome recovery of 85.6%. Around two-thirds of all sequenced samples recovered at least 50% of the full ISKNV genome. In our previous study, we identified a minimum requirement of 482 copies/μL of ISKNV to yield a genome with >50% recovery (Alathari et al. 2023). Here, in water samples with fewer than 482 copies/μL of ISKNV produced more than 50% of genome recovery, suggesting lower input requirements for water samples due to an unknown mechanism. A list of genome recovery for each sample is provided in Supplementary Table 3.

Phylogeographic analysis was performed to investigate the epidemiology of ISKNV virus and disease in Lake Volta, and as a potential indicator of transmission for which closely related genomes indicate closely related infections, shown in Figure 4. For all except one case, the tissue samples collected from farms (E) and (F) in 2023, formed a separate clade, including the two water samples collected from those farms, and the water sample from farm (C). The 2023 tissue sample from farm (C) along with the tissue and water samples from farm (D) grouped together closely though were separate to earlier samples from the same farm (2018-2022). The highest divergence was seen in samples collected in 2023 from farm F sample (F.3.2), and was related most closely to samples collected from the same farm in 2022.

A group of samples collected from (F) in 2022 diverged from a clade of samples from a previous sampling at this location, clustering separately, due to a non-synonymous mutation (T3934C) occurring in the major capsid protein (MCP) that is unique to these samples. Genome recovery and variant detection was comparable between R9 and R10 flow cells.

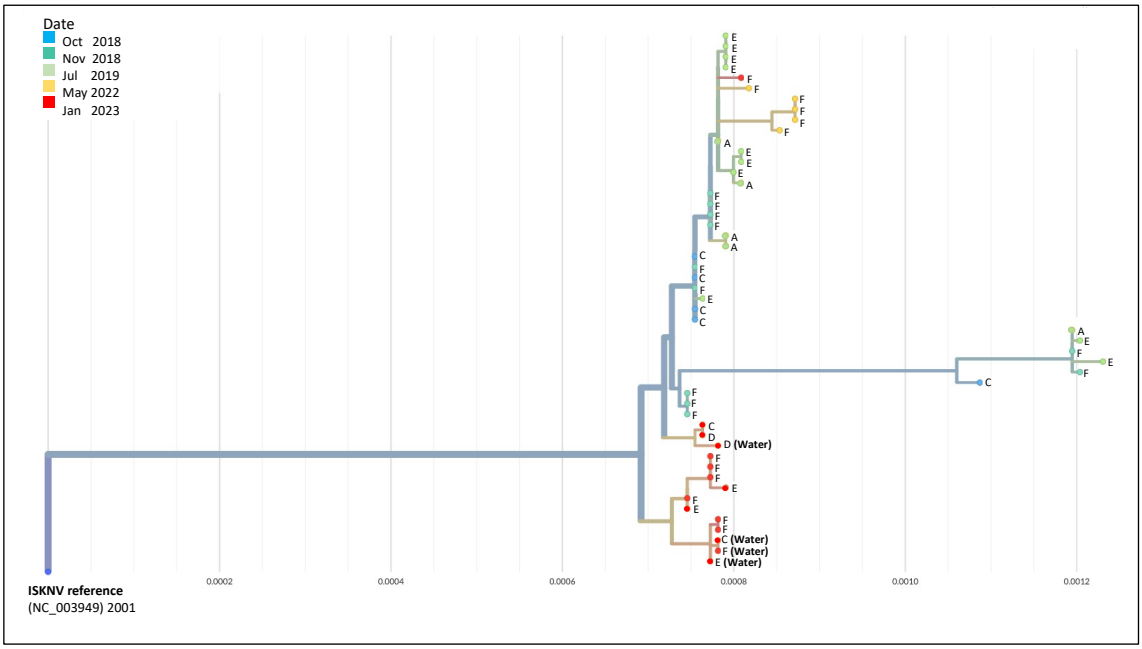


Figure 4. A phylogenetic tree of full ISKNV genomes from samples collected from Lake Volta, Ghana since 2018; Tissue and water samples collected from the latest outbreak were included and the colour represents the date of sampling. Water samples are identified in brackets. The tree was produced in Nextstrain (Hadfield et al. 2018).

To investigate differences of mutation profiles between genes across the ISKNV genome, we compared the percentage of polymorphic positions in any ORF for each of the genomes sequenced using the original ISKNV genome as a reference (Figure 5). The genomes selected were those that had 80% of genome recovery or above, compared with the reference ISKNV genome (2001), with remaining genomes removed from the analysis to avoid spurious SNPs from low coverage. Additionally, the repeat region (ORF025) was removed, as this represents a gene duplication and a potential region for circular permutation of the genome, rather than a coding region.

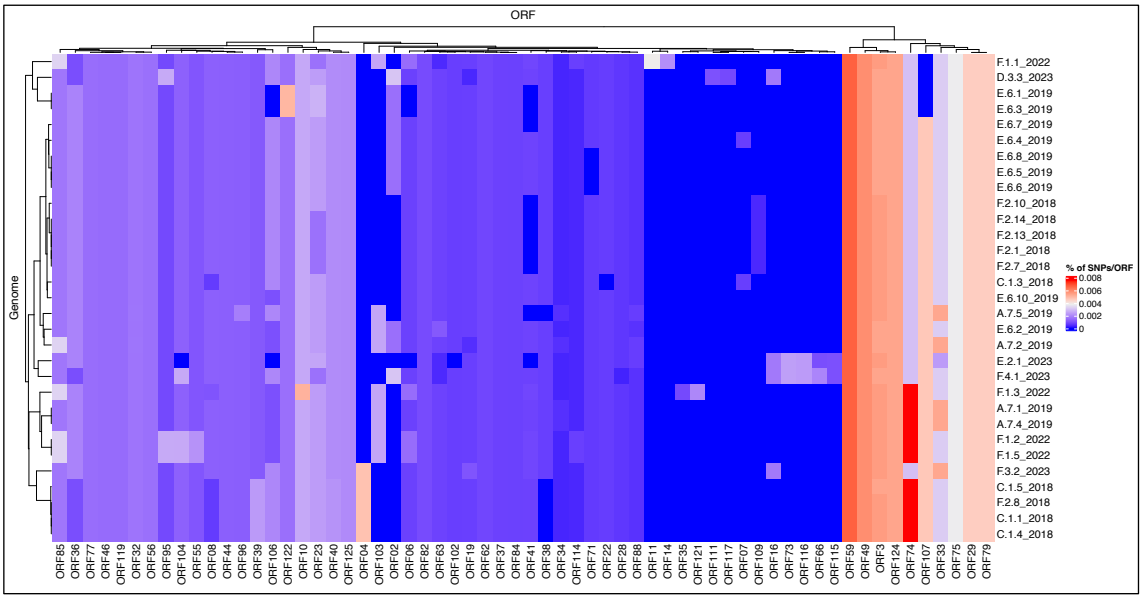


Figure 5. Mutational frequencies within the ISKNV genomes of fish tissue samples in Lake Volta, Ghana, since 2018; Heatmap shows the percentage of mutations per gene (ORF), represented on the x axis. Genomes with less than 80% genome recovery and ORFs with no mutations across all samples were removed, as well as ORF025 (repeat gene).

The highest percentage of mutations per gene were in ORF074 and ORF059, which have no assigned function. In general, ISKNV samples collected from Ghana had similar mutations, but samples collected from 2023 had mutations in samples collected from farms (E) and (F) which were not observed in any samples collected throughout previous years samplings. This was observed in Figure 4, where samples from these farms formed a separate clade. Mutations in the ORF004 were exclusively seen in four samples collected from 2018 with an outlier sample from (F.3.2) collected in 2023, which may explain its divergence from the clade of the outbreak of 2018 on the phylogenetic tree (Figure 4). All samples collected from Ghana shared a mutation in the ankyrin repeat protein (ORF125), an immunogenic gene, while another immunogenic gene (ORF117) showed a mutation only in a sample collected from farm (D). This mutation was also seen in the matching water sample. All samples had a mutation in ORF022, a proposed virulence gene, except one sample collected from 2018. Mutations in the MCP (ORF006) were higher in samples collected in 2022 than all other samples due to two mutations at this location for four out of five samples, resulting the formation of a separate clade on the phylogenetic tree (Figure 4).

Short read sequences of the water sample from farm (F), produced a total of 7,915,456 reads. Manual curation of the data using IGV (v. 2.16.2) showed the number of identified SNPs to be different to the number of SNPs detected using Geneious when using the default parameters to annotate and predict SNPs. A total of 86 SNPs were observed in IGV, while only 58 SNPs (46 SNPs with 200× coverage) were listed in Geneious, with five deletions, and two insertions. This suggests that parameters used for variant calling (such as coverage and percentage of variation at each position) for each software, may be different. A total of 27 of these were non-synonymous mutations. In comparison, the consensus sequence for the same sample generated using long read sequencing showed 46 SNPs with two insertions and four deletions. When examining the alignment in Geneious and variant/SNP calling using annotation default settings, some locations, such as a SNP in ORF058- C51,475T (coverage 4,282) was found to have a variant frequency of 90.7%, where 8.1% belonged to the original reference sequence. A similar SNP was manually detected in ORF040, location (C40,742T), however this SNP was not detected by the Geneious software, using the “annotate and predict” feature.

Long read sequencing of water sample from farm (F) showed 33 SNPs in common with short read sequences, and the same mutation at ORF058 with variant frequency of 89%, where only two fish tissue samples collected from the same farm showed the same mutation. Long read sequences from both water and tissue samples from farm (F), had 51 SNPs in common, with three extra SNPs that were unique to the water sample, and another three unique to tissue samples. Polymorphisms and substitutions were annotated in Geneious for short read and long read sequences from farm (F) listed in Supplementary Table 5.

In addition, water samples from farms (C) and (D) were sequenced using short read sequencing, and produced 21,394,234 and 16,788,272 reads, respectively. Annotation in Geneious detected 34 non-synonymous mutations out of a total of 48 SNPs in samples from farm (D), where 33 SNPs had at least 200x coverage, and four deletions. The unique SNP for ORF0117 in water and tissue samples collected from this farm using long read sequencing, was confirmed in short read sequencing. Thirty-three non-synonymous mutations out of a total of 48 SNPs were detected in the farm (C) sample. Thirty-nine out of the 48 SNPs were supported with >200x coverage. The two SNPs mentioned above in farm (F) were

detected in short reads from farm (C) but not in the farm (D). Finally, a mutation in the ISKNV MCP was confirmed by short read sequencing, and at the same location for all samples previously collected from Lake Volta outbreaks (Alathari et al. 2023). The ratio of non-synonymous and synonymous substitution rate (dN/dS ratio) was calculated for water samples C, D and F as follows: 2.2, 2.4, 0.8, respectively, suggesting positive selection at Farms C and D, but constraint at Farm F.

3.4 Discussion

This case study demonstrates the potential of using water samples for genomic surveillance of a large sized DNA virus, here for ISKNV infecting cultured tilapia in Ghana, using portable equipment in a farm setting. ISKNV was detected in both fish tissue samples and water samples collected from farm sites across Lake Volta and the in-field water sampling and sequencing method both distinguished between the different strains of the virus, and illustrated their relatedness. Sampling water within or close to the fish cages provides insight into the wider diversity of viruses on the farm than the more typical approach of tissue sampling because the latter is often based on a small number of fish, whereas the water may contain viral particles derived from many, potentially hundreds of fish, on the farm. Adopting the use of water sampling also avoids destructive sampling of fish with improved animal welfare benefits and reduced costs to the farmers.

ISKNV was detected in 81% of the fish sampled from the floating fish cages on Lake Volta in January 2023, with farm (F) having the highest viral load in both water and tissue samples of the farm sites studied. Although fish from farm (C) had very low concentrations of ISKNV in the body tissues sampled there was a relatively high concentration of viral particles in the surrounding water. Phylogenetic analysis of this water sample revealed it clustered with water samples collected from the upper region of Lake Volta. Farm (C) is surrounded by other tilapia farm cages in the Akuse region of Lake Volta and thus the likelihood is that ISKNV circulating strains may have been transported via the water from other nearby infected farms. In the current outbreak, some farmers reported a new trend of moribund tilapia, with fingerlings and juvenile fish being more susceptible than adult fish, and differing from that seen previously where no apparent age-related effect was reported. In our analysis, ISKNV positive samples were seen at all maturation stages of tilapia, but the lowest concentration and infection rates were, in general, seen in adult fish (Supplementary Table 4). This may be as a consequence of ISKNV now being practically endemic in Lake Volta, and thus fish surviving to adulthood have likely been exposed previously and thus on re-infection with new outbreaks are able to mount a more effective immune response, thus limiting viral replication.

Farms (E) and (B) showed very low concentrations of ISKNV in the water, but there were relatively high titres of virus particles detected in the tissue samples. The floating fish cages in both these two farms were located up to 12 km from the lake shoreline and from other farms, and this likely meant there was a far greater dilution of ISKNV from nearby fish in the water. Contrasting with this, water samples collected from farm (D) contained a high concentration of ISKNV, but for all of the fish, except one, sampled at this site there was a low body burden of the virus; this low viral template resulted in an inability to amplify it through our tiled PCR approach. We hypothesize this differential between the fish tissue and water titres of ISKNV might indicate a recent introduction of the virus to the farm and thus an early detection of virus presence through our water sampling approach, highlighting further the potential utility of water sampling in monitoring for this pathogen. Another explanation could be that the fish have recovered from a viral episode at the time of sampling, with the surviving fish having overcome the infection.

Integrating this data set with our previously sequenced genomes collected from Lake Volta, phylogenetic analysis groups the majority of the 2023 sequences in a separate clade indicating that the ISKNV currently infecting tilapia in Lake Volta, are not a descendent of an ongoing /previous infection but rather an emergence of a different endemic strain, or a new introduction to the Lake - most likely through fish importation. Moreover, farm (D) clustered separately from all other samples, except for a tissue sample from farm (C), and revealed an additional mutation in ORF 117 (C105,539A) in both its tissue and water samples. ORF117 is a transmembrane protein (Throngnumchai et al. 2021) which plays a vital role in viral replication and virulence (DiMaio 2014). The presence of this mutation likely explains why the water sample collected from farm (D) clustered separately from all other water samples. Further examination of the ratio between non-synonymous and the synonymous mutations could assist in estimating and identifying individual codon positions that are evolving under positive selection. For instance, it is curious that ISKNV in farms C and D are under positive selection, whereas those at Farm F are not, potentially suggesting a nuanced, localised relationship between ISKNV strains and associated environmental parameters and host phenotypes. The close relatedness in the water and tissue sample in farm (D) highlights the capability of water sampling in detecting current,

infective strains of ISKNV in fish. This was also confirmed when comparing the water and tissue samples of farm (F), where almost all SNPs were identical. It is also worth mentioning that the close relatedness of all but one of the samples collected in 2023, could indicate the same strain of ISKNV circulating in the water where this newly identified variant might be replacing the previous strain collected between 2018-2022. Sample (F.3.2) collected from farm (F) in 2023 clustered with samples collected in 2018 and might be a strain persisting from previous infections. Interestingly, water samples with less than 482 copies/µL of ISKNV, as calculated in our previous study (Alathari et al. 2023), were able to recover more than 50% of the full ISKNV genome, yet this wasn't possible for tissue samples. This could be due to an increased diversity in the environmental samples, allowing for more primer binding to extracted DNA, or the tissue DNA from tissue samples may contain inhibitors that may affect the amplification.

The heat map of mutational frequencies highlighted the presence of different SNPs in some of the samples collected in 2023 when compared to samples collected from previous years. We observed the presence of four new mutations in samples collected from farm (E.2.1) and (F.4.1), which were lacking in all the tissue samples collected previously. At least one SNP was seen in the MCP, but samples from 2022 showed two SNPs in this location. The second SNP could have become a reversed mutation in samples collected in 2023, and maybe have been corrected in the ISKNV genome due to its insufficient role in increasing the virus's fitness, or more simply the group that contains this second SNP wasn't sampled during this study. All samples collected from Ghana showed a mutation in ORF125 when compared to its reference genome. This ORF is an ankyrin repeat protein and also one of the major antigenic proteins and involved in modulating intracellular signalling networks during viral infections (Guo et al. 2011) (Throngnumchai et al. 2021).

Short read sequencing of a water sample collected from farm (F) showed a SNP in ORF40 (C40,742T) and ORF58 (C51,475T). The SNP located in ORF58 had a variant frequency of 91.6%, with 8.2% showing the original reference sequence, indicating circulation of more than one variant in the farm. The SNP in ORF40 was not detected by the Geneious software, only by manual analysis. This mutation was present in both short and long read sequences, with the short read sequencing able to show at least two strains circulating the water. Both mutations

were also seen in the water sample collected from farm (C) but not in farm (D), and could be the reason behind the (C) water sample clustering with (F) water sample. When comparing the water sample with the tissue samples, only two out of seven tissue samples collected from the same farm showed the same mutation at ORF058. This may indicate that the variant without the mutation at ORF058 derives from an historically earlier infection with a new mutation from a newly evolved variant. This is not presented in the heat map as the relevant samples, F.4.4, and F.4.3, generated a sequence recovery of less than 75% of the full ISKNV genome and were thus excluded from the analysis. Short read and long read sequencing produced a comparable number of SNPs and both approaches thus had the ability to detect the different variants.

In contrast to single-gene PCR approaches, whole genome sequencing can capture the full range of variants, providing vital information for vaccine and drug design. Other studies focusing on the MCP have shown their limitation in discriminating between viruses collected from different locations and at different time points (Ayiku et al. 2023). The portability of a next generation sequencer, and the invention of other portable technologies for amplicon generation and library preparation has led to long read sequencing being a preferred method for this analysis. These advancements have enabled performing studies like ours in remote and resource limited areas, with fast turnaround times, contrasting with that previously where the turnaround time at distant labs is in many months and likely unaffordable to many fish farm holders.

There are currently minimal disease control options for ISKNV and an urgent need for preventative measures. The approach we present in this paper for Lake Volta, show that water sampling has great potential for use in identifying the ISKNV associated with infected fish, and for determining the variants circulating within the system and infecting the fish at the time of sampling. This could assist in improving disease prevalence estimates and in the detection of emerging variants. The fact that in many cases the water for the inland ponds for hatchery stages is drawn from the lake, is likely the reason for the presence (and repeated cycling) of ISKNV infections in all fish life stages. Seeking to combat this cycle of infections and re-infections of ISKNV, encouraging farmers to seek, and pressure

for, farms designated free of ISKNV for their seeding stock would be a prudent step. Indeed, some larger farms with greater resources have already implemented this practice. Importantly, this requires that the supporting systems for aquaculture programmes in Ghana need to enable disease free hatcheries to be established and this inevitably requires also training of fisheries officers and farmers in biosecurity practice and the associated resources to deliver this.

The methods applied here to ISKNV, in addition to its capability for application to reach remote regions, could be adapted for other viral infections affecting the growth and development of aquaculture. Combining field data with in-field genomic tools can provide opportunities to understand the genetic architecture of disease resistance, leading to new opportunities for disease control in real time. Finally, there are very few available whole genome sequences for ISKNV and other important fish viruses in the database, therefore, and routine sequencing of these viruses will benefit significantly, understanding of the mutations that occur across the genome, and their role in virulence and/or transmissibility of the viral diseases in aquaculture.

Acknowledgements

SA was funded on a PhD programme through the centre for Sustainable Aquaculture Futures, a joint partnership between the University of Exeter and the Centre for Environment, Fisheries and Aquaculture Sciences (Cefas) which is gratefully acknowledged. We thank Janet Gyogluu Anchirinah, the Lake Volta local farmers and the Commission of Fisheries in Ghana involved in the investigation for their support and provision of information and samples. We also thank Joshua Quick for his insightful comments and Anke Lange, Victoria Jackson, Stephanie Andrews, Remy Chait, Exeter Microfluidics Facility, and Karen Moore for technical support and sequencing work. This project used equipment funded by the Wellcome Trust (Multi-User Equipment Grant award number 218247/Z/19/Z) to the Exeter Sequencing Centre.

Data Availability Statement

All data are deposited in NCBI BioProject ID: PRJNA935699.

References

Alathari, Shayma, Dominique L. Chaput, Luis M. Bolaños, Andrew Joseph, Victoria L. N. Jackson, David Verner-Jeffreys, Richard Paley, Charles R. Tyler, and Ben Temperton. 2023. “A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia.” *Viruses* 15 (4): 965.

Ayiku, Angela N. A., Abigail A. Adelani, Patrick Appenteng, Mary Nkansah, Joyce M. Ngoi, Collins M. Morang’a, Francis Djabeng, et al. 2023. “Molecular Epidemiology and Current Management of Infectious Spleen and Kidney Necrosis Virus (ISKNV) Infection in Ghanaian Cultured Tilapia.” *Aquaculture* , November, 740330.

Boykin, Laura, Ammar Ghalab, Bruno Rossitto De Marchi, Anders Savill, James M. Wainaina, Tonny Kinene, Stephen Lamb, et al. 2018. “Real Time Portable Genome Sequencing for Global Food Security.” *F1000Research* 7 (1101): 1101.

Brunner, Franziska S., Alexander Payne, Edward Cairns, George Airey, Richard Gregory, Natalie D. Pickwell, Myles Wilson, et al. 2023. “Utility of Wastewater Genomic Surveillance Compared to Clinical Surveillance to Track the Spread of the SARS-CoV-2 Omicron Variant across England.” *Water Research* 247 (October): 120804.

Cai, Junning. 2022. “Top 10 Species Groups in Global Aquaculture 2020.” *FAO*, June. http://dx.doi.org/.

Child, Harry T., Paul A. O’Neill, Karen Moore, William Rowe, Hubert Denise, David Bass, Matthew J. Wade, et al. 2023. “Optimised Protocol for Monitoring SARS-CoV-2 in Wastewater Using Reverse Complement PCR-Based Whole-Genome Sequencing.” *PloS One* 18 (4): e0284211.

Dharmadhikari, Tanmay, Vinay Rajput, Rakeshkumar Yadav, Radhika Boargaonkar, Dhawal Patil, Saurabh Kale, Sanjay P. Kamble, Syed G. Dastager, and Mahesh S. Dharne. 2022. “High Throughput Sequencing Based Direct Detection of SARS-CoV-2 Fragments in Wastewater of Pune, West India.” *The Science of the Total Environment* 807 (Pt 3): 151038.

DiMaio, Daniel. 2014. “Viral Miniproteins.” *Annual Review of Microbiology* 68 (April): 21–43.

FAO. 2019. “The State of the World’s Aquatic Genetic Resources for Food and Agriculture.” ReliefWeb. 2019. https://reliefweb.int/report/world/state-world-s-aquatic-genetic-resources-food-and-agriculture.

FAO. 2022. *Blue Transformation – Roadmap 2022–2030: A Vision for FAO’s Work on Aquatic Food Systems*. FAO.

Gardy, Jennifer, Nicholas J. Loman, and Andrew Rambaut. 2015. “Real-Time Digital Pathogen Surveillance - the Time Is Now.” *Genome Biology* 16 (1): 155.

Guo, Chang-Jun, Wei-Jian Chen, Li-Qun Yuan, Li-Shi Yang, Shao-Ping Weng, Xiao-Qiang Yu, and Jian-Guo He. 2011. “The Viral Ankyrin Repeat Protein (ORF124L) from Infectious Spleen and Kidney Necrosis Virus Attenuates Nuclear Factor-KB Activation and Interacts with IκB Kinase β.” *The Journal of General Virology* 92 (Pt 7): 1561–70.

Hadfield, James, Colin Megill, Sidney M. Bell, John Huddleston, Barney Potter, Charlton Callender, Pavel Sagulenko, Trevor Bedford, and Richard A. Neher. 2018. “Nextstrain: Real-Time Tracking of Pathogen Evolution.” *Bioinformatics* 34 (23): 4121–23.

Katoh, Kazutaka, Kazuharu Misawa, Kei-Ichi Kuma, and Takashi Miyata. 2002. “MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform.” *Nucleic Acids Research* 30 (14): 3059–66.

Li, Heng. 2018. “Minimap2: Pairwise Alignment for Nucleotide Sequences.” *Bioinformatics* 34 (18): 3094–3100.

Lin, Qiang, Xiaozhe Fu, Lihui Liu, Hongru Liang, Huizhi Guo, Shuwen Yin, Venkatesh Kumaresan, Zhibin Huang, and Ningqiu Li. 2017. “Application and Development of a TaqMan Real-Time PCR for Detecting Infectious Spleen and Kidney Necrosis Virus in Siniperca Chuatsi.” *Microbial Pathogenesis* 107 (June): 98–105.

Luo, Xiao, Xiongbin Kang, and Alexander Schönhuth. 2022. “Strainline: Full-Length de Novo Viral Haplotype Reconstruction from Noisy Long Reads.” *Genome Biology* 23 (1): 29.

Mueller, Jaclyn A., Alexander I. Culley, and Grieg F. Steward. 2014. “Variables Influencing Extraction of Nucleic Acids from Microbial Plankton (Viruses, Bacteria, and Protists) Collected on Nanoporous Aluminum Oxide Filters.” *Applied and Environmental Microbiology* 80 (13): 3930–42.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. “IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.” *Molecular Biology and Evolution* 32 (1): 268–74.

Quick, Joshua, Nathan D. Grubaugh, Steven T. Pullan, Ingra M. Claro, Andrew D. Smith, Karthik Gangavarapu, Glenn Oliveira, et al. 2017. “Multiplex PCR Method for MinION and Illumina Sequencing of Zika and Other Virus Genomes Directly from Clinical Samples.” *Nature Protocols* 12 (6): 1261–76.

Ragasa, Catherine, Harrison Charo-Karisa, Eugene Rurangwa, Nhuong Tran, and Kelvin Mashisia Shikuku. 2022. “Sustainable Aquaculture Development in Sub-Saharan Africa.” *Nature Food* 3 (2): 92–94.

Ramírez-Paredes, José Gustavo, Richard K. Paley, William Hunt, Stephen W. Feist, David M. Stone, Terence R. Field, David J. Haydon, et al.

2021. “First Detection of Infectious Spleen and Kidney Necrosis Virus (ISKNV) Associated with Massive Mortalities in Farmed Tilapia in Africa.” *Transboundary and Emerging Diseases* 68 (3): 1550–63.

Steward, Grieg F., and Alexander I. Culley. 2010. “Extraction and Purification of Nucleic Acids from Viruses.” In *Manual of Aquatic Viral Ecology*, 154–65. American Society of Limnology and Oceanography.

Throngnumchai, Boonyalit, Sarocha Jitrakorn, Pakkakul Sangsuriya, Sasimanas Unajak, Pongsak Khunrae, Ha Thanh Dong, Vanvimon Saksmerprome, and Triwit Rattanarojpong. 2021. “Refolded Recombinant Major Capsid Protein (MCP) from Infectious Spleen and Kidney Necrosis Virus (ISKNV) Effectively Stimulates Serum Specific Antibody and Immune Related Genes Response in Nile Tilapia (Oreochromis Niloticus).” *Protein Expression and Purification* 184 (August): 105876.

You, Weiwei, and Dennis Hedgecock. 2019. “Boom-and-bust Production Cycles in Animal Seafood Aquaculture.” Reviews in Aquaculture 11 (4): 1045–60.

3.5 Supplementary Material

A) Supplementary Tables:

Supplementary Table 1. Tissue samples collected from Lake Volta, showing the location and date of collection, size, and clinical signs observed. Sample ID represents the farm, cage, and fish, respectively.

Sample ID	*Region*	*Date*	*Stage*	*Clinical signs*	*Size*	*Organ*
A.1.1	A	09/01/2023	Adult	Healthy	240	Liver/Spleen
A.1.2	A	09/01/2023	Adult	Moribund opaque eyes	250	Liver/Spleen
A.1.3	A	09/01/2023	Adult	Moribund opaque eyes	260	Liver/Spleen
A.1.4	A	09/01/2023	Adult	Healthy cyst in gill	240	Liver/Spleen
A.2.1	A	09/01/2023	Juvenile	Healthy	170	Liver/Spleen
A.2.2	A	09/01/2023	Juvenile	Healthy	170	Liver/Spleen
A.2.3	A	09/01/2023	Juvenile	Healthy	140	Liver/Spleen
A.2.4	A	09/01/2023	Juvenile	Healthy	149	Liver/Spleen
A.3.1	A	09/01/2023	Fingerling	Healthy	80	Liver/Spleen
A.3.2	A	09/01/2023	Fingerling	Healthy	80	Liver/Spleen
A.3.3	A	09/01/2023	Fingerling	Healthy	95	Liver/Spleen
A.3.4	A	09/01/2023	Fingerling	Healthy	95	Liver/Spleen
B.1.1	B	10/01/2023	Adult	Dark eyes	160	Liver/Spleen
B.1.2	B	10/01/2023	Adult	Large white cyst	170	Liver/Spleen
B.1.3	B	10/01/2023	Adult	Healthy	190	Liver/Spleen
B.1.4	B	10/01/2023	Adult	Healthy	190	Liver/Spleen
B.2.1	B	10/01/2023	Juvenile	Healthy	65	Liver/Spleen
B.2.2	B	10/01/2023	Juvenile	Friable liver	70	Liver/Spleen
B.2.3	B	10/01/2023	Juvenile	Opaque eye	90	Liver/Spleen
B.2.4	B	10/01/2023	Juvenile	Healthy	110	Liver/Spleen
B.3.1	B	10/01/2023	Adult	Healthy	195	Liver/Spleen
B.3.2	B	10/01/2023	Adult	Healthy	220	Liver/Spleen
B.3.3	B	10/01/2023	Adult	Small liver, enlarged Gall bladder	220	Liver/Spleen
B.3.4	B	10/01/2023	Adult	Small liver, enlarged Gall bladder	190	Liver/Spleen
C.1.1	C	11/01/2023	Fingerling	Friable liver	80	Liver/Spleen
C.1.2	C	11/01/2023	Fingerling	Subserosa petechia	80	Liver/Spleen
C.1.3	C	11/01/2023	Fingerling	Healthy	75	Liver/Spleen
C.1.4	C	11/01/2023	Fingerling	Healthy	69	Liver/Spleen
C.2.1	C	11/01/2023	Juvenile	Healthy	120	Liver/Spleen
C.2.2	C	11/01/2023	Juvenile	Healthy	112	Liver/Spleen
C.2.3	C	11/01/2023	Juvenile	Healthy	120	Liver/Spleen
C.2.4	C	11/01/2023	Juvenile	Healthy	110	Liver/Spleen
C.3.1	C	11/01/2023	Adult	Darkened skin	200	Liver/Spleen
C.3.2	C	11/01/2023	Adult	Healthy	185	Liver/Spleen
C.3.3	C	11/01/2023	Adult	Healthy	165	Liver/Spleen
C.3.4	C	11/01/2023	Adult	Healthy	205	Liver/Spleen

D.1.1	D	12/01/2023	Adult	Friable liver	200	Liver/Spleen
D.1.2	D	12/01/2023	Adult	Friable liver	200	Liver/Spleen
D.1.3	D	12/01/2023	Adult	Healthy	195	Liver/Spleen
D.1.4	D	12/01/2023	Adult	Healthy	185	Liver/Spleen
D.2.1	D	12/01/2023	Juvenile	Fatty tissue	180	Liver/Spleen
D.2.2	D	12/01/2023	Juvenile	Friable liver	180	Liver/Spleen
D.2.3	D	12/01/2023	Juvenile	Healthy	180	Liver/Spleen
D.2.4	D	12/01/2023	Juvenile	Healthy	165	Liver/Spleen
D.3.1	D	12/01/2023	Fingerling	Healthy	85	Liver/Spleen
D.3.2	D	12/01/2023	Fingerling	Healthy	80	Liver/Spleen
D.3.3	D	12/01/2023	Fingerling	Darkened skin	80	Liver/Spleen
D.3.4	D	12/01/2023	Fingerling	Abrasion/lesion	80	Liver/Spleen
E.1.1	E	17/01/2023	Adult	Darkened skin	200	Liver/Spleen/Kidney
E.1.2	E	17/01/2023	Adult	Dark liver	245	Liver/Spleen/Kidney
E.1.3	E	17/01/2023	Adult	Pale liver	195	Liver/Spleen/Kidney
E.1.4	E	17/01/2023	Adult	Enlarged liver	200	Liver/Spleen/Kidney
E.2.1	E	17/01/2023	Juvenile	Healthy	90	Liver/Spleen/Kidney
E.2.2	E	17/01/2023	Juvenile	tail abrasion	90	Liver/Spleen/Kidney
E.2.3	E	17/01/2023	Juvenile	Healthy	110	Liver/Spleen/Kidney
E.2.4	E	17/01/2023	Juvenile	Healthy	95	Liver/Spleen/Kidney
E.3.1	E	17/01/2023	Fingerling	Healthy	25	Liver/Spleen/Kidney
E.3.2	E	17/01/2023	Fingerling	Healthy	20	Liver/Spleen/Kidney
E.3.3	E	17/01/2023	Fingerling	Healthy	20	Liver/Spleen/Kidney
E.3.4	E	17/01/2023	Fingerling	Healthy	20	Liver/Spleen/Kidney
F.1.1	F	18/01/2023	Adult	Moribund	270	Liver/Spleen/Kidney
F.1.2	F	18/01/2023	Adult	Louse	250	Liver/Spleen/Kidney
F.2.1	F	18/01/2023	Juvenile	White lips, small liver	75	Liver/Spleen/Kidney
F.2.2	F	18/01/2023	Juvenile	Prominent head kidney	95	Liver/Spleen/Kidney
F.2.3	F	18/01/2023	Juvenile	Enlarged liver tail erosion	95	Liver/Spleen/Kidney
F.2.4	F	18/01/2023	Juvenile	Healthy	85	Liver/Spleen/Kidney
F.3.1	F	18/01/2023	Fingerling	Small spleen, inflamed liver	55	Liver/Spleen/Kidney
F.3.2	F	18/01/2023	Fingerling	Healthy	45/46	Liver/Spleen/Kidney
F.3.3	F	18/01/2023	Fingerling	Friable liver tail erosion	45	Liver/Spleen/Kidney
F.3.4	F	18/01/2023	Fingerling	Healthy	60/44	Liver/Spleen/Kidney
F.4.1	F	18/01/2023	Fingerling	Ascites, white lips	75	Liver/Spleen/Kidney
F.4.2	F	18/01/2023	Fingerling	Ascites, pale liver	78	Liver/Spleen/Kidney
F.4.3	F	18/01/2023	Fingerling	Ascites, white lips, tail erosion, enlarged liver	75	Liver/Spleen/Kidney
F.4.4	F	18/01/2023	Fingerling	Ascites, enlarged liver, white lips	75	Liver/Spleen/Kidney

Supplementary Table 2. Table showing extracted DNA concentration and ddPCR results for filtered water and fish tissue samples. Samples with zero concentration using the ddPCR machine are highlighted in yellow.

Sample_ID	*Sample_type*	*Region*	*Qubit ng/µL*	*Input for ddPCR*	*conc. ddPCR*	*conc. x20 (copies)*	*copies /µL*
A.1.1	Tissue	A	262	52.4	0.0854	1.708	8.54
A.1.2	Tissue	A	41.2	8.24	0.612	12.24	61.2
A.1.3	Tissue	A	358	71.6	0.682	13.64	68.2
A.1.4	Tissue	A	115	23	0.152	3.04	15.2
A.2.1	Tissue	A	288	57.6	2.84	56.8	284
A.2.2	Tissue	A	104	20.8	6.75	135	675
A.2.3	Tissue	A	504	100.8	0	0	0
A.2.4	Tissue	A	230	46	0.351	7.02	35.1
A.3.1	Tissue	A	888	177.6	0.309	6.18	30.9
A.3.2	Tissue	A	480	96	1.02	20.4	102
A.3.3	Tissue	A	688	137.6	0.104	2.08	10.4
A.3.4	Tissue	A	490	98	0.308	6.16	30.8
B.1.1	Tissue	B	200	40	0	0	0
B.1.2	Tissue	B	150	30	0	0	0
B.1.3	Tissue	B	85.2	17.04	0	0	0
B.1.4	Tissue	B	456	91.2	0.107	2.14	10.7
B.2.1	Tissue	B	526	105.2	0.189	3.78	18.9
B.2.2	Tissue	B	194	38.8	0.111	2.22	11.1
B.2.3	Tissue	B	89.6	17.92	0	0	0
B.2.4	Tissue	B	632	138	1.18	23.6	118
B.3.1	Tissue	B	690	138	0.083	1.66	8.3
B.3.2	Tissue	B	862	172.4	0.37	7.4	37
B.3.3	Tissue	B	526	105.2	0	0	0
B.3.4	Tissue	B	496	99.2	0.201	4.02	20.1
C.1.1	Tissue	C	276	55.2	0.0921	1.842	9.21
C.1.2	Tissue	C	270	54	0.201	4.02	20.1
C.1.3	Tissue	C	408	81.6	0	0	0
C.1.4	Tissue	C	426	85.2	0	0	0
C.2.1	Tissue	C	670	134	0.074	1.48	7.4
C.2.2	Tissue	C	420	84	0.0778	1.556	7.78
C.2.3	Tissue	C	764	152.8	0	0	0
C.2.4	Tissue	C	1.5	0.3	0.0827	1.654	8.27
C.3.1	Tissue	C	354	70.8	0.077	1.54	7.7
C.3.2	Tissue	C	179	35.8	0.096	1.92	9.6
C.3.3	Tissue	C	179	35.8	0	0	0
C.3.4	Tissue	C	344	68.8	0	0	0
D.1.1	Tissue	D	222	44.4	0	0	0
D.1.2	Tissue	D	19.3	3.86	0.152	3.04	15.2
D.1.3	Tissue	D	69.2	13.84	0.361	7.22	36.1
D.1.4	Tissue	D	103	20.6	0.357	7.14	35.7
D.2.1	Tissue	D	418	83.6	0.171	3.42	17.1
D.2.2	Tissue	D	179	35.8	0.322	6.44	32.2
D.2.3	Tissue	D	43.4	8.68	0.296	5.92	29.6
D.2.4	Tissue	D	28.6	5.72	0.149	2.98	14.9
D.3.1	Tissue	D	204	40.8	0.0819	1.638	8.19
D.3.2	Tissue	D	356	71.2	0.351	7.02	35.1
D.3.3	Tissue	D	302	60.4	3827	76540	382700
D.3.4	Tissue	D	294	58.8	5.53	110.6	553
E.1.1	Tissue	E	326	65.2	0.152	3.04	15.2
E.1.2	Tissue	E	78	15.6	1.13	22.6	113
E.1.3	Tissue	E	168	33.6	0.336	6.72	33.6
E.1.4	Tissue	E	290	58	0	0	0
E.2.1	Tissue	E	165	33	4.87	97.4	487
E.2.2	Tissue	E	434	86.8	4.56	91.2	456
E.2.3	Tissue	E	87	17.4	3.26	65.2	326
E.2.4	Tissue	E	249	49.8	2.21	44.2	221
E.3.1	Tissue	E	328	65.6	0	0	0
E.3.2	Tissue	E	90.2	18.04	0.179	3.58	17.9
E.3.3	Tissue	E	326	65.2	0.0931	1.862	9.31

E.3.4	Tissue	E	324	64.8	0.158	3.16	15.8
F.1.1	Tissue	F	68.6	13.72	0.383	7.66	38.3
F.1.2	Tissue	F	704	140.8	0.601	12.02	60.1
F.2.1	Tissue	F	468	93.6	0.176	3.52	17.6
F.2.2	Tissue	F	960	192	100000	2000000	10000000
F.2.3	Tissue	F	36.6	7.32	111	2220	11100
F.2.4	Tissue	F	52.6	10.52	100000	2000000	10000000
F.3.1	Tissue	F	120	24	2.72	54.4	272
F.3.2	Tissue	F	606	121.2	100000	2000000	10000000
F.3.3	Tissue	F	566	113.2	11.1	222	1110
F.3.4	Tissue	F	Low	X5	0	0	0
F.4.1	Tissue	F	30.8	1.54	5521	110420	2208400
F.4.2	Tissue	F	51.2	2.56	100000	2000000	10000000
F.4.3	Tissue	F	498	24.9	100000	2000000	10000000
F.4.4	Tissue	F	262	13.1	100000	2000000	10000000
B.1(water)	Water Filter	B	2.1	10.5	0.263	5.26	1.052
C.1(water)	Water Filter	C	21.4	107	43.5	870	174
D.4(water)	Water Filter	D	5.42	27.1	90.3	1806	361.2
E.2(water)	Water Filter	E	1.85	9.25	0.296	5.92	1.184
F.4(water)	Water Filter	F	Low	X5	1890	37800	7560
Mock (pos. control)	Water Filter	F	20	10	396	7920	1584

Supplementary Table 3. Genome recovery and median length of ISKNV reads collected from tissue and water samples.

Sample_ID	***% of genome recovery***	***Median length of Reads***
A.3.1	59.56	783
B.1.2	0	619
C.1.3	21.33	678
D.3.3	97.51	1974
E.2.1	83.97	1936
E.2.2	77.83	1663
F.2.2	41.46	1007
F.2.3	5.28	523
F.3.4	81.27	1926
F.4.1	98.18	1985
F.4.2	92.53	1991
F.4.3	39.22	2001
F.4.4	72.3	1981
B.1(water)	13.14	374
C.1 (water)	66.45	1931
D.3 (water)	85.60	1952
E.2 (water)	19.46	1373
F.4 (water)	95.93	1981
Mock (pos. control)	92.06	1969

Supplementary Table 4. A summary table showing the average viral concentration for fish life stages in each farm.

Region	*Stage*	*Average viral temp/µL*
A	Adult	38.2
A	Juvenile	248.5
A	Fingerling	43.5
B	Adult	2.6
B	Juvenile	37
B	Fingerling	16.3
C	Adult	4.3
C	Juvenile	5.8
C	Fingerling	7.3
D	Adult	21.7
D	Juvenile	23.4
D	Fingerling	95824
E	Adult	40.4
E	Juvenile	372.5
E	Fingerling	10.7
F	Adult	49.2
F	Juvenile	5002779
F	Fingerling	5276222

Supplementary Table 5. List of Polymorphisms present in water and tissue samples from farm F. a) short read sequences using Novaseq; b) long reads sequencing (F.4) using ONT; c) long read sequencing using ONT, of matching tissue sample (F.4.1). SNPs were annotated and produced in Geneious Prime.

a)

Name	*Type*	*Coverage*	*product*	*Polymorphism Type*	*Min (original seq.)*	*AA Change*	*Codon Change*	*Protein Effect*
G	Polymorphism	65	ORF016L	SNP (transversion)	13195		ATA -> ATC	None
G	Polymorphism	849		SNP (transversion)	23819			
A	Polymorphism	18	ORF036R	SNP (transversion)	36128	S -> T	TCA -> ACA	Substitution
G	Polymorphism	2477	ORF040L	SNP (transversion)	40452	Q -> P	CAA -> CCA	Substitution
C	Polymorphism	2469	ORF049R	SNP (transversion)	47069	T -> P	ACT -> CCT	Substitution
C	Polymorphism	783	ORF055L	SNP (transversion)	49792	H -> Q	CAT -> CAG	Substitution
A	Polymorphism	400	ORF072R	SNP (transversion)	69075	D -> E	GAC -> GAA	Substitution
G	Polymorphism	20963	ORF075L	SNP (transversion)	70827	E -> Q	GAG -> CAG	Substitution
G	Polymorphism	3027	ORF084L	SNP (transversion)	78607	E -> D	GAA -> GAC	Substitution
G	Polymorphism	310		SNP (transversion)	84207			
T	Polymorphism	2	ORF100L	SNP (transversion)	89637	L -> Q	CTG -> CAG	Substitution
A	Polymorphism	5	ORF100L	SNP (transversion)	89673	Y -> F	TAT -> TTT	Substitution
C	Polymorphism	1001	ORF102R	SNP (transversion)	91711	K -> N	AAA -> AAC	Substitution
G	Polymorphism	74		SNP (transversion)	95137			
G	Polymorphism	3349	ORF119L	SNP (transversion)	107329	Q -> H	CAA -> CAC	Substitution

A	Polymorphism	9499	ORF001L	SNP (transition)	1180	H -> Y	CAC -> TAC	Substitution
A	Polymorphism	7486	ORF002R	SNP (transition)	1437	R -> Q	CGA -> CAA	Substitution
T	Polymorphism	139	putative major capsid protein	SNP (transition)	4328		ACG -> ACA	None
A	Polymorphism	891	ORF008R	SNP (transition)	6716		CCG -> CCA	None
G	Polymorphism	28	ORF014R	SNP (transition)	12088		CTA -> CTG	None
C	Polymorphism	67		SNP (transition)	13115			
G	Polymorphism	13		SNP (transition)	14307			
C	Polymorphism	1318	putative DNA polymerase	SNP (transition)	15309	V -> A	GTG -> GCG	Substitution
G	Polymorphism	41208	ORF023R	SNP (transition)	20742	H -> R	CAC -> CGC	Substitution
G	Polymorphism	86	ORF023R	SNP (transition)	21589		AAA -> AAG	None
T	Polymorphism	111	ORF025R	SNP (transition)	23362		CGC -> CGT	None
A	Polymorphism	344	ORF025R	SNP (transition)	23422		ACG -> ACA	None
C	Polymorphism	366	ORF025R	SNP (transition)	23425		CGT -> CGC	None
C	Polymorphism	842	ORF025R	SNP (transition)	23506		CGT -> CGC	None
C	Polymorphism	111	ORF028L	SNP (transition)	27696		GTA -> GTG	None
A	Polymorphism	208	ORF028L	SNP (transition)	28422		GAC -> GAT	None
C	Polymorphism	149	ORF029L	SNP (transition)	28746		CGA -> CGG	None
G	Polymorphism	333	putative thymidine kinase	SNP (transition)	29986		CCA -> CCG	None
C	Polymorphism	20893	ORF033L	SNP (transition)	30916	K -> R	AAA -> AGA	Substitution
C	Polymorphism	21018	ORF033L	SNP (transition)	30928	H -> R	CAC -> CGC	Substitution
G	Polymorphism	51	putative DNA-directed RNA polymerase II	SNP (transition)	33129		CCA -> CCG	None
T	Polymorphism	19	ORF036R	SNP (transition)	36598		CGC -> CGT	None
T	Polymorphism	19	0RF037L	SNP (transition)	36598		TAG -> TAA	None
A	Polymorphism	2186	ORF039R	SNP (transition)	40239		ACG -> ACA	None
T	Polymorphism	244	ORF041L	SNP (transition)	42101	E -> K	GAG -> AAG	Substitution
T	Polymorphism	1780	ORF044L	SNP (transition)	44244		CAG -> CAA	None
G	Polymorphism	83	putative cytosine DNA methyltransferase	SNP (transition)	45951		ATT -> ATC	None
T	Polymorphism	3651	ORF058L	SNP (transition)	51475		GCG -> GCA	None
C	Polymorphism	391	ORF062L	SNP (transition)	53283	Q -> R	CAG -> CGG	Substitution
G	Polymorphism	27884	putative ankyrin repeat protein	SNP (transition)	74533	N -> D	AAT -> GAT	Substitution
T	Polymorphism	27245	putative ankyrin repeat protein	SNP (transition)	74765	T -> I	ACA -> ATA	Substitution
C	Polymorphism	5612	ORF082L	SNP (transition)	77943	H -> R	CAT -> CGT	Substitution
A	Polymorphism	89	ORF088R	SNP (transition)	82637		TTG -> TTA	None
A	Polymorphism	8100	ORF104R	SNP (transition)	92428	V -> M	GTG -> ATG	Substitution

T	Polymorphism	10188	ORF104R	SNP (transition)	92798	T -> I	ACA -> ATA	Substitution
G	Polymorphism	11413	ORF104R	SNP (transition)	92936	Q -> R	CAG -> CGG	Substitution
T	Polymorphism	269	ORF107L	SNP (transition)	94693		TCG -> TCA	None
A	Polymorphism	315		SNP (transition)	94895			
A	Polymorphism	114	ORF115R	SNP (transition)	103808		ACG -> ACA	None
A	Polymorphism	114	ORF116L	SNP (transition)	103808		GCC -> GCT	None
T	Polymorphism	2130		SNP (transition)	106507			
C	Polymorphism	1106	ORF124R	SNP (transition)	110565	C -> R	TGT -> CGT	Substitution
C	Polymorphism	1496	putative ankyrin repeat protein	SNP (transition)	110889	S -> G	AGC -> GGC	Substitution

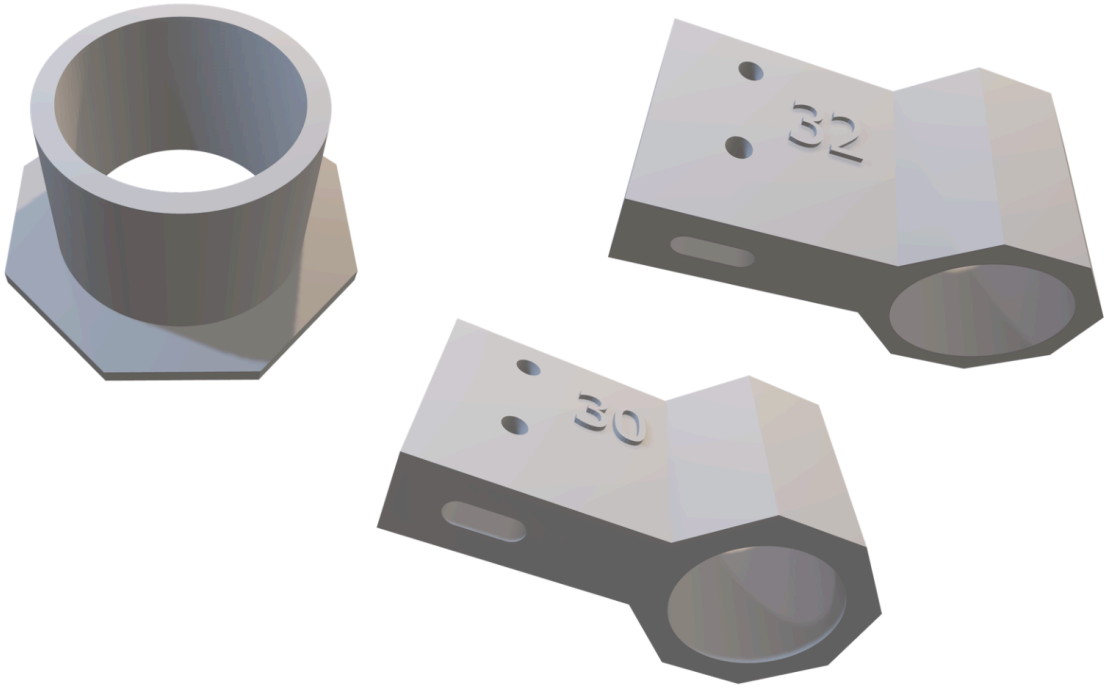
b)

Name	***Type***	***product***	***Polymorphism Type***	***Min (original sequence)***	***Amino Acid Change***	***Codon Change***	***Protein Effect***
A	Polymorphism	ORF002R	SNP (transversion)	1925	F -> I	TTT -> ATT	Substitution
G	Polymorphism	putative DNA polymerase	SNP (transversion)	16403	S -> A	TCG -> GCG	Substitution
C	Polymorphism	ORF023R	SNP (transversion)	20044	E -> D	GAA -> GAC	Substitution
A	Polymorphism	ORF023R	SNP (transversion)	21331	D -> E	GAC -> GAA	Substitution
G	Polymorphism	ORF023R	SNP (transversion)	21335	Q -> E	CAG -> GAG	Substitution
C	Polymorphism	ORF033L	SNP (transversion)	30245	Q -> E	CAG -> GAG	Substitution
A	Polymorphism	ORF036R	SNP (transversion)	36128	S -> T	TCA -> ACA	Substitution
G	Polymorphism	ORF040L	SNP (transversion)	40452	Q -> P	CAA -> CCA	Substitution
C	Polymorphism	ORF049R	SNP (transversion)	47069	T -> P	ACT -> CCT	Substitution
C	Polymorphism	ORF055L	SNP (transversion)	49792	H -> Q	CAT -> CAG	Substitution
G	Polymorphism	ORF075L	SNP (transversion)	70827	E -> Q	GAG -> CAG	Substitution
G	Polymorphism	ORF084L	SNP (transversion)	78607	E -> D	GAA -> GAC	Substitution
C	Polymorphism	ORF102R	SNP (transversion)	91711	K -> N	AAA -> AAC	Substitution
G	Polymorphism	ORF119L	SNP (transversion)	107329	Q -> H	CAA -> CAC	Substitution
A	Polymorphism	ORF122L	SNP (transversion)	109482	N -> Y	AAC -> TAC	Substitution
A	Polymorphism	ORF001L	SNP (transition)	1180	H -> Y	CAC -> TAC	Substitution
A	Polymorphism	ORF002R	SNP (transition)	1437	R -> Q	CGA -> CAA	Substitution
C	Polymorphism	ORF010L	SNP (transition)	9039	T -> A	ACA -> GCA	Substitution
C	Polymorphism	putative DNA polymerase	SNP (transition)	15309	V -> A	GTG -> GCG	Substitution
T	Polymorphism	ORF022L	SNP (transition)	18024	S -> N	AGC -> AAC	Substitution
A	Polymorphism	ORF023R	SNP (transition)	20006	E -> K	GAG -> AAG	Substitution
G	Polymorphism	ORF023R	SNP (transition)	20742	H -> R	CAC -> CGC	Substitution
C	Polymorphism	ORF033L	SNP (transition)	30916	K -> R	AAA -> AGA	Substitution

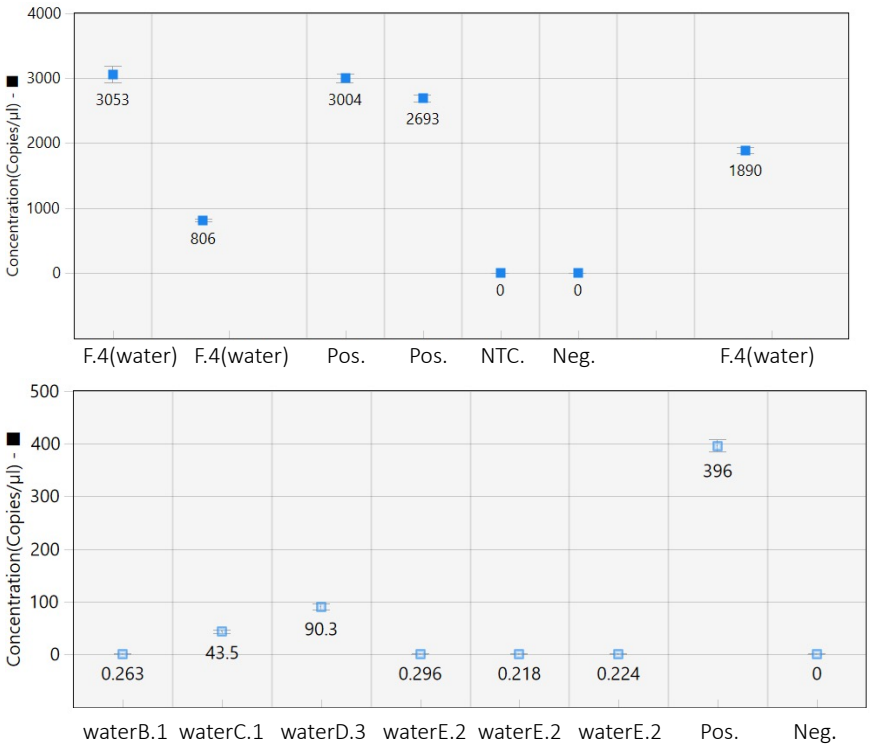
C	Polymorphism	ORF033L	SNP (transition)	30928	H -> R	CAC -> CGC	Substitution
C	Polymorphism	ORFO38L	SNP (transition)	38105	K -> E	AAA -> GAA	Substitution
T	Polymorphism	ORF040L	SNP (transition)	40812	R -> Q	CGG -> CAG	Substitution
T	Polymorphism	ORF041L	SNP (transition)	42101	E -> K	GAG -> AAG	Substitution
G	Polymorphism	ORF059R	SNP (transition)	51936	Q -> R	CAG -> CGG	Substitution
C	Polymorphism	ORF062L	SNP (transition)	53283	Q -> R	CAG -> CGG	Substitution
A	Polymorphism	ORF062L	SNP (transition)	55273	R -> C	CGC -> TGC	Substitution
C	Polymorphism	putative NTPase	SNP (transition)	59200	M -> V	ATG -> GTG	Substitution
C	Polymorphism	ORF071L	SNP (transition)	66437	I -> V	ATT -> GTT	Substitution
C	Polymorphism	ORF073R	SNP (transition)	69395	C -> R	TGT -> CGT	Substitution
G	Polymorphism	putative ankyrin repeat protein	SNP (transition)	74533	N -> D	AAT -> GAT	Substitution
T	Polymorphism	putative ankyrin repeat protein	SNP (transition)	74765	T -> I	ACA -> ATA	Substitution
A	Polymorphism	ORF079L	SNP (transition)	75893	P -> L	CCG -> CTG	Substitution
C	Polymorphism	ORF082L	SNP (transition)	77943	H -> R	CAT -> CGT	Substitution
C	Polymorphism	ORF085R	SNP (transition)	80307	C -> R	TGT -> CGT	Substitution
T	Polymorphism	ORF095L	SNP (transition)	86425	D -> N	GAT -> AAT	Substitution
A	Polymorphism	ORF104R	SNP (transition)	92428	V -> M	GTG -> ATG	Substitution
T	Polymorphism	ORF104R	SNP (transition)	92798	T -> I	ACA -> ATA	Substitution
G	Polymorphism	ORF104R	SNP (transition)	92936	Q -> R	CAG -> CGG	Substitution
G	Polymorphism	ORF114L	SNP (transition)	102888	L -> S	TTG -> TCG	Substitution
C	Polymorphism	ORF119L	SNP (transition)	107868	M -> V	ATG -> GTG	Substitution
C	Polymorphism	ORF124R	SNP (transition)	110565	C -> R	TGT -> CGT	Substitution
C	Polymorphism	putative ankyrin repeat protein	SNP (transition)	110889	S -> G	AGC -> GGC	Substitution

B) Supplementary figures

Supplementary Figure 1. The structure of an in-house created adaptor for holding syringes, to facilitate pumping water to filter.

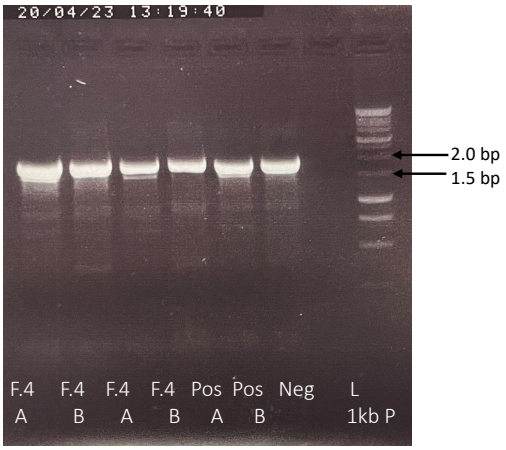


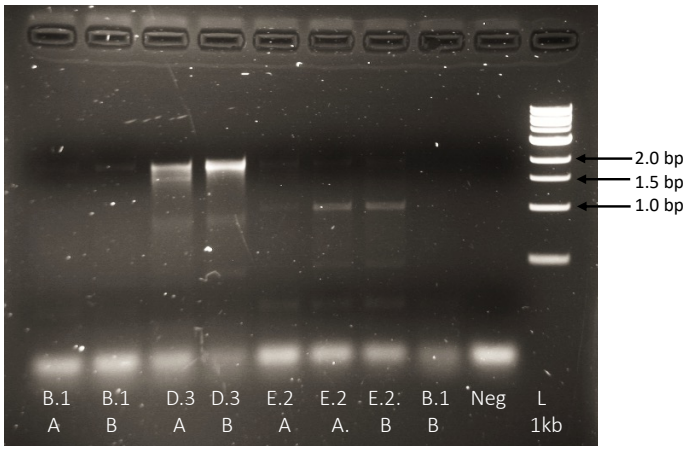
Supplementary Figure 2. Scatter plot showing concentration of ISKNV detected by ddPCR from water filters.



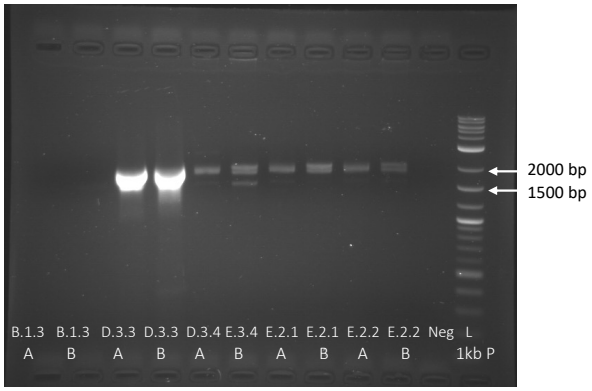
Supplementary Figure 3. Gel electrophoresis images showing amplicons produced by tiled PCR for ISKNV for: a) Water samples **b)** Tissue samples. 1.5% agarose gel was used to visualize the PCR products, with expected bands at 2kb. L, DNA ladder (1kb and1kb Plus) (New England Biolabs).

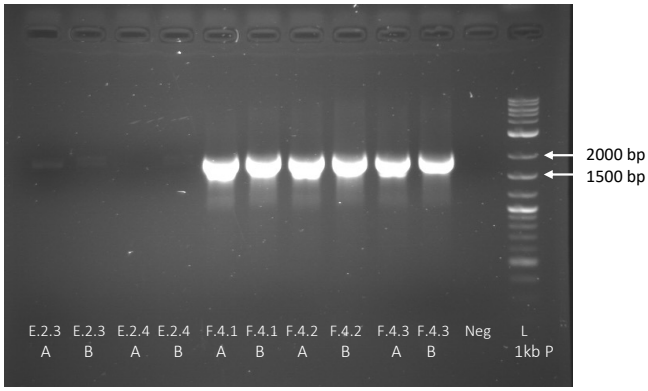
a)



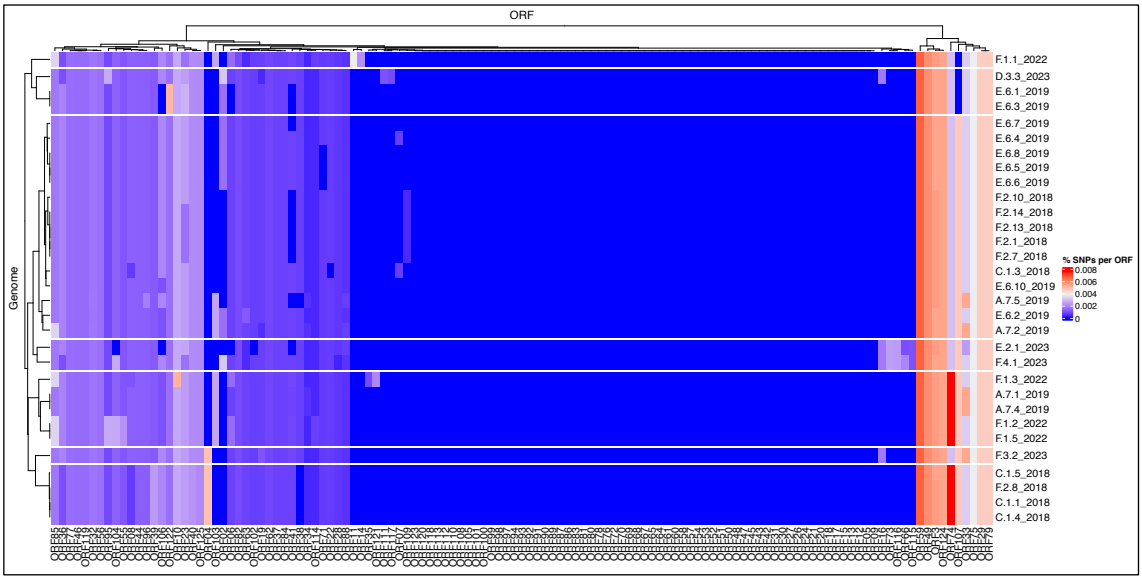


b)





Supplementary Figure 4. Mutational frequencies within the genomes of ISKNV from fish tissues sampled from Lake Volta, Ghana, since 2018 (All ORFs); Heatmap shows the percentage of mutations per gene (ORF), represented on the x axis. Genes with no mutations are included and genomes with less than 80% genome recovery and including all the ORFs.



4. Evaluation of a tiled PCR method for a segmented RNA virus affecting the growth of tilapia aquaculture for more than a decade- Tilapia Lake Virus

Abstract:

Tilapia Lake Virus (TiLV) is a segmented, negative sense single stranded RNA virus, belonging to the family Amnoonviridae. Since its first reporting, TiLV has been identified in 16 tilapia-producing countries, and it is believed that more than 45 countries are at high risk of TiLV (FAO 2017; Debnath et al. 2020; Aich et al. 2022). Determining the provenance of TiLV and tracking its movement across borders are crucial elements for minimising the impact of this disease on farmed and wild fish populations, where effective diagnostics and rapid sequencing methods for surveillance can play an essential role for reducing its spread (Chaput et al. 2020; Delamare-Debouteville et al. 2023). However, despite the huge socio-economic impact of TiLV, there are currently only a few published whole genomes of this virus, severely affecting the prediction of its origin, evolution, and epidemiology (Abbadi et al. 2023).

Here we developed a tiled PCR approach for full whole genome sequencing of TiLV by generating tiles of amplicons for each segment, in a similar approach for ISKNV previously described in Chapter 3. We compared this method to short-read sequencing approaches through the (re)sequencing of samples previously collected from Bangladesh (Chaput et al. 2020), and to samples collected from two farms in Thailand. The tiled PCR approach was able to differentiate between samples according to country and farm, highlighting its potential for use for infield real-time genomic surveillance tool for the tracking and in turn control and containment of TiLV.

4.1 Introduction:

Tilapia are increasingly important to domestic and global food security, yet there has been ~260 000 tonnes of decline in world produce, primarily reflected production drops in Indonesia and Egypt which were, respectively, the second and third largest producers, accounting for nearly 40 percent of world production in 2020 (Cai 2022). A wide range of bacteria, fungi, protozoa, and viruses have been described as a challenge to tilapia's growth. While bacterial and fungal infections have been addressed through the use of antibiotics or topical treatments, no specific therapy has been described for viral infections of tilapia. Viruses were not implicated as substantive threats until 2009, when massive losses were described in Israel and Ecuador (Bacharach et al. 2016).

4.1.1 Outbreaks of TiLV

The first recorded outbreak of TiLV occurred in Israel, when Eyngor and colleagues reported a syndrome comprising lethargy, endophthalmitis, skin erosions, renal congestion, ocular alterations, and encephalitis, with transmissibility of disease from affected to naïve fish (tilapia). TiLV was identified as a novel orthomyxo-like virus and it now poses a global threat to tilapia aquaculture (Eyngor et al. 2014; Bacharach et al. 2016). Since the first recorded outbreak, TiLV has now been reported in other major tilapia producing countries, including Ecuador (Bacharach et al. 2016), Egypt (Nicholson et al. 2017; Fathi et al. 2017), and Thailand (Pulido et al. 2019). A number of tilapia species have been reported to be affected by TiLV, such as Nile tilapia (*O. niloticus*); red tilapia (*Oreochromis* sp.); hybrid tilapia (*O. niloticus* × *O. aureus*), and various species of wild tilapia, causing mortality up to 90%. Moreover, all life stages of tilapia have been shown to be susceptible (Aich et al. 2022).

The outbreak of TiLV had not been reported to affect any other species in polyculture systems incorporating tilapia, suggesting the specificity of the disease only to tilapia, even after long-term cohabitation (Eyngor et al. 2014; Fathi et al. 2017; Behera et al. 2018; Chaput et al. 2020). However, in 2017, TiLV disease was reported from river carp in Malaysia, increasing concerns that TiLV could

infect other freshwater fish (Abdullah et al. 2018; Chaput et al. 2020). Moreover, an experimental challenge with TiLV showed that giant gourami is an additional susceptible species, with cohabitation between tilapia and giant gourami causing this cross-species transmission (Jaemwimol et al. 2018). Co-infection of different bacterial pathogens in TiLV infected fish has been reported, that include but are not restricted to *Aeromonas* spp., *Flavobacterium* spp. and *Streptococcus* spp, with *Aeromonas* spp. being more frequent than with other bacterial species (Abdullah et al. 2018; Aich et al. 2022).

In Thailand, in 2016 & 2017, severe die-offs were observed in red tilapia fingerlings during the first month after being transferred into floating cages, and within hatcheries. Bacterial infection was initially suspected, but disease surveillance confirmed that the disease outbreaks in farmed tilapia was associated with TiLV (H. T. Dong, Siriroob, et al. 2017). In the same year (2017), following a severe tilapia mortality event in Bangladesh, the presence of TiLV was confirmed (Chaput et al. 2020). Between 2020-2021 a study by Piewbang et al, described a coinfection of Tilapia parvovirus TiPV with TiLV, in multiple independent farms in Thailand causing significant losses (Piewbang et al. 2022).

4.1.2 TiLV Characteristics:

TiLV is an enveloped, segmented, negative sense single stranded RNA virus, which was initially proposed to belong to the family Orthomyxoviridae, due to similarities in the structure of its segment termini. Subsequently TiLV has been placed in a new family, Amnoonviridae, with the Linnaean classification *Tilapia tilapinevirus*. Its genome length is 10,323 bp and it contains 10 genome segments, encoding 14 predicted proteins (Aich et al. 2022; Chaput et al. 2020). The largest segment, segment 1, contains a predicted protein with weak homology to the PB1 subunit of influenza C virus, an orthomyxovirus (Bacharach et al. 2016). The other nine segments show no recognizable homology to other viruses but have conserved, complementary sequences at their 5' and 3' termini, consistent with the genome organisation found in other orthomyxoviruses (Bacharach et al. 2016). Electron microscopy has revealed that TiLV are enveloped icosahedral particles of 55 to 75 nm (Figure 1), with sensitivity to

organic solvents (ether and chloroform) indicating that TiLV is an enveloped virus (Eyngor et al. 2014). The first TiLV genome was sequenced using a shotgun transcriptome approach on an Illumina sequencing platform (Bacharach et al. 2016; Al-Hussinee et al. 2018).

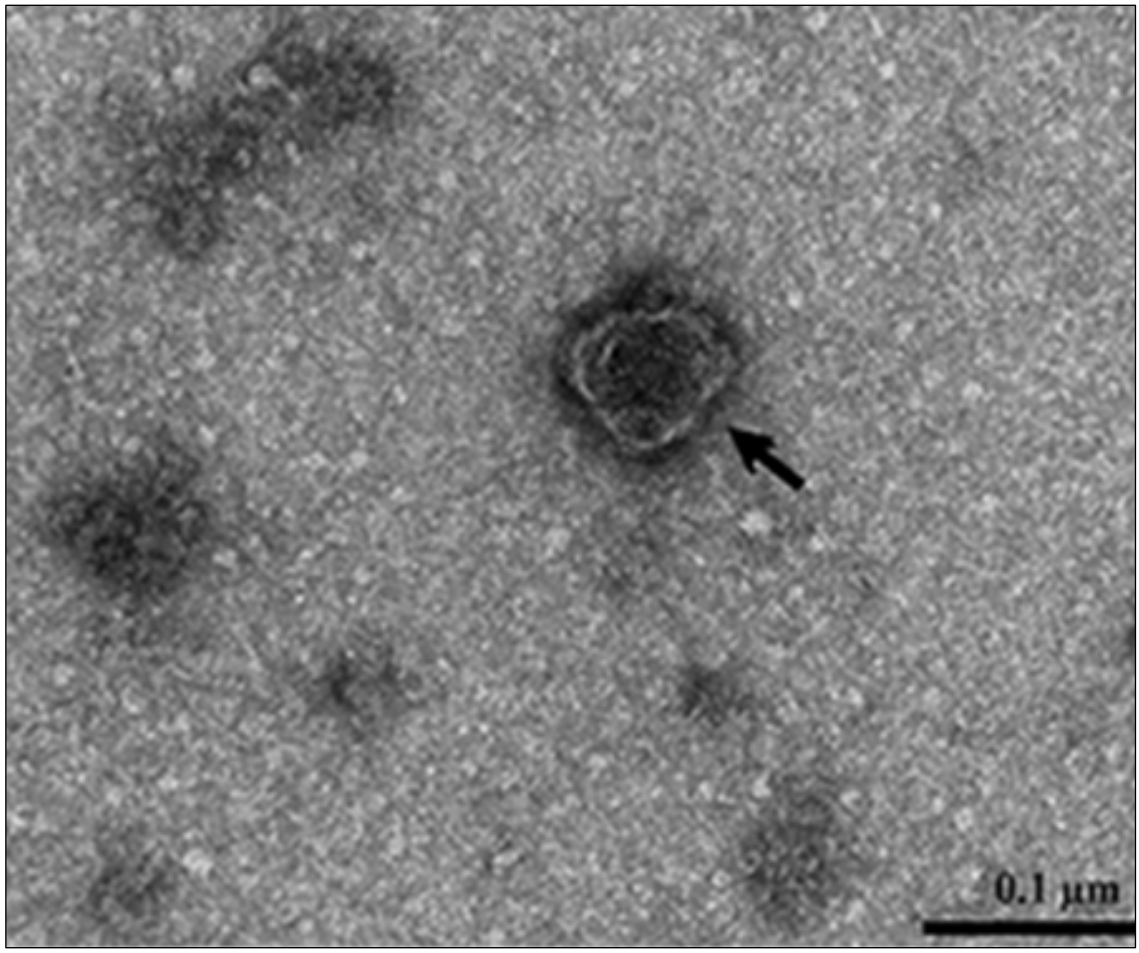


Figure 1. Transmission electron micrograph of TiLV infected E-11 cells. High magnification of a free virion showing a round enveloped viral particle with 60 to 80 nm diameter (Tattiyapong, Dachavichitlead, and Surachetpong 2017).

4.1.3 Clinical symptoms:

Many of TiLV's clinical signs are similar to other viral infections in tilapia and subclinical infection of TiLV has also been reported, creating challenges for early diagnosis. Mortality in tilapia populations with corneal opacity can be considered one of the common signs of this disease (Aich et al. 2022). Other clinical symptoms include anorexia, poor body condition, abnormal swimming, bilateral exophthalmia, congestion, scale protrusion, severe anaemia, skin erosion, pale

coloration of gills, and swollen abdomen (Eyngor et al. 2014; Tattiyapong, Dachavichitlead, and Surachetpong 2017). The most common microscopic lesions associated with TiLV infections include hepatitis and encephalitis lesions (Al-Hussinee et al. 2018). In situ hybridization indicates TiLV replication and transcription at sites of pathology in the liver and central nervous system of diseased tilapia (Bacharach et al. 2016).

4.1.4 Transmission

TiLV viral particles are detected in the reproductive organs, serum, and eggs of Tilapia (fertilised and unfertilised). An intragastric route is predicted to be the prime route of infection, as the intra peritoneal route needs to pass the first line of defence before entry into the body (Aich et al. 2022). The cohabitation mode of transmission described by Eyngor and et al, demonstrates the ability of TiLV to spread via the water environment. Relatively high mortality rates have been observed for both the intraperitoneal and waterborne routes. Fish surviving initial mortality events however have been shown to then be immune to further TiLV infections, suggesting the mounting of an adaptive immune response (Eyngor et al. 2014).

TiLV disease is highly contagious and spreads through both horizontal and vertical transmission. Adult tilapia may have asymptomatic infections and act as asymptomatic carriers that pass the virus to their offspring (Aich et al. 2022). Vertical transfer has been demonstrated by Dong et al, suggesting that TiLV causes systemic infection in tilapia broodstock, with the virus able to spread into the reproductive organs. Subsequently, the fertilised eggs from infected broodstock tested positive for TiLV (Ha Thanh Dong et al. 2020). Rapid and accurate detection of this virus is crucial for selection of fish broodstock, in view of the vertical transmission possibilities from parents to offspring (Taengphu et al. 2020). This is particularly important for farmers importing tilapia seed (eggs/embryos) from TiLV reported countries, increasing the risk of cross-country and potential cross-continents spread of the virus (Aich et al. 2022).

4.1.5 Diagnostic tools:

TiLV has been successfully isolated and propagated using a variety of cell lines, and electron microscopy has been used to directly demonstrate the presence of TiLV virions or nucleic acid. Despite the strength of using cell culture methods as a primary diagnostic tool for TiLV, it is laborious, time consuming and requires specialist training and facilities. Conventional PCR assays require post-PCR processing steps, such as gel electrophoresis and PCR product purification, and are not as sensitive and specific as RT-qPCR approaches (Waiyamitra et al. 2018). Several sensitive and rapid molecular diagnostic tools have been published for early detection of the virus, including reverse transcriptase polymerase chain reaction (RT-PCR), reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) (Tsofack et al. 2017), and loop-mediated isothermal amplification (LAMP) (Yin et al. 2019; Aich et al. 2022).

RT-qPCR for the detection of viruses is advantageous because of its quantitative nature, high sensitivity, specificity, scalability and its rapid time to result (Waiyamitra et al. 2018; Aich et al. 2022) (Aich et al. 2022). Various nested and semi-nested RT-PCR assays have been developed focusing on segment 3 of TiLV, and this is becoming the most widely sequenced segment due to its use for detection of the virus. To avoid amplification of fish genes, Dong et al proposed a semi-nested PCR, which was a modification of the nested protocol of Kembou Tsofack et al. (H. T. Dong, Siriroob, et al. 2017; Tsofack et al. 2017; Waiyamitra et al. 2018; Chaput et al. 2020). In addition, Waiyamitra et al created a sensitive and specific TaqMan probe-based RT-qPCR assay targeting segment 3 of TiLV, for the detection of TiLV in field samples (Waiyamitra et al. 2018). Other studies have developed new semi-nested RT-PCR methods by designing primers from highly conserved regions of TiLV genome segment 1 for disease diagnosis and surveillance (Taengphu et al. 2020).

Virus genomics have been used to investigate infectious disease outbreaks for decades, with the high rates of mutation and replication creating novel variants across short timescales (Grubaugh, Ladner, et al. 2019). A total of 548 genome sequences of TiLV segments have been reported, and analyses of these sequences have led to a better understanding of how TiLV evolves and spreads

across regions (Thawornwattana et al. 2021). Nucleotide substitution rate per site per year for TiLV is comparable with other RNA viruses (Verma et al. 2022). RNA viruses undergo rapid evolutionary changes due to the absence of proofreading in their RNA polymerases (Steinhauer, Domingo, and Holland 1992). Phylogenetic analysis can reconstruct chains of transmission, and evidence of TiLV's global spread has been based on phylogenetic analysis of short sequences from a single segment.

As a segmented virus, TiLV is capable of undergoing reassortment, where multiple strains of viruses with segmented genomes co-infect the same host cell and exchange their genetic materials (Chaput et al. 2020; Thawornwattana et al. 2021). Previous studies with limited sequence data have stated the absence of reassortment in this virus. However, most of the TiLV reported sequences are partial genomes, and analysis of individual genomic segments may limit interpretation of how TiLV evolves (Thawornwattana et al. 2021). Chaput *et al* performed phylogenetic analysis of the ten segment coding regions of TiLV collected from Bangladesh, placing the circulating strain in a clade with two isolates from Thailand, separate from the Israeli and South American isolates. Phylogenetic analysis of individual segments gave conflicting results, sometimes clustering the Bangladesh strain with one of the Israeli isolates, and splitting pairs of isolates from the same region. This suggests that the predicted phylogeny of TiLV isolates depend on the segment sequenced, and that reassortment is common in TiLV (Chaput et al. 2020), and is the dominant force for its evolution (Verma et al. 2022). Due to the inherent nature of segmented viruses, it is impossible to generate a single, complete viral genome with few small overlapping PCR amplified regions (Delamare-Debouteville et al. 2023). Additionally, conventional NGS techniques cannot determine the 5′ and 3′ terminal sequences of the RNA viral genome. Therefore, in whole genome sequencing approaches, there is an additional loss of unsequenced regions, and this is proportional to the number of segments (Misu et al. 2023). Tools, such as the Primal Scheme software, which are used to generate primers for recovering full viral genomes, have yet to be adapted for segmented viruses.

Given the significant impact of TiLV on the tilapia aquaculture industry, there is a critical need for more robust genomic surveillance to facilitate better management

and tracking of this virus (Delamare-Deboutteville et al. 2023). Here, a tiled PCR for WGS was developed for TiLV using samples from two different countries for a phylogeographic study. To our knowledge, this is the first attempt to use the Artic pipeline for a segmented virus, with tiles crossing individual segments, by creating two separate pools. This tool enables real-time tracking of TiLV to improve genomic surveillance and control, especially in remote regions.

4.2 Methods and Materials

4.2.a TiLV Samples from Bangladesh:

1. *Samples and total RNA extraction*

As an initial development and testing for the tiled PCR method for TiLV, we selected samples collected from Bangladesh, from one affected farm. Samples were collected by members of our lab group in July 2017, in response to reports of high tilapia mortality in a village in Trishal Upazila, Mymensingh District (Chaput et al. 2020). The farmer reported that over the previous 20 days, 15 tonnes of tilapia had been lost across a 28-hectare farm. Following sample collection, fish were terminated via Schedule 1 process, and dissected on-site. Samples were stored in RNAlater (Ambion Inc., Austin TX, USA) and kept at ambient temperature until arrival in the UK, where they were stored at −20°C until processing. Tissue Samples were processed by RNA extraction from < 20 mg fish tissues (Fish1 R2-heart, Fish1 R3-liver, Table 1), using the RNeasy Mini kit (Qiagen, Venlo, Netherlands), following the manufacturer's protocol for RNAlater-fixed animal tissues. RNA was eluted in 50 µL RNase-free water, quantified by spectrophotometry on a NanoDrop ND-1000 (NanoDrop Technologies Inc, Wilmington DE, USA) and stored at −80°C.

Table 1. List of infected TiLV fish, collected from Bangladesh (Chaput et al. 2020); organs used for RNA extraction, weight and total RNA concentration are listed for each sample.

Fish farm	*Sample*	*Organ*	*Tissue (mg)*	*RNA conc (ng/µL)*	*A260/280 ratio*	*A260/230 ratio*
F1	F1R2	heart	6	286.6	2.14	1.93
F1	F1R3	liver	8	944.0	2.15	2.09

2. *Reverse transcription*

To prepare cDNA from the TiLV RNA samples, 1 µg of RNA was initially treated with RQ1 RNase-Free DNase (Promega) in a total volume of 10 µL for each reaction, and incubated for 30 min at 37°C. DNase was inactivated with 1 µL of RQ1 Stop Solution and a 10 minute incubation at 65°C. Reverse transcription of TiLV RNA was carried out with M-MLV reverse transcriptase (Promega, Madison WI, USA), by adding 1 µl of random hexamers to each reaction. This was followed by a 5 min incubation at 70°C (melts secondary structures within the template), with immediate cooling on ice for 2 min (prevents secondary structures from reforming). A mastermix was created by adding 5 µl M-MLV 5x reaction buffer, 2 µl dNTP mix (10 mM), 5 µl water, and 1 µl M-MLV reverse transcriptase (@ 200 units/µl). This was added to the treated sample, with gentle mixing followed by brief centrifugation. Samples were incubated for 60 min at 37°C and the reaction was inactivated by heating for 10 min at 70°C.

3. *Primer design for a tiled PCR for TiLV:*

Primers were designed to produce amplicons that span the full TiLV genome. PrimalScheme (v 1.3.2) was used to produce primers for each segment individually, using a multiple sequence alignment generated by (Chaput et al. 2020). This alignment was generated from six complete or near-complete TiLV genomes that were publicly available from NCBI, listed in Table 2. Genomes were used to construct alignments of the full coding region of each segment, missing the segment termini. Primers were designed to target 325 bp amplicons for each

TiLV segment, except segment 7 (330 nt amplicons), seg 9 (352 nt), and seg 10 (271 nt). Size selection was adjusted according to the software, as it failed to produce the exact size selection for each segment. This produced 68 individual primers (34 primer pairs) spanning the full TiLV genome (Supplementary Table 1). A schematic diagram of the primers' locations on the TiLV reference genome are also shown in Figure 2.

Table 2. List of sequences used for generating a sequence alignment to produce tiled PCR primers; the NCBI number and the countries they were collected from are listed.

Sample ID	*Accession no.*	*Country*	*Reference*
TIL-4-2011	KU751814-82	Israel	(Bacharach et al. 2016)
AD-2016	KU552131-142	Israel	(Bacharach et al. 2016)
TV1	KX631921-936	Thailand	(Surachetpong et al. 2017)
WVL18053-01A	MH319378-387	Thailand	(Al-Hussinee et al. 2018)
EC-2012	MK392372-381	Ecuador	(Subramaniam et al. 2019)
F3-4	MK425010-019	Peru	(Pulido et al. 2019)

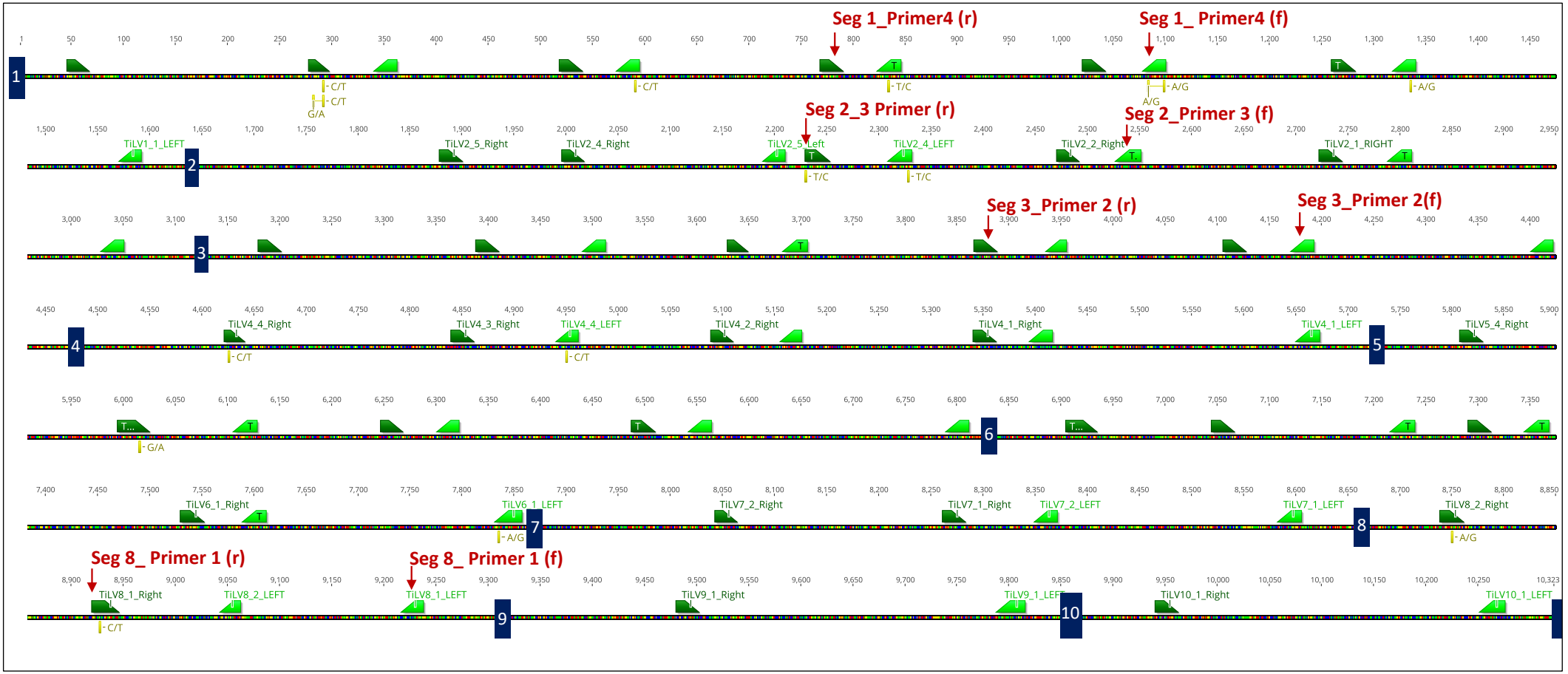


Figure 2. A schematic diagram of the location of the multiplex primers for TiLV shown on the full reference genome. Primers were generated for all 10 segments using the Primal Scheme software and are visualised in Geneious Prime (*v. 2022.1.1)*. Failing primer pairs are listed in red, while the segments are separated by blue rectangles across the genome.

Reversed transcribed TiLV was used to individually test each primer by following the nCoV-2019 sequencing protocol v3: https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye?step=6 (accessed on 29 August 2020). Primers were prepared at 10 µM working stocks, and two primer pools were generated (Pool A and Pool B), containing odd and even numbered primer pairs, respectively, at 15 nM concentration per primer. Individual primer pairs were tested using a PCR reaction of 12.5 µL of 2x NEBNext Mastermix (NEB); 1.28 µL of each 10 µM pool (final conc. of each primer 15 nM) and 3 µL of cDNA. Finally, 5.5 µL of NFW was added for a total reaction of 25 µL. PCR conditions were: 98°C for an initial heat activation for 30 s, 15 s at 98°C for denaturation, followed by a 65°C for annealing and extension step for 5 min for 30 cycles. Amplicons were visualised by gel electrophoresis, using a 1.5% agarose, at 100 V for 50 min. Four out of 34 primer pairs failed to produce a product of appropriate size. The Agilent TapeStation system was used to confirm the failing of the four primer pairs, with primer 2 segment eight as a positive control.

Tiled PCR was performed for pools A and B for each sample, using 0.5 µL of Q5 Hotstart High-Fidelity Polymerase (NEB), 5 µL Q5 Reaction buffer (NEB), 0.5 µL of dNTP (10µM), and a final concentration of primer A or B at 15 nM for each primer. A total of 2.5 µL of cDNA was added to each pool, and NFW was added for a total reaction mix of 25 µL. PCR conditions were the same as above. The concentration of the amplicons of each pool are listed in Table 3.

4. *Library Preparation and Sequencing:*

Generated amplicons for pools A & B for each sample were quantified using a Qubit dsDNA broad range kit (Invitrogen, Waltham, MA, USA), combined and assigned a barcode for multiplexing. Library preparation was carried out following the native barcoding of amplicons with EXP NBD104 and SQK-LSK109 protocol: version *NBA_9093_v109_revD_12Nov2019*. Equimolar amounts of each barcoded sample were pooled and taken forward for the adaptor ligation step using a total volume of 60 µL of DNA. An amount of 5 µL of Adapter Mix II (AMII), 25 µL of Ligation Buffer (LNB) and 10 µL of T4 DNA Ligase were all added to the barcoded DNA. The reaction was incubated for 10 min at room temperature, and a 0.5× AMPure XP bead clean-up was performed, followed by 2 × 250 µL of Short

fragmented buffer (SFB) (ONT) washes. The pellet was resuspended in 15 µL of Elution Buffer (EB) (ONT) for 10 min at 37°C, and quantified (Table 3).

Table 3. Concentration of TiLV following tiled PCR and library preparation for sequencing of samples collected from Bangladesh.

Fish Farm	*Sample*	*Barcode*	*Pool*	*Amplicon conc. (ng/µL)*	*Final Library (ng/µL) (A & B)*
F1	F1R2	10	A	55.6	41.8
			B	46	
F1	F1R3	11	A	29.6	25
			B	40.8	
Neg	-	12	-	Too low	Too low

A total of 585.2 ng of the adapter ligated cDNA was loaded onto a MinION flow cell with 785 detected pores. Sequencing was performed for 72 h, with real-time super high accuracy base calling using the Oxford Nanopore Guppy tool (v. 6.5.7). Adapter trimming was performed, and samples were demultiplexed using guppy_barcoder. Reads below 200 bp in length and above 450 bp were removed. A complete TiLV reference genome was generated from concatenating each reference genome segment from NCBI (Table 4), and amplicons were mapped against it using minimap2 (v.2.17). Genome coverage was visualised in Tablet (v. 1.21.02.08).

Table 4. List of the concatenated reference genomes for TiLV, to generate a full reference genome. Size of each segment and their reference number from NCBI are listed.

Segment	*Size*	*Reference*	*Location on genome*
Segment 1	1641	NC_029926.1	1-1641
Segment 2	1471	NC_029921	1642-3112
Segment 3	1371	NC_029927.1	3112-4483
Segment 4	1250	NC_029922.1	4484-5733
Segment 5	1099	NC_029923.1	5734-6832
Segment 6	1044	NC_029928.1	6833-7876
Segment 7	777	NC_029924.1	7877-8653
Segment 8	657	NC_029929.1	8654-9310
Segment 9	548	NC_029925.1	9311-9858
Segment 10	465	NC_029930.1	9860-10323

4.2.b TiLV Samples from Thailand

1. Samples

Tilapia tissue samples of the liver and brain were collected in December 2022, from two different farms in Thailand (Table 5, Figure 3 & 4). Tissue samples were preserved in RNALater and kept at -20°C and shipped to the University of Exeter on dry ice.

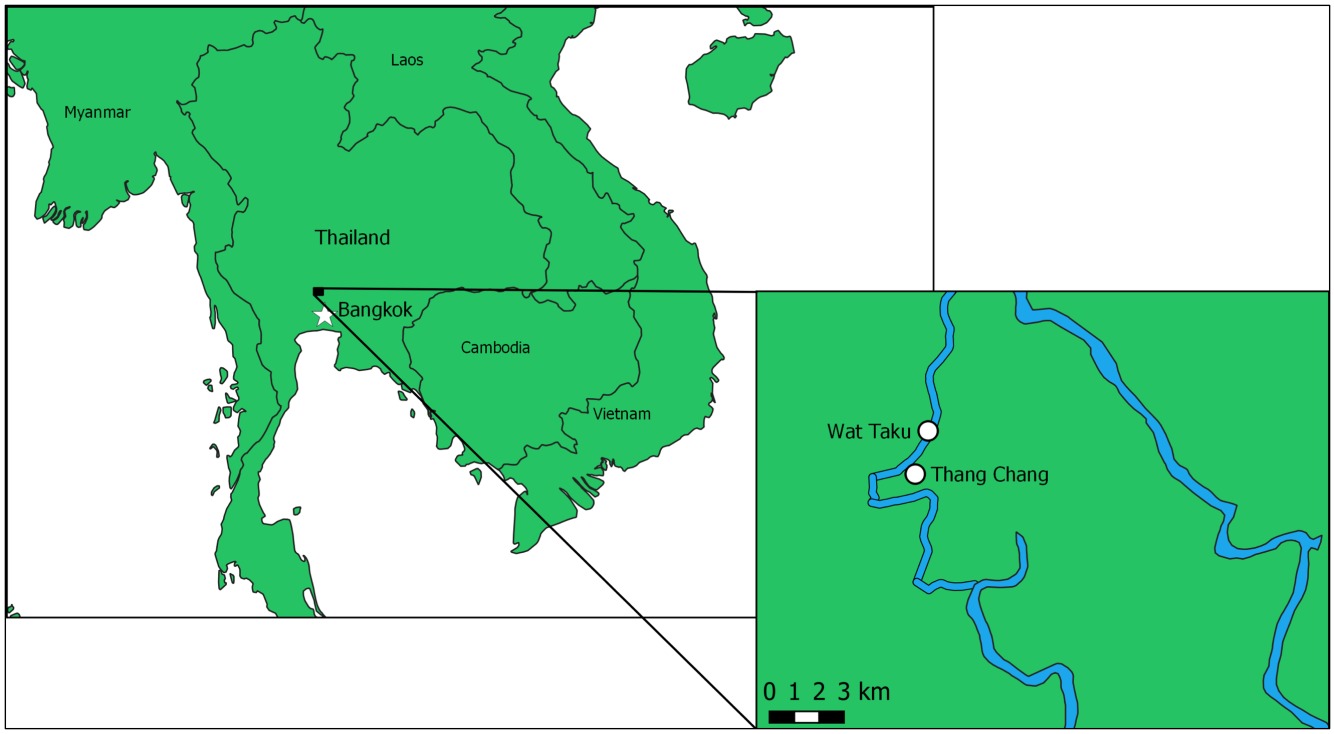


Figure 3. Location of tilapia fish farms in Thailand. Samples collected from two different farms in shown in white circles. Photo credit: Stephanie Andrews.

Table 5. Location of fish farms in Thailand, fish stage and clinical symptoms of sampled fish.

Farm	*Fish ID*	*Pond/ Case*	*Region*	*Sample details*	*Species*	*Clinical Sign*
1	*ID1*	Pond	Thang Chang, Bang ban, Ayuthaya, Thailand	Nursing pond, fish aged 30-40 days	Red Tilapia	Brain necrosis, tail rot, eye shrinkage and hemorrhagic opercula
	ID2	Pond				
	ID3	Pond				
2	*ID4*	Case	Wat Taku, Bang Ban, Ayuthaya, Thailand	Grow out, Fish aged 45-60 days	Red Tilapia	Scale protrusion, liver pallor, excess fluid in abdominal cavity, swollen gallbladder with 2-3% daily mortality
	ID5	Case				
	ID6	Case				
	ID7	Case				
	ID8	Case				

a)

Figure 4. a). Fish farms and sampled fish from Thailand, showing clinical symptoms. Farm 1: Thang Chang region; ponds (1, 2); fish showing clinical symptoms of eye shrinkage (3), and tail rot (4).

b)

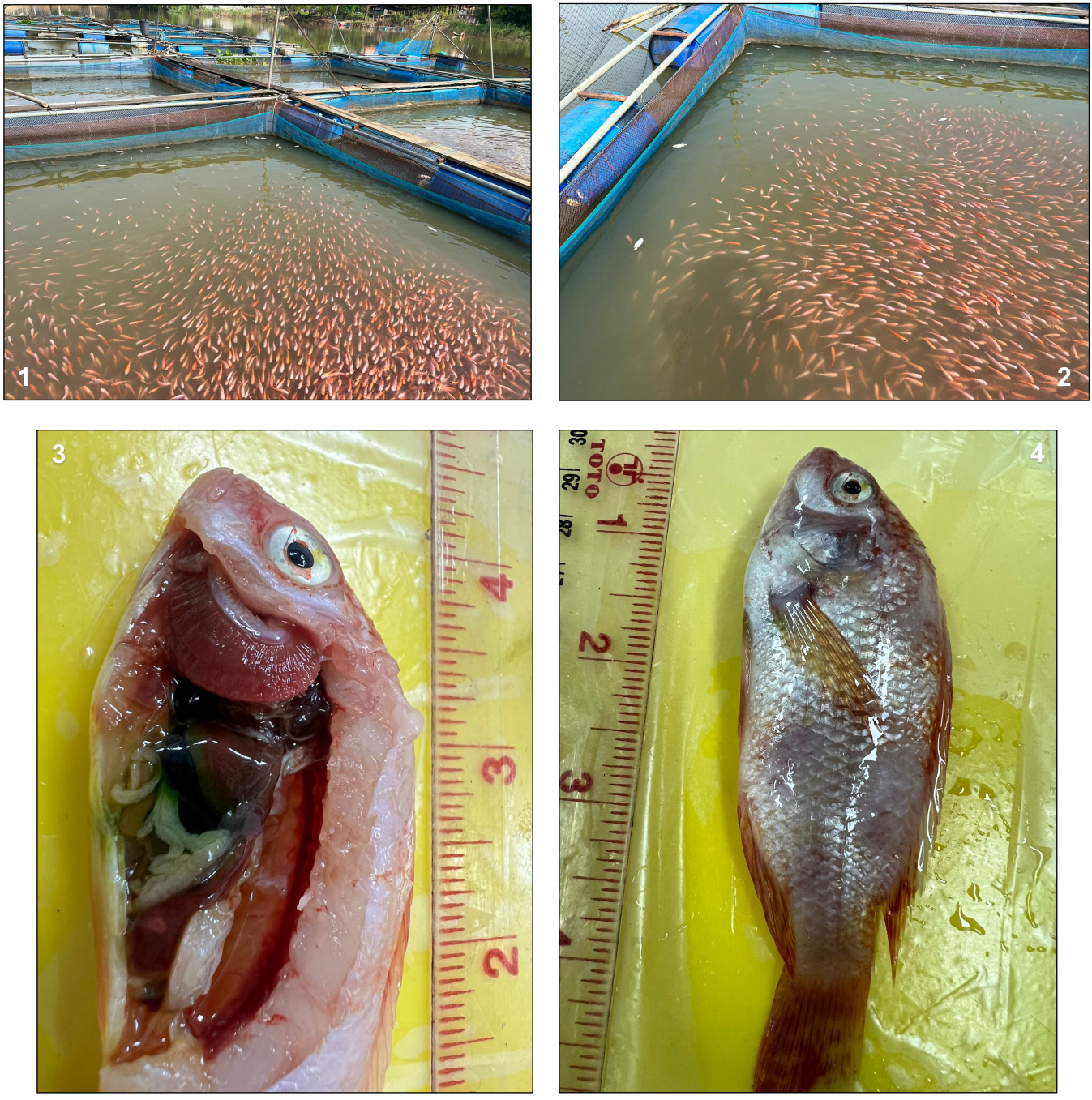


Figure 4. b). Fish farms and sampled fish from Thailand, showing clinical symptoms. Farm 2: Wat Taku region; case in river (1,2); clinical symptoms are shown as: excess fluid (3), and scale protrusion (4); (Photo credit: Partho Debnath).

5. *RNA extraction and reverse transcription*

For lysis and homogenisation of samples, a TissueLyser II instrument (Qiagen, Venlo, Netherlands) was used for an approximate of 10 gm of tissue (liver and brain) and was carried out in a 2 mL tubes with a single 5 mm diameter steel bead for 2 x 2 min at 30 Hz. A proteinase K step was necessary for successful RNA extraction. Following lysis in 600 µL buffer, 400 µL lysate was mixed with 787 µL RNase-free water and 13 µL proteinase K solution (>600 mAU/mL, Qiagen, Venlo, Netherlands), then incubated at 55°C for 10 minutes. Following centrifugation at 10000 x g for 3 minutes, supernatant was moved to a clean tube, mixed with 0.5 vols 100% molecular grade EtOH (Thermo Fisher Scientific, Inc., Waltham MA, USA), and loaded onto a micro-RNeasy spin column, with the subsequent washing and elution steps as described in the RNeasy Micro protocol. Samples were kept at -80 °C until further processing. Luna script (NEB) was used to convert 1 µg of total RNA to cDNA, and the samples collected from this outbreak were quantified using a NanoDrop.

6. *TiLV Semi-nested PCR protocol:*

A semi-nested PCR was performed using primers designed from highly conserved regions of TiLV segment 1: TiLV/nSeg1F; 5′-TCTGATCTATAGTGTCTGGGCC-3′ and TiLV/nSeg1R; 5′-AGTCATGCTCGCTTACATGGT-3′ (Taengphu et al. 2020). The expected amplified product was 620 bp. Primers TiLV/nSeg1F and TiLV/nSeg1R; 5′- CCA CTT GTG ACT CTG AAA CAG −3′ with an expected product of 274 bp were employed in the second round PCR. The first RT-PCR reaction of 25 µL composed of 100 ng of RNA template, 400 nM of each primer, 0.5 µl of SuperScript III RT/Platinum Taq Mix (Invitrogen), and 1× of supplied buffer. Amplification profiles consisted of a reverse transcription step at 50 °C for 30 min; a denaturation step at 94 °C for 2 min, 30 PCR cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s; and a final extension step at 72 °C for 2 min. 5 µl of product from the first round PCR was then used as template in the second round PCR reaction of 25 µl containing 500 nM of primer TiLV/nSeg1F, 600 nM of primer TiLV/nSeg1RN, 0.16 mM of each dNTP, 0.8 mM MgCl2, 1 unit of Platinum Taq DNA polymerase (Invitrogen), and 1.2× supplied buffer. Thermocycling conditions consisted of a 5 min initial denaturation step at 94 °C followed by 30

cycles and a final extension step described above. Finally, 10 µL of the amplified products were analysed by 1.5% agarose gel electrophoresis stained with RedSafe DNA staining dye.

7. *Tiled PCR, Library Preparation and Sequencing:*

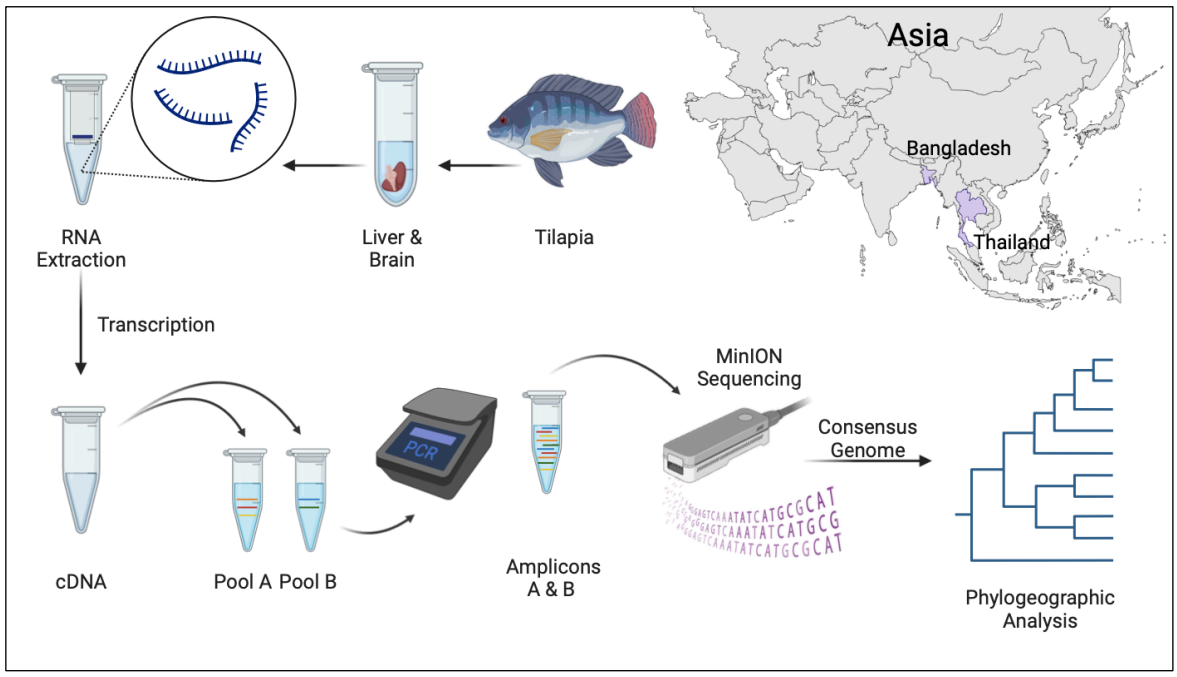
To generate amplicons that span the full genome, each sample was added to pools A & B for a tiled PCR, using the same primers mentioned previously, in Supplementary Table 1.

Equimolar amounts of generated amplicons for pools A & B for each sample were quantified using a Qubit dsDNA broad range kit, shown in Table 6, and combined, with each sample assigned to a single barcode. Library preparation was carried out following the native barcoding of amplicons with EXP NBD104 and SQK-LSK109 protocol, as mentioned previously. A total of 15 µL of the elute was retained and adaptor ligated DNA was taken forward for priming and loading onto the flow cell.

Initially, to test the success of the library preparation, samples were loaded on a flongle, where 68.6 ng were loaded. This flongle had a low number of pores (60 pores). Subsequently, 1 µg of adaptor ligated cDNA library was loaded onto a R9.4 MinION flowcell for better results. This flow cell had 1416 pores and the sequencing run lasted 72 hours. Reads were basecalled with super high accuracy base calling, using the Oxford Nanopore Guppy tool (v. 6.5.7). Adapter trimming was performed, and samples were demultiplexed using `guppy_barcoder`. Read demultiplexing was performed by requiring barcodes at both ends, and reads below 200 bp and above 600 bp were removed. Reads were mapped to the full TiLV reference genome from the NCBI (NC_029926.1) using minimap2 (v.2.17). The bam file was visualised in Tablet. The Artic pipeline was then evaluated to produce a consensus genome for all segments for TiLV. The locations of the primers were added manually to the files produced in the Primal Scheme software. The sequences of all the left primers were reverse transcribed to match the reference genome. The Artic pipeline produced a sorted bam file, using minimap2 (v.2.17), but failed to produce a consensus sequence. The bam file was subsequently transferred to Geneious Prime to create a consensus genome. Genome loci with coverage <20× were identified using

`samtools depth` to calculate per-locus coverage followed by parsing with a python script.

A Schematic Diagram of the Workflow for Sample Collection and Processing. Figure was generated with BioRender (https://biorender.com/).



4.3 Results

4.3.a TilV Samples from Bangladesh

1. Testing individual primer pairs:

Individual primers were tested using the samples collected from Bangladesh. Following amplification, 30 out of 34 primers generated a faint band at the expected location. The failing primer pairs were primer pair 4 of segment 1, 3 of segment 2, 2 of segment 3 and 1 from segment 8, where two of these were located near the end regions. Gel images are not provided as they were faint, and bands were quickly lost due to exposure to UV light. The four primer pairs that failed to show a band using gel electrophoresis, were revaluated using the TapeStation for each primer pair, and primer pair 2 of segment 8 was used as a positive control, shown in figure 5. This showed a faint band for primer 1 of segment 8 only, and an obvious band for the positive control at the expected regions.

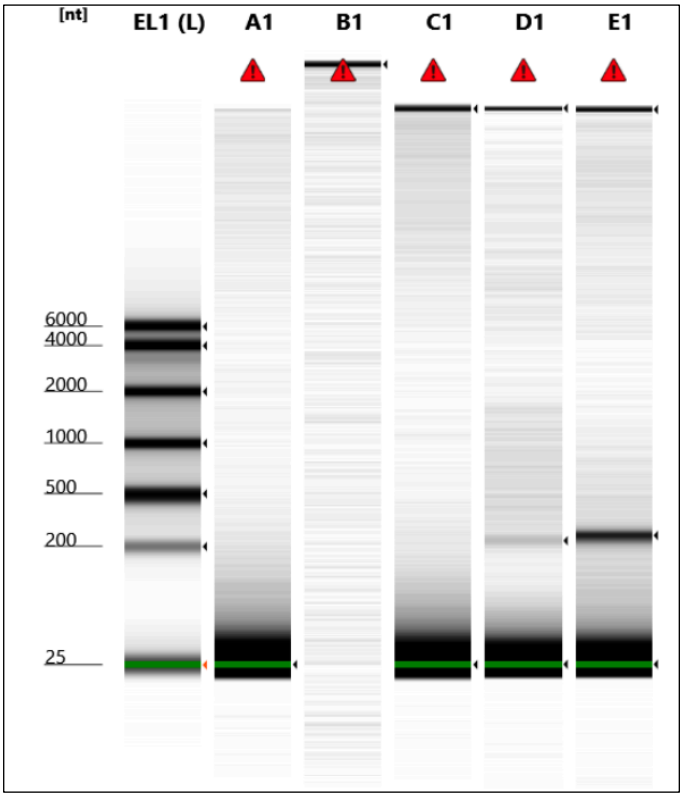


Figure 5. TapeStation image for TiLV RNA individual primer pairs failing to amplify. Primer pairs 1of segment 4, 3 of segment 2, 2 of segment 3 and 1 of segment 8; 2 of segment 8 was used as a positive control to the right.

2. *Tiled PCR for TiLV*

A tiled PCR was performed using primer pool A and pool B, and bands were seen with gel electrophoresis at the expected sizes (271 nt - 359 nt).

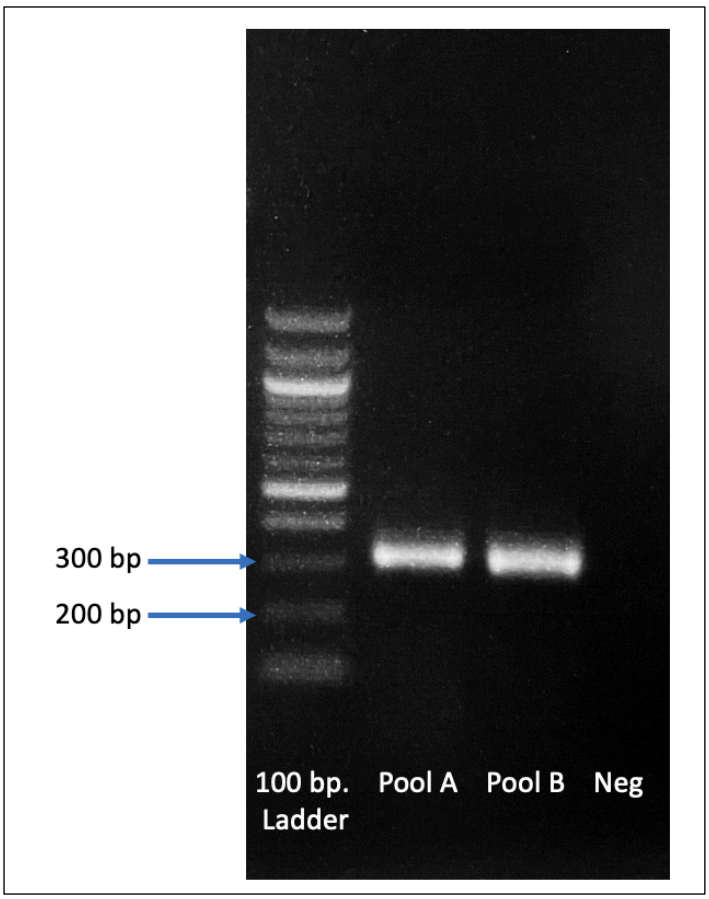


Figure 6. Gel electrophoresis image of TiLV tiled amplicons of sample R2 collected from Bangladesh; size of amplicons at the expected region, between 271-359 bp. First lane: 100 bp ladder.

3. *Sequencing Results:*

A total of 2,745,146 reads were produced. Following basecalling, trimming was performed, and samples were demultiplexed. The number of reads following trimming were 641,985 and 712,043 for F1R2 and F1R3, respectively. Reads were mapped to the concatenated TiLV reference genome (Table 4), using minimap2 (v.2.17). Genome coverage was visualised in Tablet (Figure 7), and regions without coverage (60 -120 nt) were seen at the ends of each segment.

Figure 7. TiLV genome visualised in tablet, showing reads aligned to the TiLV concatenated reference genome for all segments.

4.3.b TiLV Samples from Thailand

Three farms in Thailand were tested for TiLV, TiPV, and ISKNV, as part of their testing routine. Two of the three farms tested positive for TiLV (Figure 8), while being negative for TiPV and ISKNV. Following RNA extraction, concentration of total RNA for each sample was measured on a NanoDrop (Table 6), before a tiled PCR was performed.

Table 6. Total RNA concentration of TiLV, cDNA template added for the tiled PCR for each pool and the concentration of the final library prepared for each sample.

Fish no.	*Total RNA ng/µL*	*Template Added µL*	*Barcode*	*Library prepared cDNA conc. (ng/µL)*
ID1	222.3	5	1	9.76
ID2	652.3	1.53	2	14.7
ID3	515	1.9	3	Too low
ID4	2974.6	1	4	21.8
ID5	279.9	3.57	5	16.1
ID6	903.1	1	6	25.8
ID7	1889.5	1	7	20.4
ID8	388.9	2.57	8	8.54
Neg.	-	5	9	Too low

A semi-nested PCR was performed on all samples collected from three farms in Thailand. TiLV was detected in farms 1 & 2, where gel electrophoresis showed clear bands at the expected sizes for all samples (Figure 8). However, sample ID3 showed only one band for the nested PCR, missing the band generated from the first round of the nested PCR.

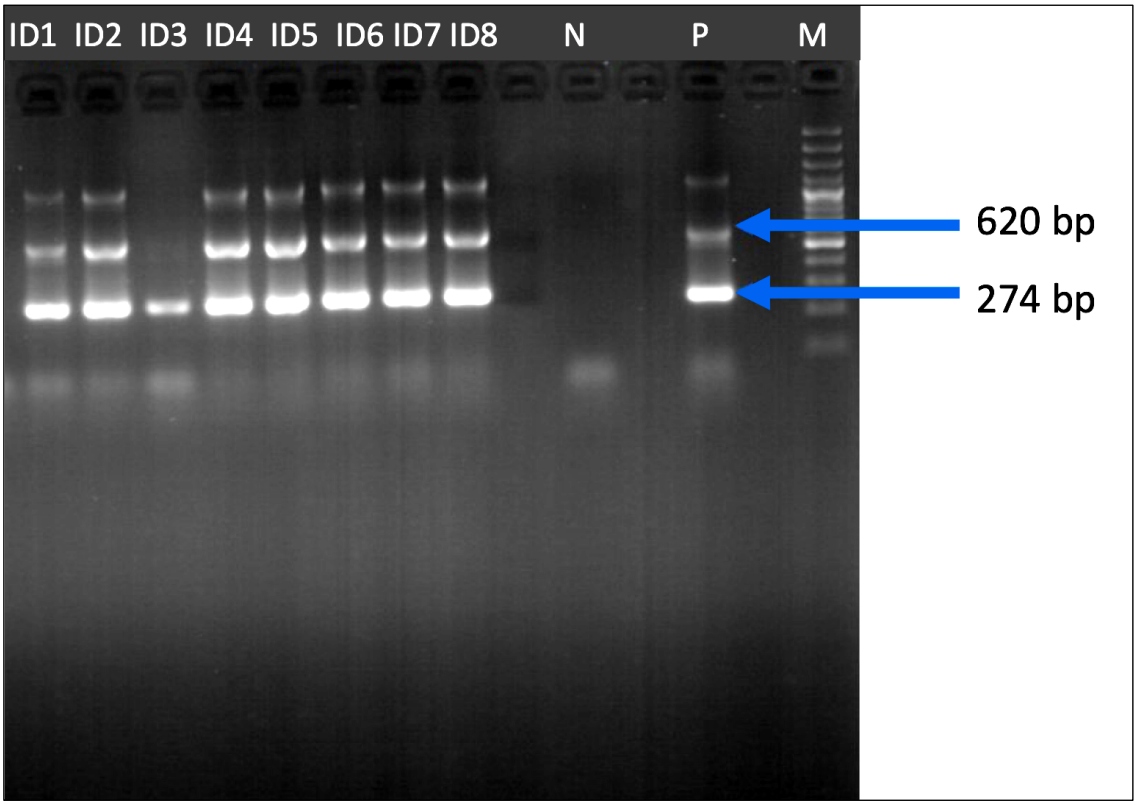


Figure 8. Gel electrophoresis showing amplicons generated for the detection of TiLV, from samples collected from Thailand. Samples (ID1 - ID3) for farm 1; samples (ID4 - ID8) for farm 2; and samples (1-8) for a sampled farm that where TiLV was under the limit of detection. (N) for negative control; and (P) for positive control, and (M) for the ladder.

Sequencing Results:

A total of 12.37 million reads were produced using the MinION flowcell R9.4. Following basecalling and demultiplexing, the mean and median length for each sample were calculated (Table 7). Sample ID3 showed to have the lowest number of reads, confirming low detection of TiLV by semi-nested PCR.

Table 7. Long read sequencing results of TiLV samples collected from Thailand; showing the number of reads, median length, and the number of reads after trimming for each sample.

Fish no.	*Barcode*	*No._reads*	*Median_length (bp)*	*Reads 200-600bp*	*% Genome covered > 20x*
ID1	BC01	647891	448	504785	75.1
ID2	BC02	469845	451	339774	78.2
ID3	BC03	99664	445	92037	69.5
ID4	BC04	514615	450	444993	82.1
ID5	BC05	298187	456	189277	77.9
ID6	BC06	353454	452	271739	74.9
ID7	BC07	321543	453	214717	78.2
ID8	BC08	500906	451	359338	64.0

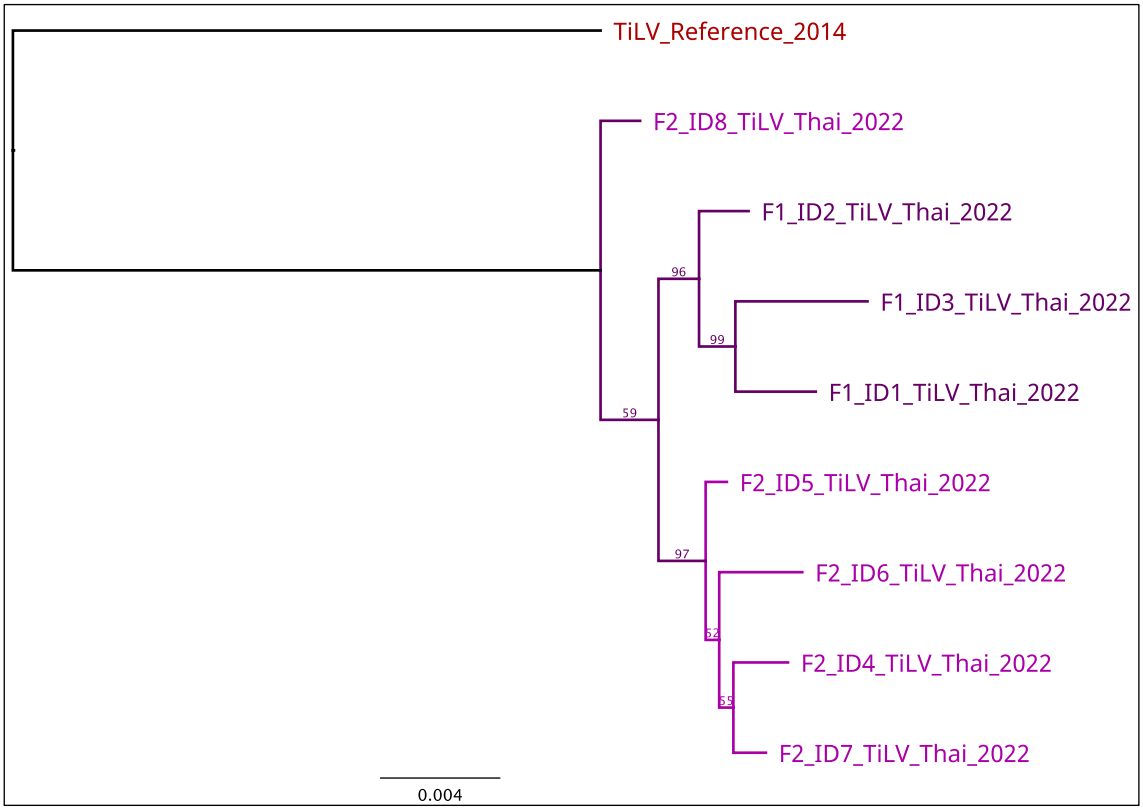
The Artic pipeline was evaluated to produce a consensus genome for all segments for TiLV. In comparison to previous work with ISKNV, the pipeline required significant manual editing to run successfully. Locations of the primers were manually adjusted and sequences of all left primers needed to be reverse transcribed to match the reference genome. Even with these edits, the Artic pipeline produced a sorted bam file, using Minimap2, but failed to produce a consensus sequence. The bam file was used to generate a consensus genome in Geneious Prime. The Artic pipeline in brief, subsamples the reads to 200x, trims the primers, and generates a consensus sequence, calling Ns at regions with less than 20x read depth. On the other hand, Geneious Prime software uses all the provided reads to generate the consensus genome. Consequently, consensus sequences from Geneious Prime were manually curated to identify regions with <20x coverage.

The consensus genomes from both farms from Thailand were aligned to the concatenated reference full genome using Mafft alignment (Katoh et al. 2002). A

Neighbour Joining consensus tree was generated in Geneious Prime, using Geneious Tree Builder, with the Tamura-Nei substitution model with 100 bootstrap branch support.

Farms 1 and 2 showed a clear separation on the phylogenetic tree, forming two clades clustering according to farm, except for sample ID8 from farm 2, which clustered separately (Figure 9 (a)). Sample ID8 was removed when generating the phylogenetic tree in Figure 9 (b), showing two clades clustering according to farm, without affecting the placement of other samples. All samples were rooted to the concatenated TiLV reference genome.

a)



b)

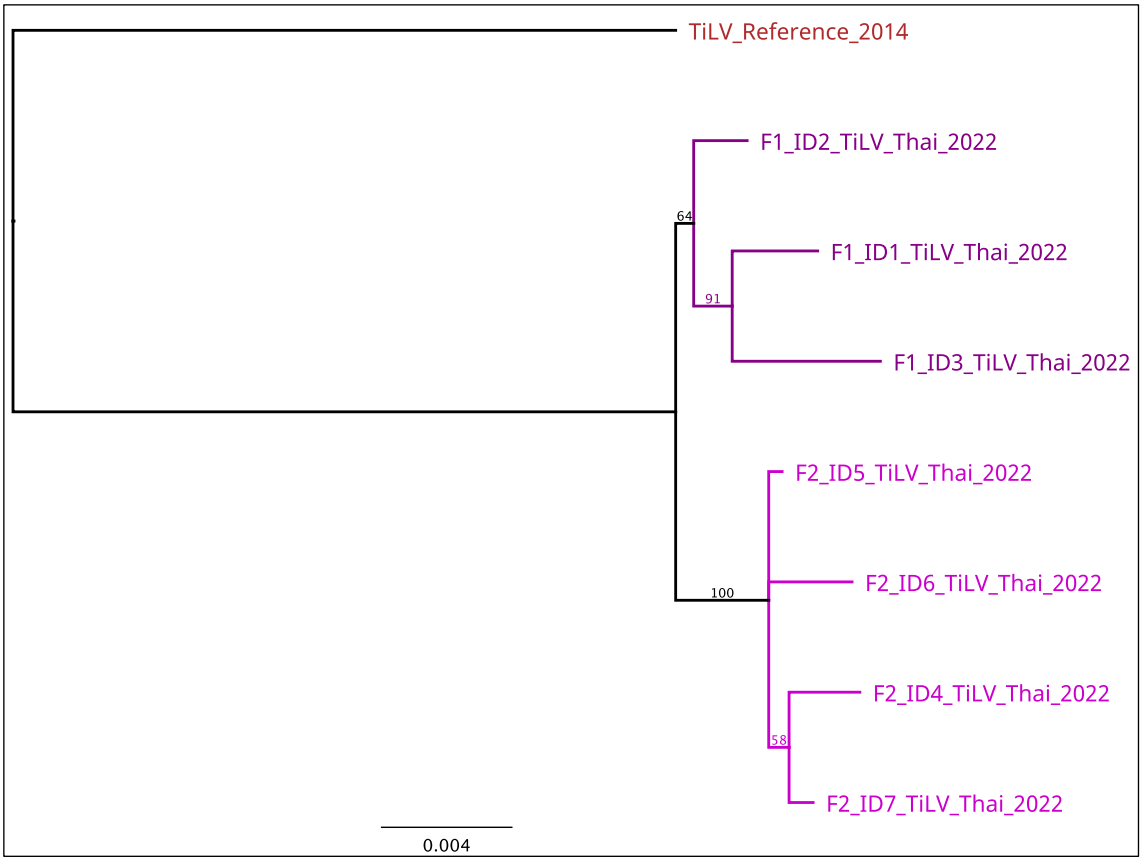


Figure 9. A phylogenetic tree generated in Geneious for the full TiLV genome for samples collected from two farms in Thailand. Two clades are formed, Farm 1 & 2 are in dark purple and light purple, respectively. a) All genomes were aligned to the TiLV concatenated reference genome; b) all genomes except sample (ID8). The scale bar represents the number of substitutions per site.

Additionally, the consensus genomes for TiLV samples collected from Bangladesh, were generated in a similar way as mentioned previously and aligned to TiLV samples from Thailand. Whole genome TiLV samples from Bangladesh showed a clear separation from all other genomes collected from Thailand (Figure 10). A short read sequenced sample by (Chaput et al. 2020), from Bangladesh was aligned to the consensus sequences generated by long read sequencing Samples (F1R2, F1R3) collected from the same farm, using our tiled PCR approach. This sample formed a monophyletic group with TiLV samples collected from the same farm, confirming long read sequencing in accurately capturing similar strains as to short read sequencing methods, shown in figure 10.

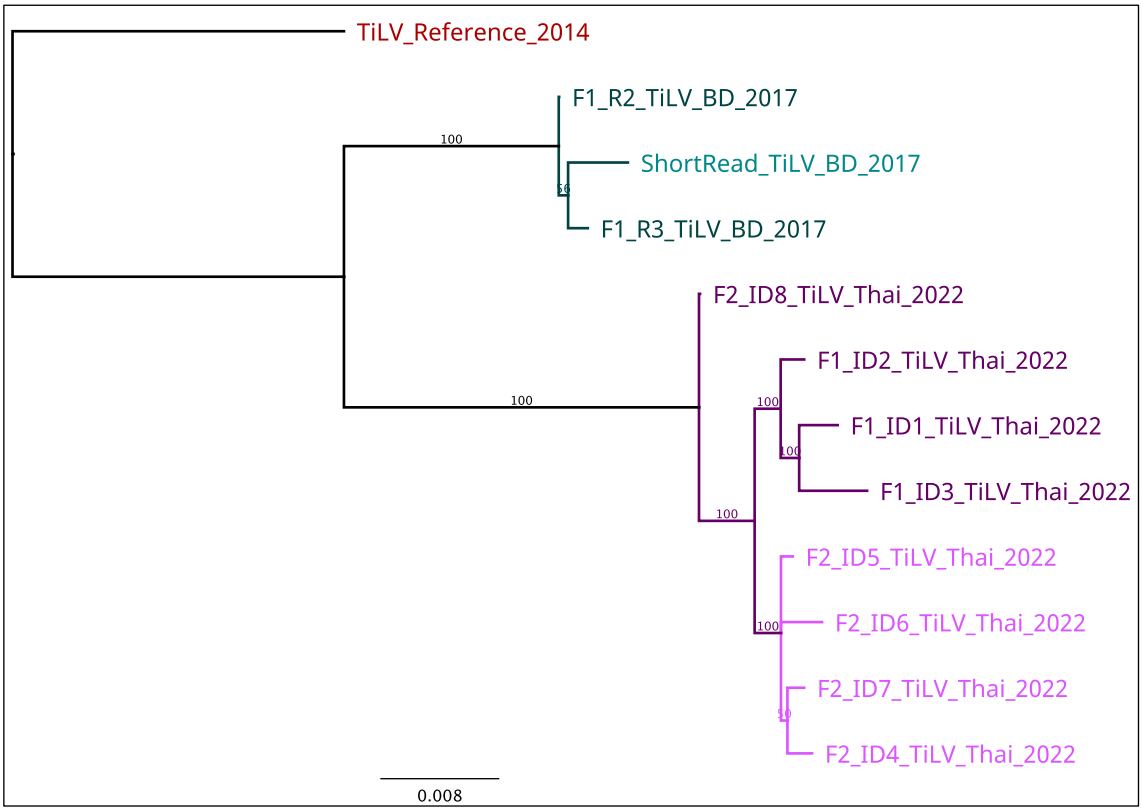


Figure 10. A phylogenetic tree generated in Geneious for full TiLV genome for samples collected from Bangladesh (in green) and Thailand (in purple); all samples were sequenced using long read sequencing. In addition, a short read sequenced genome of a sample collected from the same farm in Bangladesh was added**.** All genomes were aligned to their reference genome. The scale bar represents the number of substitutions per site.

A total of 48 SNPs were seen when aligning the consensus genomes generated by short read and long read sequenced samples from one farm in Bangladesh. These consensus genomes were initially generated from bam files aligned to the TiLV concatenated reference genome. When examining the alignment, SNPs were mainly located at the termini of the segments. Locations across the genome with < 20x coverage, were excluded from the analysis. Loci of regions across the TiLV genome with less than 20x coverage are listed in Supplementary Table 2.

4.4 Discussion

Following a successful developed method for tiled PCR to produce near complete ISKNV genomes for a better understanding of its epidemiology, work in this chapter sought to explore applying this method to TiLV as a very different and segmented RNA virus, also causing mass mortalities in tilapia since its formal identification in 2014 (Eyngor et al. 2014).

A total of 34 primer pairs were generated to span the full ten segments of the TiLV genome. Samples chosen for this study were samples collected in 2017 from Bangladesh, from one farm (Chaput et al. 2020) and samples collected from Thailand, specifically for this project. Not all amplicons generated had the same length, as some segments were larger or smaller than the selected size of (325 bp), and therefore size selection was adjusted accordingly. Additionally, when primers were individually tested, four of these failed to amplify, where one was located near the segment ends. This was expected, as the primers generated from the sequence alignment recovered the ORFs of each segment but did not include the termini of the segments. The other failing primers were located within the ORFs, and further examination showed a missing region around ~3900-4100 bp (segment 3-primer pair 2) in all TiLV sequenced genomes when using the tiled PCR method. One SNP was found in both the forward and reverse primers, located at T4184C and T3869C, respectively, and could be the reason behind the dropping of this region. On the other hand, in Segment 2-primer pair 3, despite failing when individually tested, seemed to be recovered when generated with the tiled PCR. This illustrates that primers need to be continually adjusted as new mutations may arise and can decrease the efficiency and disable amplification of some regions (Brejová et al. 2021). In general, samples collected from Bangladesh had the highest percentage of genome recovery of all sequenced samples, reaching 84.9%, (lacking mainly coverage at the non-coding termini). In some TiLV studies, these regions are excluded from the analysis, due to some similarity to the fish ribosomal sequences (Chaput et al. 2020).

In Thailand, two out of three farms were positive for TiLV. All samples collected were tested with a semi-nested PCR for segment one. Sample (ID3) from farm one, showed an absence of a visible band for the first round of semi-nested PCR

at 620 bp when visualised on a gel. Nevertheless, the tiled PCR protocol for this samples was able to recover 69.5% of its full genome, including the drop-out regions of the termini. MinION sequencing of both positive farms showed that sample ID3 yielded the lowest number of reads. This could be due to the lower concentration of TiLV within this sample, and needs to be further addressed when testing for TiLV. For this protocol a gel electrophoresis might not be a necessary step, as previous work for ISKNV showed a sufficient number of sequences were still produced with no visible band at the expected location (Alathari et al. 2023). A quantitative PCR or ddPCR could assist in detecting the exact number of viral templates needed for good genome recovery.

Challenges were faced when trying to generate a full consensus genome for TiLV using the Artic pipeline. Primers were generated using the *Primal Scheme* software for each individual segment, therefore files needed to be concatenated manually for each produced file for all the ten segments. Primers were then aligned to the TiLV reference genome in Geneious, and showed to be descending from the end of the genome to the beginning. This in turn has caused the location of these primers to be listed incorrectly. Moreover, the left primer pair needed to be reverse transcribed to match the reference genome. This produced several files while raising an error: *List index out of range.* Finally, the produced compressed sorted bam file was used to generate a consensus genome in Geneious. To test if both generated consensus genomes by the Artic pipeline and Geneious were similar, a bam file and consensus sequence previously produced by the Artic pipeline for ISKNV were compared. The consensus genome produced in Geneious had higher genome recovery, as the Artic pipeline masks genome positions with < 20x read depth and changes these positions to “N” (an ambiguous nucleotide) when generating the consensus genome.

When calculating the mean of the total genome recovery when aligned to its reference genome, this was higher for isolates from Bangladesh than isolates from Thailand. This could be due to additional mutations in the TiLV genome, during the five-year sampling gap, as the alignment sequence used to generate the primers were selected for TiLV genomes were from previous year (between 2016-2019), and therefore may be similar to the TiLV genomes infecting fish within that period.

Despite the loss of the termini of the segments, phylogenetic analysis revealed a clear separation of clades according to location and country. Initially TiLV collected from farms from Thailand were long read sequenced and aligned to the reference genome. Two clades clustered according to farm, except for sample ID8 that clustered separately. Removal of this sample had no effect on separation according to farm. Although most genomes had drop-out regions in similar locations, samples of Farm 1 seemed to have a distinct mutation at segment 1 (1099-1259 bp), which contains most predicted proteins. Moreover, sample ID8 was found to have more regions with low read depth, and only 64% of genome was recovered, which was the lowest in samples collected from Thailand. Most of the missing regions were in segments 1, 5 and 6, and this may be a result of mutations occurring at the primer binding locations. Segments 5 and 6 are predicted to include a signalling peptide, suggesting that the encoded proteins of these two segments may exist as part of the virus envelope. Vaccine development has been recently explored focusing on these two segments (Lueangyangyuen et al. 2022), which may highlights their role in host interactions.

A monophyletic group was formed for samples collected from Bangladesh, and this group was closer to the reference genome (from 2014) than the samples collected from Thailand (2022). This highlights the difference in the TiLV genome according to location and time. Samples from Bangladesh lacked a mutation in segment 1, located at 368-518 bp, and may be a consequence for clade separation. A TiLV genome produced from short read sequencing, collected from Bangladesh, clustered with genomes produced by long read sequencing, collected from the same farm. This emphasises the equivalence of our developed method to previously published results. Aligning long read and short consensus genomes that were generated by mapping to the concatenated reference genome, produced 48 SNPs. This is expected due to the higher error rate in long read sequencing. The Artic pipeline overcomes this issue by filtering reads with < 20x coverage to generate a consensus genome for accurate variant calling, and therefore the depth of reads with < 20x coverage were identified, and taken into consideration when calculating the percentage of genome recovery. These were seen across the full genome, but mainly at the termini of each segment.

This is the first attempt to use the Artic pipeline for a segmented virus, with tiles crossing individual segments, by creating two separate pools, which has been an additional difficulty to generate consensus genomes. The Artic pipeline has been recently described as a detection method for the clades of another segmented RNA virus in wild birds in Chile, known as Avian Influenza, using one pool of primers only (Ariyama et al. 2023). Yet, using one pool of primers could increase the risk of primer interactions. Therefore, further adjustment to primers will be needed, and the Artic Network group has suggested adjustments to their software to include the addition of segmented viruses. Our tiled PCR method has successfully produced almost complete TiLV genomes, which has been used to predict its phylogeography, and relatedness to previous TiLV infections. This is a potential tool to understand the spread and evolution of TiLV. Clustering patterns cannot be predicted through focusing on a single segment, as it could be misleading, as previous studies of a single segment showed that it predicts its phylogeny differently, compared to whole genome analysis (Chaput et al. 2020). Additionally, this may lead to missing key information of variants of concern (VOC) that cause increased virulence and transmissibility across non-sequenced segments. Therefore, it is paramount to compare the full TiLV genomes when trying to predict their epidemiology.

However, it is important to mention the limitations of our tiled PCR protocol. First, amplicon sequencing usually generates reads of the most abundant variant at the time of sampling, with the additional stochastic effect of which templates amplify in the first few rounds of the PCR (Delamare-Deboutteville et al. 2021). Second, it generates a consensus sequence that overlooks within-host diversity. TiLV as an RNA segmented virus results in high genetic diversity, allowing for rapid evolution and adaptation to local conditions and ecosystems (Skornik et al. 2021). Therefore, Intra-host single nucleotide variation (iSNV) of the virus is more likely, and it is essential to detect these variations at the population level for determining rates of adaptation and patterns of transmission.

Finally, whole genome sequencing is undoubtedly the way forward to evaluate new variants arising, increasing the suitability of genomic surveillance, which is now feasible and easy to access through portable technologies. Suggesting easier viral collection methods, such as water concentration using filters,

mentioned in chapter 3, could be a next step to help contain and control viral outbreaks in tilapia aquaculture, especially in remote farms with limited access to routine testing. TiLV collected from water sample could reveal its wider population circulating the farms, specifically useful for a fast-evolving RNA virus.

References

Abdullah, Azila, Rimatulhana Ramly, Mohammad Syafiq Mohammad Ridzwan, Fahmi Sudirwan, Adnan Abas, Kamisa Ahmad, Munira Murni, and Beng Chu Kua. 2018. “First Detection of Tilapia Lake Virus (TiLV) in Wild River Carp (Barbonymus Schwanenfeldii) at Timah Tasoh Lake, Malaysia.” *Journal of Fish Diseases* 41 (9): 1459–62.

Aich, Nilav, Anirban Paul, Tanmoy Gon Choudhury, and Himadri Saha. 2022. “Tilapia Lake Virus (TiLV) Disease: Current Status of Understanding.” *Aquaculture and Fisheries* 7 (1): 7–17.

Alathari, Shayma, Dominique L. Chaput, Luis M. Bolaños, Andrew Joseph, Victoria L. N. Jackson, David Verner-Jeffreys, Richard Paley, Charles R. Tyler, and Ben Temperton. 2023. “A Multiplexed, Tiled PCR Method for Rapid Whole-Genome Sequencing of Infectious Spleen and Kidney Necrosis Virus (ISKNV) in Tilapia.” *Viruses* 15 (4): 965.

Al-Hussinee, Lowia, Kuttichantran Subramaniam, Mohammad Shamim Ahasan, Bill Keleher, and Thomas B. Waltzek. 2018. “Complete Genome Sequence of a Tilapia Lake Virus Isolate Obtained from Nile Tilapia (Oreochromis Niloticus).” *Genome Announcements* 6 (26). https://doi.org/10.1128/genomeA.00580-18.

Bacharach, Eran, Nischay Mishra, Thomas Briese, Michael C. Zody, Japhette Esther Kembou Tsofack, Rachel Zamostiano, Asaf Berkowitz, et al. 2016. “Characterization of a Novel Orthomyxo-like Virus Causing Mass Die-Offs of Tilapia.” *MBio* 7 (2): e00431-16.

Behera, B. K., P. K. Pradhan, T. R. Swaminathan, N. Sood, Prasenjit Paria, Abhishek Das, D. K. Verma, et al. 2018. “Emergence of Tilapia Lake Virus Associated with Mortalities of Farmed Nile Tilapia Oreochromis Niloticus (Linnaeus 1758) in India.” *Aquaculture* 484 (February): 168–74.

Brejová, Broňa, Kristína Boršová, Viktória Hodorová, Viktória Čabanová, Askar Gafurov, Dominika Fričová, Martina Neboháčová, Tomáš Vinař, Boris Klempa, and Jozef Nosek. 2021. “Nanopore Sequencing of SARS-CoV-2: Comparison of Short and Long PCR-Tiling Amplicon Protocols.” *PloS One* 16 (10): e0259277.

Cai, Junning. 2022. “Top 10 Species Groups in Global Aquaculture 2020.” *FAO*, June. http://dx.doi.org/.

Chaput, Dominique L., David Bass, Md Mehedi Alam, Neaz Al Hasan, Grant D. Stentiford, Ronny van Aerle, Karen Moore, John P. Bignell, Mohammad Mahfujul Haque, and Charles R. Tyler. 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.

Debnath, Partho Pratim, Jerome Delamare-Debouteville, Mona Dverdal Jansen, Kornsunee Phiwsaiya, Afsana Dalia, Md Abir Hasan,

Saengchan Senapin, Chadag Vishnumurthy Mohan, Ha Thanh Dong, and Channarong Rodkhum. 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): 1381–89.

Delamare-Deboutteville, Jérôme, Watcharachai Meemetta, Khaettareeya Pimsannil, Pattiya Sangpo, Han Ming Gan, Chadag Vishnumurthy Mohan, Ha Thanh Dong, and Saengchan Senapin. 2023. “A Multiplexed RT-PCR Assay for Nanopore Whole Genome Sequencing of Tilapia Lake Virus (TiLV).” *Scientific Reports* 13 (1): 20276.

Delamare-Deboutteville, Jerome, Suwimon Taengphu, Han Ming Gan, Pattanapon Kayansamruaj, Partho Pratim Debnath, Andrew Barnes, Shaun Wilkinson, et al. 2021. “Targeted Gene Sequencing with Nanopore Enables Rapid and Accurate Confirmatory Diagnostic of Tilapia Lake Virus.” *BioRxiv*. https://doi.org/10.1101/2021.03.29.437503.

Dong, H. T., S. Siriroob, W. Meemetta, W. Santimanawong, W. Gangnonngiw, N. Pirarat, P. Khunrae, T. Rattanarojpong, R. Vanichviriyakit, and S. Senapin. 2017. “Emergence of Tilapia Lake Virus in Thailand and an Alternative Semi-Nested RT-PCR for Detection.” *Aquaculture* 476 (July): 111–18.

Dong, Ha Thanh, Saengchan Senapin, Warachin Gangnonngiw, Vuong Viet Nguyen, Channarong Rodkhum, Partho Pratim Debnath, Jerome Delamare-Deboutteville, and Chadag Vishnumurthy Mohan. 2020. “Experimental Infection Reveals Transmission of Tilapia Lake Virus (TiLV) from Tilapia Broodstock to Their Reproductive Organs and Fertilized Eggs.” *Aquaculture* 515 (January): 734541.

Eyngor, Marina, Rachel Zamostiano, Japhette Esther Kembou Tsofack, Asaf Berkowitz, Hillel Bercovier, Simon Tinman, Menachem Lev, et al. 2014. “Identification of a Novel RNA Virus Lethal to Tilapia.” *Journal of Clinical Microbiology* 52 (12): 4137–46.

FAO. 2017. “Outbreaks of Tilapia Lake Virus (TiLV) Threaten the Livelihoods and Food Security of Millions of People Dependent on Tilapia Farming.” https://www.fao.org/fileadmin/user_upload/newsroom/docs/GIEWS%20Special%20Alert%20338%20Global%20I7326EN.pdf.

Fathi, Mohamed, Cathryn Dickson, Malcolm Dickson, William Leschen, Johanna Baily, Fiona Muir, Kristina Ulrich, and Manfred Weidmann. 2017. “Identification of Tilapia Lake Virus in Egypt in Nile Tilapia Affected by ‘Summer Mortality’ Syndrome.” *Aquaculture* 473 (April): 430–32.

Grubaugh, Nathan D., Jason T. Ladner, Philippe Lemey, Oliver G. Pybus, Andrew Rambaut, Edward C. Holmes, and Kristian G. Andersen. 2019. “Tracking Virus Outbreaks in the Twenty-First Century.” *Nature Microbiology* 4 (1): 10–19.

Jaemwimol, Phitchaya, Pattarasuda Rawiwan, Puntanat Tattiyapong, Pattrawut Saengnual, Attapon Kamlangdee, and Win Surachetpong.

2018. “Susceptibility of Important Warm Water Fish Species to Tilapia Lake Virus (TiLV) Infection.” *Aquaculture* 497 (December): 462–68.

Katoh, Kazutaka, Kazuharu Misawa, Kei-Ichi Kuma, and Takashi Miyata. 2002. “MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform.” *Nucleic Acids Research* 30 (14): 3059–66.

Lueangyangyuen, A., S. Senapin, H. T. Dong, S. Unajak, E. Wangkahart, and P. Khunrae. 2022. “Expression and Purification of S5196-272 and S6200-317 Proteins from Tilapia Lake Virus (TiLV) and Their Potential Use as Vaccines.” *Protein Expression and Purification* 190 (February): 106013.

Misu, Masayasu, Tomoki Yoshikawa, Satoko Sugimoto, Yuki Takamatsu, Takeshi Kurosu, Yukiteru Ouji, Masahide Yoshikawa, Masayuki Shimojima, Hideki Ebihara, and Masayuki Saijo. 2023. “Rapid Whole Genome Sequencing Methods for RNA Viruses.” *Frontiers in Microbiology* 14 (February): 1137086.

Nicholson, P., M. A. Fathi, A. Fischer, C. Mohan, E. Schieck, N. Mishra, A. Heinimann, J. Frey, B. Wieland, and J. Jores. 2017. “Detection of Tilapia Lake Virus in Egyptian Fish Farms Experiencing High Mortalities in 2015.” *Journal of Fish Diseases* 40 (12): 1925–28.

Piewbang, Chutchai, Puntanat Tattiyapong, Matepiya Khemthong, Sitthichok Lachroje, Suwimon Boonrungsiman, Tanit Kasantikul, Win Surachetpong, and Somporn Techangamsuwan. 2022. “Dual Infections of Tilapia Parvovirus (TiPV) and Tilapia Lake Virus (TiLV) in Multiple Tilapia Farms: Their Impacts, Genetic Diversity, Viral Tropism, and Pathological Effects.” *Aquaculture* 550 (March): 737887.

Pulido, Liseth Lucero Huamancha, Chìo Marcela Mora, Armando Luis Hung, Ha Thanh Dong, and Saengchan Senapin. 2019. “Tilapia Lake Virus (TiLV) from Peru Is Genetically Close to the Israeli Isolates.” *Aquaculture* 510 (August): 61–65.

Skornik, Revital, Adi Behar, Marina Eyngor, Michal Perry Markovich, Natan Wajsbrot, Eyal Klement, and Nadav Davidovich. 2021. “Temporal Trends of Tilapia Lake Virus Disease in Israel, 2017-2018.” *Transboundary and Emerging Diseases* 68 (6): 3025–33.

Steinhauer, D. A., E. Domingo, and J. J. Holland. 1992. “Lack of Evidence for Proofreading Mechanisms Associated with an RNA Virus Polymerase.” *Gene* 122 (2): 281–88.

Subramaniam, Kuttichantran, Hugh W. Ferguson, Richard Kabuusu, and Thomas B. Waltzek. 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, Win, Taveesak Janetanakit, Nutthawan Nonthabenjawan, Puntanat Tattiyapong, Kwanrawee Sirikanchana, and Alongkorn Amonsin. 2017. “Outbreaks of Tilapia Lake Virus Infection, Thailand, 2015-2016.” *Emerging Infectious Diseases* 23 (6): 1031–33.

Taengphu, Suwimon, Pakkakul Sangsuriya, Kornsunee Phiwsaiya, Partho Pratim Debnath, Jerome Delamare-Debouttеville, Chadag Vishnumurthy Mohan, Ha Thanh Dong, and Saengchan Senapin. 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 (September): 735423.

Tattiyapong, Puntanat, Worawan Dachavichitlead, and Win Surachetpong. 2017. “Experimental Infection of Tilapia Lake Virus (TiLV) in Nile Tilapia (Oreochromis Niloticus) and Red Tilapia (Oreochromis Spp.).” *Veterinary Microbiology* 207 (August): 170–77.

Thawornwattana, Yuttapong, Ha Thanh Dong, Kornsunee Phiwsaiya, Pakkakul Sangsuriya, Saengchan Senapin, and Pakorn Aiewsakun. 2021. “Tilapia Lake Virus (TiLV): Genomic Epidemiology and Its Early Origin.” *Transboundary and Emerging Diseases* 68 (2): 435–44.

Tsofack, Japhette Esther Kembou, Rachel Zamostiano, Salsabeel Watted, Asaf Berkowitz, Ezra Rosenbluth, Nischay Mishra, Thomas Briese, et al. 2017. “Detection of Tilapia Lake Virus in Clinical Samples by Culturing and Nested Reverse Transcription-PCR.” *Journal of Clinical Microbiology* 55 (3): 759–67.

Verma, Dev Kumar, Neeraj Sood, Anutosh Paria, T. R. Swaminathan, C. V. Mohan, K. V. Rajendran, and P. K. Pradhan. 2022. “Reassortment and Evolutionary Dynamics of Tilapia Lake Virus Genomic Segments.” *Virus Research* 308 (January): 198625.

Waiyamitra, Pitchaporn, Puntanat Tattiyapong, Kwanrawee Sirikanchana, Skorn Mongkolsuk, Pamela Nicholson, and Win Surachetpong. 2018. “A TaqMan RT-QPCR Assay for Tilapia Lake Virus (TiLV) Detection in Tilapia.” *Aquaculture* 497 (December): 184–88.

Yin, Jiyuan, Qing Wang, Yingying Wang, Yingying Li, Weiwei Zeng, Jiexing Wu, Yan Ren, et al. 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): 817–24.

4.5 Supplementary Material

Supplementary Table 1:

Multiplex Primers for TiLV (generated for all 10 segments) using the Primal Scheme software; the primer melting temperature (tm) for each primer and the percentage of the GC are listed.

Name	*Pool*	*Sequence*	*Size (nt)*	*%gc*	*tm °C*
TiLV1_1_LEFT	1	TGTGGGCATTTCAAGAAGGAGT	22	45.4	60.3
TiLV1_1_RIGHT	1	TGAGCTCAGATGGAGAACTACAC	23	47.8	59.8
TiLV1_2_LEFT	2	TTGTGTGCAAAAGTACCACGCT	22	45.4	61.7
TiLV1_2_RIGHT	2	ACTAGCAGTAGCGGTGTACACT	22	50.0	61.0
TiLV1_3_LEFT	1	GGTTGCTTTCCTCTGTGGTAGA	22	50.0	60.1
TiLV1_3_RIGHT	1	CAAGGTGGCAGTTGCGTTAAAC	22	50.0	61.2
TiLV1_4_LEFT	2	TGCGTTCTCATCAAAGACACTCT	23	43.4	60.5
TiLV1_4_RIGHT	2	GAGTCACAAGTGGTCCCCTGAA	22	54.5	62.0
TiLV1_5_LEFT	1	CCTTACTAGGGACGGTGACCTA	22	54.5	60.2
TiLV1_5_RIGHT	1	CCATCAGGGTGCACTTGGTA	20	55.0	59.6
TiLV1_6_LEFT	2	GCCTGGGAGGGAAGACTGTAAT	22	54.5	61.6
TiLV1_6_RIGHT	2	GGTGGGCTGGACTGCTTTATAA	22	50.0	60.5
TiLV2_1_LEFT	1	CATTCAAGGGCAGAACTGAGGT	22	50.0	60.7
TiLV2_1_RIGHT	1	GGCTCATTGACTACCCTGTCTT	22	50.0	59.9
TiLV2_2_LEFT	2	ACTCAGGAGGAAGCAATTGATCT	23	43.4	59.7
TiLV2_2_RIGHT	2	AACTTAGCATCCTCGACAGCGA	22	50.0	61.9
TiLV2_3_LEFT	1	TGGATTTGTGATAAAAATCGGCGA	24	37.5	59.9
TiLV2_3_RIGHT	1	CTCCTGTGATATGTGGTTAGCTGT	24	45.8	60.4
TiLV2_4_LEFT	2	CCCACTGCACTAGATCCATTTGG	23	52.1	61.5
TiLV2_4_RIGHT	2	CTCGGGAGAACGTAATGCCTTT	22	50.0	60.8
TiLV2_5_LEFT	1	TTTCAGTGAGGAACCTTGGACC	22	50.0	60.3
TiLV2_5_RIGHT	1	GTCCTCGTAACCCATCCACTTT	22	50.0	60.2
TiLV3_1_LEFT	1	TTGCACAGCTAACTGGGGTTTT	22	45.4	61.1
TiLV3_1_RIGHT	1	TGCCTTGTGCCCAACTTTAACA	22	45.4	61.4
TiLV3_2_LEFT	2	TCTGAGCAAGAGTACCAGCAGA	22	50.0	61.0
TiLV3_2_RIGHT	2	TCTCTATCACGTGCGTACTCGT	22	50.0	61.1
TiLV3_3_LEFT	1	TGCTCAAAGTTCCTCGCCTG	20	55.0	60.6
TiLV3_3_RIGHT	1	GAGGCGGTTGGTCTCCTTTT	20	55.0	60.2
TiLV3_4_LEFT	2	AGCGATAATACCAGCATACTAGCT	24	41.6	59.6
TiLV3_4_RIGHT	2	GATGACGTCCCATCTTGTCTCA	22	50.0	60.0
TiLV3_5_LEFT	1	AGGGGAGCAAGACTTTGTGAGT	22	50.0	61.7
TiLV3_5_RIGHT	1	CTCGCAAATGGGTGTACTGTCA	22	50.0	61.0
TiLV4_1_LEFT	1	ACAAAGACTAGTATGGCAGCTGC	23	47.8	61.1
TiLV4_1_RIGHT	1	TATAAGGCTCCTTCCGACCCTC	22	54.5	60.9

TiLV4_2_LEFT	2	GAAGAGAGAGAGGGAGAACGCT	22	54.5	61.1
TiLV4_2_RIGHT	2	AGCACGATAGGAATCCCCACTC	22	54.5	61.7
TiLV4_3_LEFT	1	CGTCAGTTTGGTTGCTCTCGA	21	52.3	60.9
TiLV4_3_RIGHT	1	AACAACACCAATACTCCCGTCC	22	50.0	60.7
TiLV4_4_LEFT	2	TGCAGATAGGTGACCAGGTCA	21	52.3	60.6
TiLV4_4_RIGHT	2	TATCTTCCAACAGCCCCTGC	20	55.0	59.7
TiLV5_1_LEFT	1	GACTCCAATAGCTATGCAGGCG	22	54.5	61.3
TiLV5_1_RIGHT	1	TCAGTAGCTCTCCAATCACCTCT	23	47.8	60.5
TiLV5_2_LEFT	2	GCAGTACCTAACAGCTTCCCAG	22	54.5	60.8
TiLV5_2_RIGHT	2	AACACCCATGCCAATTGCTACT	22	45.4	61.0
TiLV5_3_LEFT	1	GCCGAAGTCGTTTGTAGTAGCA	22	50.0	61.1
TiLV5_3_RIGHT	1	CACCAGGTAATAGACAAACTTATATTTCTCT	31	32.2	60.2
TiLV5_4_LEFT	2	GCTCGATTAAATCCCTCGTCTGT	23	47.8	60.7
TiLV5_4_RIGHT	2	GAAGCAGAGGGACTTCGTCATC	22	54.5	60.9
TiLV6_1_LEFT	1	TGCATTTTTATCTACAGGATTGTCCA	26	34.6	59.5
TiLV6_1_RIGHT	1	CAGTTCAGATGATGGAGTTCCCC	23	52.1	60.9
TiLV6_2_LEFT	2	AACGAAGTCATAGACTCAGGTGG	23	47.8	60.0
TiLV6_2_RIGHT	2	CACGCGACATTAGCATACAGGT	22	50.0	61.2
TiLV6_3_LEFT	1	GCTCTAAACTGTTTGAGACATCGC	24	45.8	60.7
TiLV6_3_RIGHT	1	AAGCAACTTCATCCTGCATCGC	22	50.0	62.2
TiLV6_4_LEFT	2	ACGGCTTCAGAACTGAATACAAGT	24	41.6	60.8
TiLV6_4_RIGHT	2	TCACATGTATTTATTGATTTTACAGCAGGA	30	30.0	60.7
TiLV7_1_LEFT	1	TGTCCTACAAGATTGGTGAGCTT	23	43.4	59.9
TiLV7_1_RIGHT	1	ACAGAGATGCATGTCCCCTTTG	22	50.0	60.8
TiLV7_2_LEFT	2	AGTATGAAGTGAGCCCCGGATT	22	50.0	61.4
TiLV7_2_RIGHT	2	AGGGATTGGCACTAACCCAACT	22	50.0	61.8
TiLV8_1_LEFT	1	CAACACTAAGAGAGGGCCAAGG	22	54.5	60.8
TiLV8_1_RIGHT	1	AGGTACTGTTTCCGATTGAATTCAAA	26	34.6	59.9
TiLV8_2_LEFT	2	CTTGTTAAGCACGCCGGCAT	20	55.0	61.9
TiLV8_2_RIGHT	2	TTCACGGAAATGGTTGATAGCAG	23	43.4	59.5
TiLV9_1_LEFT	1	TGTCACGATGGATAGAAAATACAGATTC	28	35.7	60.1
TiLV9_1_RIGHT	1	GCCAGCCATGTCAGATATCCTC	22	54.5	60.8
TiLV10_1_LEFT	2	AGTGTGGCAGATTATTTGTCAAGTG	25	40.0	60.4
TiLV10_1_RIGHT	2	AGACTGCACGTCAAGAGACTTC	22	50.0	60.4

Supplementary Table 2:

List of the location of reads with less than 20x coverage across the TiLV genome.

Sample ID	*F1R2_TiLV*	*F1R3_TiLV*	*ID1_TiLV*	*ID2_TiLV*	*ID3_TiLV*	*ID4_TiLV*	*ID5_TiLV*	*ID6_TiLV*	*ID7_TiLV*	*ID8_TiLV*				
Country	BD	BD	Thai	Thai	Thai	Thai	Thai	Thai	Thai					Thai
Regions with <20 read depth	1-30	1-32	1-33	1-33	1-35	1-35	1-35	1-35	1-33	1-45	5179-5188	5275	6607-6612	6753-6755
	1593-1878	1593-1878	378-518	368-518	370-518	372	368-518	369-392	878	368-518	5191-5194	5281	6618	6758-7035
	3052-3174	3052-3167	1099-1259	1099-1259	847-1259	424	847-1020	410-482	1593-1878	529	5197-5199	5283	6628-6629	7373-7528
	4084	3956-4105	1593-1878	1593-1877	1593-1878	1593-1878	1593-1878	484-485	3052-3176	531-532	5202	5678-5992	6631-6635	7860-8037
	4424-4621	4424-4621	3052-3175	3052-3165	3053-3178	3052-3165	3052-3168	488-489	3956-4621	538-540	5204-5205	6330-6505	6639	8606-8910
	5677-5803	5677-5801	3958	3956-4105	3956-4621	3956-4105	3956-4107	847-1020	5677-5807	558	5208-5210	6512-6514	6649-6650	9244-9478
	6812-6904	6131-6225	3965-4620	4325	4962-5080	4424-4619	4325	1341-1878	6566-6904	568-569	5212	6520	6653-6655	9816-9940
	7858-8034	6239	5677-5807	4423-4620	5084	5678-5807	4422-4621	3052-3170	7858-8037	572-573	5214-5216	6523-6527	6659	10277-10323
	8606-8733	6571	6566-6904	5677-5807	5676-5807	6596-6853	5677-5807	3956-4105	8606-8737	586-768	5219-5223	6533	6661	
	9244-9475	6812-6904	7858-8035	6566-6904	6566-6904	7858-8037	6231	4424-4619	9244-9478	1099-1259	5225-5226	6535	6675	
	9817-9935	7761	8606-8740	7858-8036	7858-8038	8606-8733	6566-6904	5676-5805	9816-9937	1593-1875	5228-5229	6540	6696-6697	
	10277-10323	7816	9244-9478	8606-8737	8606-8917	9243-9478	7858-8037	6132-6244	10277-10323	2559-2715	5240	6545-6546	6699	
		7858-8037	9816-9937	9244-9478	9242-9480	9816-9937	8606-8737	6568-6904		3052-3178	5242	6558-6559	6702-6703	
		8454	10277-10323	9816-9937	9816-9940	10277-10323	9242-9480	7858-8037		3956-4621	5251	6562-6563	6713	
		8602-8733		10277-10323	10277-10323		9816-9934	8606-8737		4744	5253-5255	6565-6566	6719	
		9241-9478					10277-10323	9244-9480		4762	5259-5261	6586	6722-6723	
		9817-9936						9816-9936		4803-4804	5265	6595	6737-6740	
		10277-10323						10277-10323		4827	5269	6603-6605	6751	

Supplementary code 1: The code used to determine read depths less than 20x coverage.

```
lines = !samtools depth -a [sorted.bam File produced by the Artic pipeline]

coverage = [(int(y[2]), int(y[2])) for y in [x.split('\t') for x in lines]]

cautious_loci = [x[0] for x in coverage if x[1] <20]

def group_consecutive_numbers(numbers):
    groups = []
    current_group = []

    for num in sorted(set(numbers)):
        if not current_group or num == current_group[-1] + 1:
            current_group.append(num)
        else:
            groups.append(current_group)
            current_group = [num]

    if current_group:
        groups.append(current_group)

    return groups

def format_groups(groups):
    formatted_groups = []
    for group in groups:
        if len(group) == 1:
            formatted_groups.append(str(group[0]))
        else:
            formatted_groups.append(f"{group[0]}-{group[-1]}")
    return formatted_groups

consecutive_groups = group_consecutive_numbers(cautious_loci)

formatted_groups = format_groups(consecutive_groups)

formatted_groups
```

5. *The experience of an epidemiological study for in-field genomic surveillance for an ISKNV outbreaks in Ghana*

Low-Income Food-Deficient Countries are particularly vulnerable to threats to food security and are most likely to benefit from the capacity for real-time epidemiology. Yet, they are the least likely to have access to centralised sequencing. This shortfall means that there are long delays in obtaining results, where rapid analysis of the sequencing information is essential for early interventions. Since the development of the MinION sequencer, the potential for field-based sequencing has been promising, with the tagline of Oxford Nanopore Technologies (ONT) stating "To enable the analysis of anything, by anyone, anywhere". Despite this, the adoption of field-based sequencing of diseases has been relatively slow, with a primary focus on human infections, with much of the work carried out by a handful of research groups.

The potential barriers to field-based sequencing include lack of expertise, access to lab equipment in remote settings, unreliable power, difficulty in maintaining cold-chain for reagents, no access to internet connection and cutting-edge computers for processing large data sets. Also, a barrier of cost and lingering doubt over error rates has hindered its adoption. Until these issues are resolved, and a greater number of case studies are published to guide future efforts, it is unlikely that real-time genomic surveillance using portable sequencing will achieve its full potential.

In this chapter, I outline my efforts to evaluate real-time sequencing of Infectious spleen and kidney necrosis virus (ISKNV) in the field during a trip to Ghana. I describe the issues through first-hand experience of the challenges of field-based sequencing and offer a perspective on how best to overcome these challenges in the future.

5.1 Section 1 - Designing a portable field kit

The equipment used for preparing samples for DNA extraction, library preparation and sequencing are usually available in most molecular labs. This includes a thermocycler, vortex, centrifuge, and cold storage. To create a portable genomic surveillance system, smaller size equipment was transported to Ghana using a Pelicase (Peli Products, UK), carrying a portable thermocycler machine (Mini16 Thermal Cycler) (Figure 1), and a miniGel electrophoresis device, both by MiniPCR (Amplyus, Cambridge, MA, USA), a mini centrifuge and mini vortexer. The Pelicase included reagents transported at ambient temperature: PEG, NaCl, syringe filters as back up plans for viral concentration of water samples; a handheld Quick grip for water filtering; a DNA extraction kit (Qiagen); dissecting kits, Tricaine (MS-222), and different sizes of LoBind microcentrifuge tubes. The following reagents were transported on ice blocks to maintain temperatures < 0°C during transport: pre-designed primers, polymerase and library prep reagents. The following reagents were transported on cool packs to maintain temperatures ~4°C: AMPure XP beads; a 1kb ladder (NEB) and positive controls for ISKNV extracted DNA.

A laptop, portable compact sequencing devices (an MK-1B: a small sequencing device that operates on a laptop; and an MK-1C with its fully integrated computer) (Figure 2), flowcells and flongles by Oxford Nanopore Technologies (ONT), were all carried in a personal backpack. A Field Sequencing kit (SQK-LRK001) (ONT) comprising lyophilised reagents for minimising the need for a cold-chain was also taken, although it is only designed for one sample (no multiplexing through barcodes) and is announced to be discontinued by the company in 2024. All ONT reagents had to be taken to perform field work, as there is no supply chain for ONT products in Ghana, and the nearest provider is in Cape Town in South Africa.

Figure 1. A MiniPCR machine connected to a mobile phone, performing a tiled PCR on ISKNV isolates collected from Lake Volta, Ghana.

5.2 Section 2 - Establishing a lab in a resource limited setting

The Commission of Fisheries provided a physical workplace near Lake Volta, a water chemistry lab based on a grow-out farm for tilapia fingerlings, and was sampled during our sampling campaign. This was situated in the Akosombo region, which is around two hours away from Accra, and was used as a base for easier access to the fish farms selected for sampling. An electric power source was provided, allowing access to a 4°C fridge, a −20°C freezer, and a water bath which was used for DNA extractions. In general, the laboratory had no molecular equipment. After approximately 14 hrs (six hours of flight travel), the Pelicase was opened and the contents were placed in a temperature appropriate unit upon arrival at the destination. Most reagents thawed during travel, but remained cold, with minimal apparent impact on their performance.

Following our first sampling day at the nearest farm (Farm A), DNA extractions were performed on fish tissue samples and the extracted DNA was taken to the hotel to perform the 3-hour tiled PCR in the evening to ensure samples were ready for library preparation and sequencing the following day. This was done due to the miniPCR's reliance on a personal mobile connectivity for its performance. When amplicons were visualised on the miniGel electrophoresis device, we were not able to see bands at the 2kb region (amplicon target size), except for the positive control. Access to a Qubit fluorometer could have assisted in confirming the success of the DNA extraction, and quantifying the concentration of DNA to be added to the multiplex PCR. The small form factor of the Qubit would have added little to the shipping requirements.

Library preparation was approximately three hours, and samples were prepared for sequencing on Lake Volta near the floating cages, shown in figure 2. Wifi connection was established through mobile device hotspots (4G), and we were able to perform a sequencing run. Early sunsets reduced time available for sequencing to prevent working in the dark and far from facilities. This was not originally planned for, and the sequencing was cut short after the beginning of the first MinION run.

The rest of the five farms were located at distances of less than 100 km, with approximate travel times of 2-3 hours by car. The journey was usually difficult and longer than expected due to unpaved and broken roads. Accompanied by a team belonging to the Commission of Fisheries, we managed to reach all the farms that were selected for this study. All fish sampling took place on farms and tissue samples were kept in either RNALater or/and ethanol. Water samples collected from each farm were taken back to the provided laboratory for filtration. A range of 200-500 mL of water was also taken back to the lab and kept at 4°C. Prior to the trip, we were assured that there would be access to a centrifuge suitable for 50 mL Falcon tubes, but this was not available on arrival. Consequently, plans to concentrate viral particles from water samples using PEG were abandoned and our backup plan of concentration via filtration was enacted. Water samples were filtered using sequentially decreasing sized syringe filters assisted with an Erwin® quick-grip minibar 102 clamp (6") to facilitate the pumping of the water, with a custom 3D-printed adaptor for the syringe. A small team of volunteers was used to process samples due to the difficulty in pushing the water through the filters, especially when processing through the 0.1 µm filters.

Figure 2. In-field sequencing, in a hotel and near the fish cages on Lake Volta, Ghana. Sequencing tissue samples using the MK-1C (ONT) on the left, and the MK-1B (ONT), on the right.

5.3 Section 3 - Training farmers and local experts

To test the practicality of on-site training, a medical veterinarian, with limited knowledge of molecular biology, and who was also a member of the Commission of Fisheries, was trained to run a tiled PCR, perform library preparation and MinION sequencing. After five hours of one-to-one training, the trainee was able to successfully complete all steps. Sequencing by this member was performed using a library prepared prior to departure in Exeter, as a positive control (Figure 3).

Figure 3. Training a member of the commission of fisheries; showing Kwaku Duodu performing sequencing using the MinION sequencer on ISKNV samples.

Sequencing results within the first hour were similar to those obtained in Exeter, with the length of reads at expected range (2 kb), yet the time limitation and our inability to leave our devices behind meant we had to cut the sequencing short. To reduce time, another library was prepared using the Field Sequencing kit, where samples are prepared in less than 20 min, opposed to 3 hours, but this kit falls short being designed for one sample. Although sequencing of the sample was successful, I was not able to train the veterinarian in appropriate methods of data analysis in the time available. Accurate interpretation of outputs from Artic require a grounding in knowledge of bioinformatics and command-line Linux. Therefore, a future recommendation would be a week-long course following sampling where trainees are provided appropriate training. Trainees could accelerate their training prior to the workshop by gaining a fundamental understanding of sequence analysis through available online bioinformatics training modules and specific educational guides for nanopore sequencing (Salazar et al. 2020). Despite this, the portability of the MinION sequencer has made it possible to perform training in the field, making this a revolutionary feature and particularly appealing for resource-limited countries. A short campaign to seed countries with externally trained experts could accelerate in-country training through a domestic program, reducing costs and avoiding entry restrictions for foreign visitors. Interestingly, during the trip I attended a meeting with the Director of the Commission of Fisheries in Ghana, where new aquaculture development plans were discussed to increase tilapia production. Their central focus was the

need for molecular training to assist in the diagnosis and the control of diseases affecting the growth of aquaculture in Ghana. Therefore, it is likely that government policy would align with the proposed training needs if a suitable source of funding could be identified.

5.4 Discussion

This chapter aims to provide an overall view of our fieldwork experience in Ghana, demonstrating in-field feasibility and repeatability for monitoring and tracking viral outbreaks in aquaculture in resource limited areas, highlighting the technical and economic challenges and restrictions that were faced during this study.

During sample preparation an initial technical issue to arise was DNA quantification. When testing our generated amplicons on a gel, we were not able to see bands at the amplicon target size. The presence of a Qubit fluorometer could have been useful to confirm the amount of DNA to be added to the multiplex PCR. Yet, it was noticed that despite high quantification of DNA, the presence of ISKNV was low in most samples when tested using the ddPCR upon our return to Exeter. The presence of a portable qPCR machine could have been a solution for quantifying ISKNV, and to avoid non-specific DNA targets of all microorganisms as well as fish tissue. In addition, due to performing sequencing that lasts 24-72 hrs, we were unable to leave our laptop and sequencing device in the field to complete the sequencing, and failed to maintain an internet connection during transfer. Therefore, we suggest creating a set up for overnight sequencing, or early sequencing during the day to avoid moving the devices during sequencing. A portable electric generator is essential to maintain charging the laptop while sequencing, and a portable light source could assist for evening sequencing runs. Finally, the time spent in the field and in Ghana needs to be extended for troubleshooting, resolving the technical issues faced during our field trip, and for providing suitable training for downstream analysis of data.

Economic issues were highlighted during our discussions with farmers. They spoke of their experiences since the initial outbreak in 2018, where some were explaining events dated long before that year, due to bacterial infections. For the Commission of Fisheries and authorities, difficulties arose between the rules they would like to impose on farmers, such as unauthorised farms and the undocumented import of fingerlings, and the difficulties the farmers and consequently their families would face due to their losses in tilapia production. They mentioned that the major bottleneck to the increase of tilapia production is the high cost of fish feed, and one of the farmers mentioned a rapid increase in

feed price, from 6 GH₵ to 25 GH₵ per kilogram in the past year (equivalent to £1.64 per kg). Due to a considerable reduction in tilapia produce, mainly caused by ISKNV outbreaks, farmers couldn't afford this increase, reporting extremely narrow profit margins, and were doubtful of any proposed solutions.

Despite the promising performance of the MinION sequencing device, and its ability to carry out real-time epidemic surveillance for understanding the way viruses transmit, spread, and evolve, it still falls short of its potential due to several factors. Proposing an epidemiological study for the management of ISKNV using the MinION sequencer, as a cheaper and more accessible approach than other sequencing platforms, this technology remains costly for farmers in Ghana. The capital costs for purchasing MinION platforms is not insignificant, with the Mk-1B and Mk-1C costing $1,000 and $4,900, respectively. Furthermore, the lifespan of these devices can be short, with rapid obsolescence. For example, the Mk-1C sequencer was introduced in 2019, with an integrated screen and GPU to remove the need for a laptop for sequencing. However, increases in sequencing throughput and advances in sequencing chemistry now exceed the capabilities of the Mk-1C for processing data and on Dec 8th 2023, ONT announced that it is being discontinued in early 2024, with hardware and software support removed by 2026 ("Oxford Nanopore Technologies" 2023). Similarly, automated sample preparation devices such as the Voltrax, were rapidly replaced (at cost to the user) by the Voltrax2, which is soon to be replaced again by the TraxION. For early adopters of the technology, these continual updates of capital equipment are a source of frustration as experimental protocols need to be re-evaluated and re-written. However, for establishing a programme of standardised field-based genomic surveillance, the rapid obsolescence, and associated costs are a significant barrier to adoption. Even for the cheapest device for sequencing (the flongle at $70 per unit), the reagents used for library preparation are ~ $600. Although multiplexing samples (using up to 96 barcodes) is now feasible, reducing the cost to approximately $7 per sample still remains high for low-income countries. Even when accounting for these costs, other factors such as a lack of representatives of Oxford Nanopore Technologies in certain areas in the world limit access to products and associated support. Shipping to these areas or countries is often delayed, with significant impact on the performance of reagents. Upon its release, the lyophilised sequencing kit (SQK-LRK001) was

touted as a solution to cold-chain issues with field sequencing in remote settings. However, its planned obsolescence in 2024 is perhaps an indicator that there are more significant challenges such as local infrastructure to support supply chains, trained expertise, and expertise in analytics that currently hinder adoption.

Unless a collaborative effort is proposed to generate a plan for disease management on Lake Volta, such as centralised sequencing, there are major barriers to small farms adopting this approach for routine monitoring. Centralised whole genome sequencing (WGS) has been successfully adapted for human outbreak investigations, such as SARS-Cov2, where isolates are sequenced at a centralised lab, and a generic analysis report is produced (Beukers, Jenkins, and van Hal 2021; Grant et al. 2018). Governmental and financial support will be needed for the success of this approach, and would be a welcome and impactful investment, considering its potential to deliver food security, livelihoods, and reduce the overall distress to the region. If such an approach were to be adopted to get the ISKNV outbreak under control, localised sequencing could be achieved by creating a mobile lab for routine monitoring. During the Zika outbreak sequencing was carried out in real-time on a mobile laboratory bus, where results were obtained from patient samples in less than 48hrs (Loman 2017). This has great potential, especially with the drive to farms located on Lake Volta being a few hours apart. In our work we have shown that water sampling can efficiently replace destructive tissue sampling for monitoring outbreaks (see Chapter 4), further reducing the costs to the farmer, the need for sampling by veterinary specialists and associated permits and permissions to routinely monitor the lake.

The expense of this approach remains the biggest limiting factor, as the technical issues for applying it for epidemiological studies can be easily addressed. Yet, MinION sequencing is constantly evolving, including more rapid turn-around time, and a price decrease is continuously witnessed per sample. Increasing use of cloud computing and decreasing costs of data transport and storage, coupled with improving global availability of trained bioinformaticians will further expand the possibility of sharing data for processing, reducing the expenses needed to ship samples to countries with established sequencing facilities, or travelling abroad for training purposes.

Lake Volta

References

Beukers, Alicia G., Frances Jenkins, and Sebastiaan J. van Hal. 2021. "Centralised or Localised Pathogen Whole Genome Sequencing: Lessons Learnt From Implementation in a Clinical Diagnostic Laboratory." Frontiers in Cellular and Infection Microbiology 11 (May): 636290.

Grant, K., C. Jenkins, C. Arnold, J. Green, and M. Zambon. 2018. "Implementing Pathogen Genomics: A Case Study." London: Public Health England.

Loman, Nick. 2017. "Genomic Sequencing Offers New Clues to Zika's Spread from Brazil to the Caribbean to Miami." Fred Hutch. May 24, 2017. https://www.fredhutch.org/en/news/center-news/2017/05/using-genomic-sequencing-to-track-zika-spread-from-brazil-to-carribean-to-miami.html.

"Oxford Nanopore Technologies." 2023. Oxford Nanopore Technologies. 2023.https://community.nanoporetech.com/posts/minion-mk1c-discontinuatio?utm_campaign=Discontinued%20kits%202023.

Salazar, Alex N., Franklin L. Nobrega, Christine Anyansi, Cristian Aparicio-Maldonado, Ana Rita Costa, Anna C. Haagsma, Anwar Hiralal, et al. 2020. "An Educational Guide for Nanopore Sequencing in the Classroom." PLoS Computational Biology 16 (1): e1007314.

6. General discussion

Viruses tend to be considered the most important potential threats in aquaculture (Debnath et al. 2023). Their easy transmission through water has facilitated their spread and transfer to different parts of the world. Rapid detection is essential to contain and control their spread, and many methods have been described for their identification. These methods usually involve trained personnel for fish sampling, observing clinical symptoms, and a large number of fish sacrificed. Yet, remote fish farm establishments lack access to diagnostic laboratories and an increased difficulty arises in identifying the cause of infection, due to the similarity of clinical symptoms of viral diseased fish. Furthermore, subclinical viral infections in fish emphasises the importance of the presence of an established routine testing programme for fish farms.

In this thesis, we were able to develop a tiled-PCR method for an in-field whole genome sequencing (WGS) approach for a virus affecting tilapia fish, known as ISKNV. The Artic pipeline tool, designed specifically for nanopore data generated from tiling amplicon schemes, was employed for this study, providing insight into circulating strains during outbreaks. Detected mutations allowed for the prediction of ISKNVs evolutionary rate, and locate non-synonymous mutations in important structural proteins. We successfully detected ISKNV in water samples collected from floating cages, by concentrating viral fractions on filters. Both tissue and water isolates were comparable, exhibiting similar mutations, with an advantage of water samples revealing the presence of more than one circulating strain. Additionally, our interest expanded to testing the practicality of this method for other viruses in aquaculture, by applying this protocol to Tilapia lake virus (TiLV), an equally important virus hindering the growth of tilapia globally.

My primary focus on ISKNV was mainly due to its capability to infect a wide range of fish species, causing mass mortalities globally, and being listed in new countries almost every year. Devastating outbreaks affecting fish farms located across Lake Volta in Ghana have been the focus of this study, due to the livelihoods and food security being at stake in these communities. A wide range of large dsDNA viruses belonging to the family *Iridoviridae*, are known to be serious impediments to the expansion of global aquaculture, where similar

symptoms are noticed during infection (Qin et al. 2023). The lack of a universal gene marker for viral detection has increased the challenge to monitor their presence. The efficiency of targeted tests comes at the cost of focussing on a single part of the genome, lacking the ability to discover mutations occurring across the genome, which may be involved in developing more virulent and transmissible strains. Nevertheless, mutations in targeted genes used for detection, may prevent their amplification and consequently fail to identify novel variants of concern.

Key findings

In Chapter 2, the Arctic pipeline to generate consensus genome was jointly used with the tiled PCR method for phylogeographic analysis of ISKNV isolates collected from Ghana during a series of outbreak events, since its first discovery in Lake Volta, in 2018. Our developed method was tested on isolates harvested from cell lines and applied to extracted DNA obtained directly from tissue samples (archival samples). An additional set of samples were collected in 2022 from one of the previously sampled farms, in search of mutational changes present in the genome over time. This was successful in determining the number and location of mutations that have taken place in the ISKNV genome and also predicting its evolutionary rate, with mutations occurring in the major capsid protein (MCP). Confirming non-synonymous mutations located at coding regions of the genome is necessary for genomic surveillance by relating to changes observed in the host. Furthermore, this information is essential for the development of diagnostics, vaccines and drugs for treatment. In addition, the number of ISKNV templates needed to recover at least 50% of the full ISKNV genome was determined through the use of ddPCR, identifying a practical limit under which tiled PCR is not cost- or time-efficient. Similar approaches could be used to identify ddPCR-derived limits for other viruses, detecting low quantities of viral template without the need for a standard curve to be customised for each virus. The full ISKNV genome circulating Lake Volta has been annotated and submitted to the NCBI, and compared to other ISKNV samples collected from different hosts and countries. Samples collected from Ghana created a separate clade to all other samples collected from different locations, despite some sampled being collected within the same year.

In Chapter 3 two research aims were established for fieldwork. The first was to perform a phylogeographic study for ISKNV samples collected from six geographically distant farms, which included previously sampled sites. Our second aim was to evaluate the possibility of concentrating ISKNV from water samples from floating fish cages in Lake Volta for non-destructive sampling, comparing sites at different distances from the shore. We combined this approach with developed genomic surveillance for ISKNV in Ghana, using portable equipment in a farm setting. ISKNV was detected in both water and fish tissue samples collected from farm sites across Lake Volta, and similar strains were identified in paired water and tissue samples. Additionally, water samples were able to provide insight into the wider diversity of strains in the water by detecting the presence of at least two circulating strains, confirmed by short read sequencing. This highlights the limitations of focusing on mutations observed in consensus genomes that represent the dominant variant of the virus, and was unexpected for a slow evolving dsDNA virus. Another assumption of this protocol is that individual fish are infected by a single strain, and consensus genomes will not be able to resolve intra-host single nucleotide polymorphisms (iSNVs). While this might be an acceptable approximation for dsDNA viruses, it is critical to determine intra host variation in fast evolving RNA viruses which may produce multiple variants in every round of replication. This could be misleading in epidemiological studies when resolving transmission chains and predicting evolution rates. Nevertheless, water sampling avoids destructive sampling of fish and reduces costs to the farmers, providing an overview of fish viruses circulating the farms. In one fish farm, despite detecting high concentrations of ISKNV in the water, only one out of 12 sampled fish showed to be infected when testing with conventional PCR, which could have been missed with a less robust sampling campaign that relied on fish sampling alone. In general, ddPCR was able to detect reasonably low numbers of viral templates from both tissue and water samples. Farms situated at further distances from the shore had very low viral template in the water surrounding the cages. This information is valuable to both farmers and the Commission of Fisheries to advise farmers for optimal placement of cages to reduce cross-infection or to minimise impact of an ongoing infection when coupled with improved understanding of water flow in the lake. However, most farms had inland ponds, used for fish as grow-out farms for fish fry and

fingerlings; these grow-out ponds were usually supplied with water from the lake. Detecting ISKNV in the surrounding waters could encourage farmers to refrain from such practice to minimise exposure of fish to viral pathogens.

We evaluate the success of our developed tiled PCR method (Chapter 4) for an equally important ssRNA virus threatening tilapia aquaculture, known as TiLV. This virus has been known as an emerging threat to tilapia aquaculture worldwide, for more than a decade. In 2017, FAO issued an alert for tilapia-producing countries to initiate an active TiLV surveillance programme. Here, we followed the same steps as used in for ISKNV (in Chapters 2 & 3) to generate a TiLV consensus genome for genomic surveillance of fish tissue samples collected from Thailand. Samples available in our lab from a previous study (Chaput et al. 2020), collected from Bangladesh, were used as a control to test the efficiency of this protocol, and to compare to the most recent TiLV isolates. Applying this method to TiLV (a negative sense ssRNA virus) had unique challenges, as a final consensus genome was not generated with the Artic pipeline, rather a bam file produced by this pipeline was used to form a consensus genome. Hopefully improvements to the Artic-network pipeline to better evaluate segmented ssRNA viruses will be incorporated in future updates. Consensus genomes generated by short read sequencing clustered with the long read sequenced isolates collected from the same farm, confirming the equivalence of our selected method to produce similar results to other sequencing methods. Despite the small number of samples included in this study, and their phylogenetic placement close to sequences from farms in Bangkok, Thailand, there was a clear cluster separation between farms, indicating the capability of this method to differentiate between samples according to location. Only one sample clustered separately from both farms, and is most likely due to lowest genome recovery in comparison to all TiLV sequenced samples, when aligned to its reference genome. Further research is recommended to generate a consensus genome through the Artic pipeline for negative sense ssRNA viruses, to reduce the time needed to produce robust results.

Difficulties faced during our sampling campaign have been addressed in (Chapter 5). This includes the challenges experienced for performing our tiled PCR method in the field, using a portable carrier for equipment and a basic lab to process

samples. Some advice is provided here to improve results in a future sampling campaign: Despite the technical challenges experienced, we can consider this as a successful attempt and approach, given it was the first time to be applied in the field, encountering new circumstances (such as heat and an unfamiliar working space). Additionally, the length of time needed for sequencing could be reduced, as sufficient data is usually available within the first few hours from the start of sequencing (Quick et al. 2016). Training of a member of the Commission of Fisheries was successful, and proved the feasibility of performing field-based training, to reduce travel cost for trainees. In addition, suitable training for downstream analysis of data needs to be considered, and/or a collaboration with in-country informatics specialists. A collaborative effort is essential for adopting this approach for managing ISKNV on Lake Volta, and future routine monitoring. We suggested a centralised whole genome sequencing (WGS) system to be utilised, and a generic analysis report to be produced and assessed by the Commission of Fisheries, frequently.

During our sampling trip to Ghana, we took the opportunity to speak to farmers at each sampling site, where they expressed their concerns and the measurements taken to reduce the impact of the deadly episodes experienced in tilapia fish production since 2018. The overall view was described as 50-90% of fish losses, mainly experienced in adult fish, and usually when transferred from grow-out pond to the lake. Some mitigation plans included food regimes in the affected farms (known as break feeding). They stated that high quality feed with high fat and protein, may cause to stress on the organs, yet reducing these nutrients and fasting has not shown any noticeable change. Overall, farmers had treated the pond water with bleaching powder, salt, oxygen flow, antibiotics, and heat shock, with no noticeable improvement in fish survivorship. The farmers also reported that water tested in the lake had a similar chemistry usually, yet its quality deteriorated when the water tide was high. A growing catfish sector was being established in some farms, to compensate for the losses in tilapia, with more research required as to the impact of co-localisation of tilapia and catfish on pathogen transmission. One extreme solution may be a future of culturing different species that are less susceptible to diseases which are endemic in the lake, although tilapia is considered a delicacy in Ghana, and so any such proposal might meet cultural resistance.

Future monitoring plan

For a future surveillance strategy, a rapid workflow for sample processing should be developed. The first step should be creating a detection tool for ISKNV in the water, such as investing in a rapid lateral flow test, specific for ISKNV. This will assist in the initial identification of the virus in the farms. Additionally, it will be useful to collect water away from the fish farms, for a better understanding of how far ISKNV travels through water, and provide feedback to policy makers on a specified safe distance between farms. Using an App to connect the Commission of Fisheries to the farmers will allow for quick notification and assist the authorities to guide the farmers, and share a list of recommended measures to control the outbreak. Routine monitoring systems could also be applied and followed up in this app. Later on, to help identify the source of infection and mutations linked to VOCs, the portable sequencing device will be used, and a trained team of scientists and/or veterinarians could carry out the sequencing at the farms using a minibus, for transportation and carrying lab equipment. Here, the Ghanaian authorities will be able to integrate metadata, molecular genetic data and surveillance systems to better understand the links between outbreaks. Developing a new primer scheme, biannually or annually, will be required for updated primers that may include additional mutations, and more regularly for fast-evolving viruses. Fortunately, all the current limitations from our field study can be addressed and further developed, yet economic issues will require a collaborative effort from governmental and funding institutes. Molecular training for diagnostic testing needs to be a priority to overcome diseases in Lake Volta, stepping away from relying on other countries. All the above aims will initially need assistance from governments, other countries and/or organisations, with considerable effort and commitment, and avoid these solutions to be an added burden to a financially struggling country.

Advancements in technology such as portable next generation sequencing (NGS) and artificial intelligence (AI) are increasingly being tested for early diagnosis of disease (MacAulay et al. 2022). This has led to easier detection of different organisms without impacting the environment or the organism itself. Currently, environmental DNA (eDNA) and environmental RNA (eRNA) methods

are being extensively used in aquaculture and fisheries settings to understand the presence of different pathogens in water, following targeted or passive methods (Bohara, Yadav, and Joshi 2022; MacAulay et al. 2022). Although DNA within water or sediment samples may not be indicative of active infectious stages of a pathogen, eRNA detection can indicate active gene transcription (MacAulay et al. 2022). Common viruses found in freshwater aquaculture which have been detected using eDNA and eRNA methods include red seabream virus (RSV), and salmon alphavirus (Bohara, Yadav, and Joshi 2022).

More recently, Kellner et al. designed a CRISPR-based diagnostic tool that combines nucleic acid preamplification with CRISPR–Cas enzymology for specific recognition of desired DNA or RNA sequences (Kellner et al. 2019). It is termed specific high-sensitivity enzymatic reporter unlocking (SHERLOCK), and allows multiplexed, portable, and ultrasensitive identification of RNA or DNA from clinical samples. SHERLOCK is a potential method for rapid detection and identification of infectious diseases, which is characterized by sensitivity and specificity comparable to traditional PCR-based methods, but does not require sophisticated equipment and has a very low estimated cost. Embedding CRISPR-Cas into molecular diagnostics may provide a step-change in global diagnostics programs (Mustafa and Makhawi 2021).

The direct diagnosis of diseased fish underwater requires a high level of technology, and the diversity and heterogeneity of fish diseases increases the difficulty of diagnosis, and currently the accuracy of diagnosis using these various physiological indicators is low. Image-based disease-diagnosis techniques have been widely used in the diagnosis of fish diseases (Li et al. 2022). Image processing combined with computer vision may provide a real-time, non-invasive, and economical technique for disease diagnosis. Camera images can detect disease on the surface of the fish, and microscopic images can provide details of minor changes in tissues within the fish to diagnose pathogens (Li et al. 2022). AI has the potential to considerably reduce the time required to survey fish for disease whilst simultaneously allowing for higher throughput but requires significant input in “teaching” the AI to detect specific diseases. These methods can revolutionise remote diagnostic testing (MacAulay et al. 2022; Li et al. 2022). Although setting up AI-driven image analysis per site will be cost-prohibitive, as

these new technologies evolve and become more available, a price decrease is usually witnessed.

Finally, there is an increase in reported incidences of viral disease in aquaculture, due to several factors, such as intensified farming, transporting fish with subclinical infections and global warming. For a global growing population, with 44 countries considered as Low-income Food-deficient countries (LIFDC) (FAO 2023), this is a major concern. This study presents an early diagnostic tool of viral infections for timely containment, for an improved productive and economic performance in aquaculture systems. Rapid genome sequencing during ongoing outbreaks has only been possible in recent years with the invention of new sequencing technologies, accompanied by real time data sharing. Taking advantage of these technologies can reshape disease diagnosis and monitoring in aquaculture, enabling researchers to confirm their dynamics, probable transmission routes and other information crucial for interventions. The quick turn-around time for results for this adapted method is a critical aspect for epidemiological studies, predicting how cases are related and identifying new introductions. Our water sampling method has further reduced sampling time, providing an accurate insight into circulating strains, and an attractive adaptation for non-destructive sampling, dispensable of reagents or electrical equipment. The cost of the technologies proposed in this study are in constant decline, and many ONT providers are expanding globally. This approach will in turn increase the availability of whole genome sequences for important fish viruses in the database, for a better understanding of their mutations, virulence, and moreover fundamental for vaccine and drug design.

References

Bohara, Kailash, Amit K. Yadav, and Pabitra Joshi. 2022. “Detection of Fish Pathogens in Freshwater Aquaculture Using EDNA Methods.” *Diversity* 14 (12): 1015.

Chaput, Dominique L., David Bass, Md Mehedi Alam, Neaz Al Hasan, Grant D. Stentiford, Ronny van Aerle, Karen Moore, John P. Bignell, Mohammad Mahfujul Haque, and Charles R. Tyler. 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.

Debnath, S. C., Jamie McMurtrie, Ben Temperton, Jérôme Delamare-Deboutteville, Chadag Vishnumurthy Mohan, and Charles R. Tyler. 2023. “Tilapia Aquaculture, Emerging Diseases, and the Roles of the Skin Microbiomes in Health and Disease.” Aquaculture International: Journal of the European Aquaculture Society 31 (5): 2945–76.FAO. 2023. “FAO Country Profiles.” https://www.fao.org/countryprofiles/lifdc/en/.

Kellner, Max J., Jeremy G. Koob, Jonathan S. Gootenberg, Omar O. Abudayyeh, and Feng Zhang. 2019. “SHERLOCK: Nucleic Acid Detection with CRISPR Nucleases.” *Nature Protocols* 14 (10): 2986–3012.

Li, Daoliang, Xin Li, Qi Wang, and Yinfeng Hao. 2022. “Advanced Techniques for the Intelligent Diagnosis of Fish Diseases: A Review.” *Animals : An Open Access Journal from MDPI* 12 (21). https://doi.org/10.3390/ani12212938.

MacAulay, Scott, Amy R. Ellison, Peter Kille, and Joanne Cable. 2022. “Moving towards Improved Surveillance and Earlier Diagnosis of Aquatic Pathogens: From Traditional Methods to Emerging Technologies.” *Reviews in Aquaculture* 14 (4): 1813–29.

Mustafa, Mujahed I., and Abdelrafie M. Makhawi. 2021. “SHERLOCK and DETECTR: CRISPR-Cas Systems as Potential Rapid Diagnostic Tools for Emerging Infectious Diseases.” *Journal of Clinical Microbiology* 59 (3). https://doi.org/10.1128/JCM.00745-20.

Qin, Pan, Hetron Mweemba Munang’andu, Cheng Xu, and Jianjun Xie. 2023. “Megalocytivirus and Other Members of the Family Iridoviridae in Finfish: A Review of the Etiology, Epidemiology, Diagnosis, Prevention and Control.” Viruses 15 (6). https://doi.org/10.3390/v15061359.

Quick, Joshua, Nicholas J. Loman, Sophie Duraffour, Jared T. Simpson, Ettore Severi, Lauren Cowley, Joseph Akoi Bore, et al. 2016. “Real-Time, Portable Genome Sequencing for Ebola Surveillance.” Nature 530 (7589): 228–32.

Appendix

Awards and publications

Global Partnerships
University of Exeter
20th October 2023

Middle East and North Africa Fund

Dear Ms Alathari,

On behalf of the Middle East and North Africa Development Fund, I would like to thank you for applying to the 2023/24 round.

We are delighted to advise you that your application has been successful, and the panel has allocated in total £5,000 to support this project. Funding will be provided in accordance with the breakdown of costs indicated on your application form.

The panel were excited by the proposed collaboration between the University of Exeter and the University of Baghdad.

Funds from Exeter must be spent before 31st July 2024; all according to the budget submitted with the project proposal. Any major changes to the budget plan should be submitted to the Global Partnerships team as soon as possible.

Please email Global Partnerships (GP-Funding@exeter.ac.uk) by Friday 20th October to confirm acceptance of the award and that you have also read and agree to the conditions of funding as set out below. Your email in response will be taken as acceptance of the award on the terms stated.

We wish you the very best with your project and look forward to hearing further details about the impact and outcomes from this work. Please note a post-project report will be due 30th August 2024, as indicated in the conditions stated. We trust your project will be successful and look forward to seeing its fruitful results.

Conditions of Funding:

- Awardees have until 31st July 2024 to complete initiatives and expend the funding.
- Funding is only provided for the activities as detailed in your application, unless by prior agreement.
- Expenditure of the award must not exceed the value of the award.
- Details of the award may be listed on the Global Partnerships website.
- The Global Partnerships team may, from time to time, contact award holders to monitor progress and expenditure.
- Post-project reports will be due by 30th August 2024, with a follow up report assessing project outcomes (publications, funding outcomes, mobility outcomes etc) 18 months later (31st March 2026)

Aquaculture 558 (2022) 738367

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aquaculture

Relationships between pond water and tilapia skin microbiomes in aquaculture ponds in Malawi

Jamie McMurtrie [a,b,*,1], Shayma Alathari [a,b,1], Dominique L. Chaput [a], David Bass [b,c,d], Camerson Ghambi [e], Joseph Nagoli [f], Jérôme Delamare-Deboutteville [g], Chadag Vishnumurthy Mohan [g], Joanne Cable [h], Ben Temperton [a], Charles R. Tyler [a,b]

[a] Biosciences, Geoffrey Pope Building, University of Exeter, Exeter, Devon EX4 4HB, UK
[b] Sustainable Aquaculture Futures, University of Exeter, Exeter, Devon EX4 4QD, UK
[c] Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Weymouth Laboratory, Weymouth, Dorset DT4 8UB, UK
[d] Department of Life Sciences, The Natural History Museum, Cromwell Rd, London SW7 5BD, UK
[e] Department of Aquaculture and Fisheries Science, Lilongwe University of Agriculture and Natural Resources (LUANAR), P.O. Box 219, Lilongwe, Malawi
[f] WorldFish Malawi, Lilongwe, P.O. Box 30294, Malawi
[g] WorldFish, Bayan Lepas, Penang, Malaysia
[h] School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK

ARTICLE INFO

Keywords:
Aquaculture
Skin microbiome
Tilapia
Pond
Bacterial community
Eukaryotic community

ABSTRACT

Intensification of fish farming practices is being driven by the demand for increased food production to support a rapidly growing global human population, particularly in lower-middle income countries. Intensification of production, however, increases the risk of disease outbreaks and thus the likelihood for crop losses. The microbial communities that colonise the skin mucosal surface of fish are poorly understood, but are important in maintaining fish health and resistance against disease. This skin microbial community is susceptible to disruption through stressors associated with transport, handling and the environment of intensive practices, and this risks the propagation of disease-causing pathogens. In this study, we characterised the microbial assemblages found on tilapia skin — the most widely farmed finfish globally — and in the surrounding water of seven earthen aquaculture ponds from two pond systems in distinct geographic regions in Malawi. Metabarcoding approaches were used to sequence the prokaryotic and microeukaryotic communities. We found 92% of prokaryotic amplicon sequence variants were common to both skin and water samples. Differentially enriched and core taxa, however, differed between the skin and water samples. In tilapia skin, *Cetobacterium*, *Paucibacter*, *Pseudomonas* and Comamonadaceae were enriched, whereas, the cyanobacteria *Cyanobium*, *Microcystis* and/or *Synechocystis*, and the diatom *Cyclotella*, were most prevalent in pond water. Ponds that clustered together according to their water prokaryotic communities also had similar microeukaryotic communities indicating strong environmental influences on prokaryotic and microeukaryotic community structures. While strong site-specific clustering was observed in pond water, the grouping of tilapia skin prokaryotes by pond site was less distinct, suggesting fish microbiota have a greater buffering capacity against environmental influences. The characterised diversity, structure and variance of microbial communities associated with tilapia culture in Malawi provide the baseline for studies on how future intensification practices may lead to microbial dysbiosis and disease onset.

1. Introduction

Capture fisheries will not be able to satisfy the demand for seafood products from an ever-increasing human population with rising living standards (Henchion et al., 2017) combined with plateauing, and in some cases declining, wild fish stocks due to overfishing and ecosystem degradation (Link and Watson, 2019). Seeking to meet this demand for aquatic products, many aquaculture farming practices are undergoing intensification. Shifting from extensive to intensive and semi-intensive practices in aquaculture, however, is often associated with increased

* Corresponding author at: Biosciences, Geoffrey Pope Building, University of Exeter, Exeter, Devon EX4 4HB, UK.
E-mail address: jm1106@exeter.ac.uk (J. McMurtrie).
[1] JM and SA listed as co-first authors.

https://doi.org/10.1016/j.aquaculture.2022.738367
Received 7 December 2021; Received in revised form 11 April 2022; Accepted 15 May 2022
Available online 21 May 2022

microorganisms

an Open Access Journal by MDPI

Indexed in: PubMed | CITESCORE 6.4 | IMPACT FACTOR 4.5

CERTIFICATE OF ACCEPTANCE

Certificate of acceptance for the manuscript (**microorganisms-2789885**) titled:

The Citizen Phage Library: Rapid isolation of phages for the treatment of antibiotic resistant infections in the UK

Authored by:

Julie Fletcher; Robyn Manley; Christian Fitch; Christina Bugert; Karen Moore; Audrey Farbos; Michelle Michelsen; Shayma Alathari; Nicola Senior; Alice Mills; Natalie Whitehead; James Soothill; Stephen Michell; Ben Temperton

has been accepted in *Microorganisms* (ISSN 2076-2607) on 18 January 2024

MDPI Academic Open Access Publishing since 1996

Basel, January 2024

Shu-Kun Lin

Dr. Shu-Kun Lin
Publisher & President