Uncover aquaculture pathogens identity using Nanopore MinION

Virtual Nanopore Day, Southeast Asia, Wednesday 10th November 2021

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

WILDERLAB

MAHIDOL UNIVERSITY BIOTEC-NSTDA

WorldFish

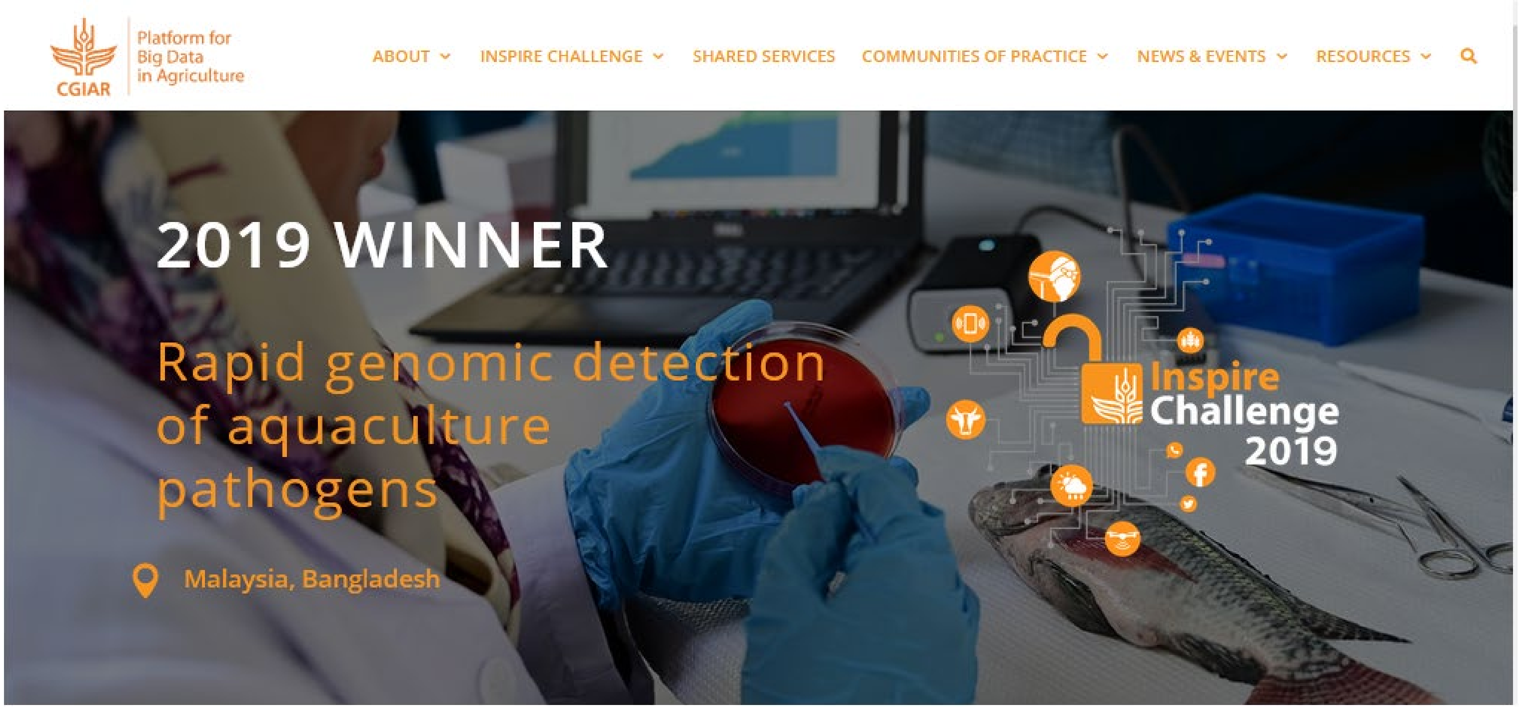
Aquaculture growth and future challenges

- Currently, aquaculture is one of the **fastest growing food producing sector** in the world.
- Global demand for aquatic foods will roughly **double by 2050** - aquaculture is predicted to meet most of this demand **complementing capture fisheries.**
- With the intensification, **the incidence of fish diseases** has also increased, which hinder the development and sustainability of aquaculture industry.

Public health' questions with aquatic foods

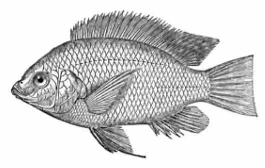
- Risk of zoonotic diseases?
- Risk of food-born diseases?
- Bacteria with antimicrobial resistance genes?
- Residues?

Inspire Challenge project



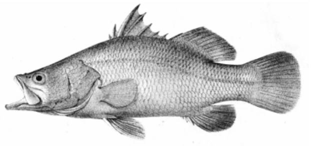
Hosts & pathogens from around the world

Tilapia
Oreochromis sp.



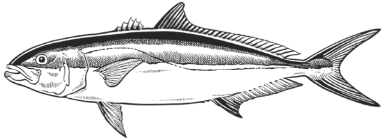
Streptococcus sp; Edwardsiella sp.; Aeromonas sp; Vibrio sp.; Infectious Spleen and Kidney Necrosis Virus (ISKNV); Tilapia lake virus (TiLV)

Barramundi
Lates calcarifer



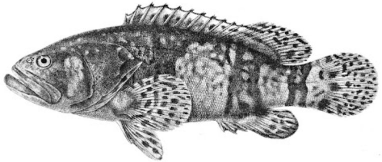
Vibrio sp; Streptococcus iniae

Kingfish
Seriola lalandi



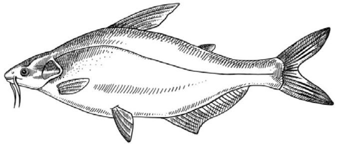
Photobacterium damselae

Qld giant grouper
Epinephelus lanceolatus



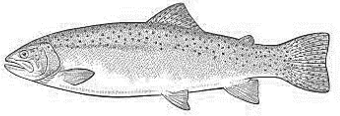
Streptococcus agalactiae

Pangasius
Pangasius sp.



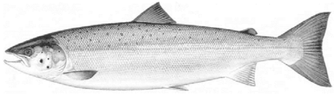
Aeromonas hydrophila

Rainbow trout
Oncorhynchus mykiss



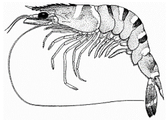
S. iniae; Yersinia ruckeri

Atlantic salmon
Salmo salar



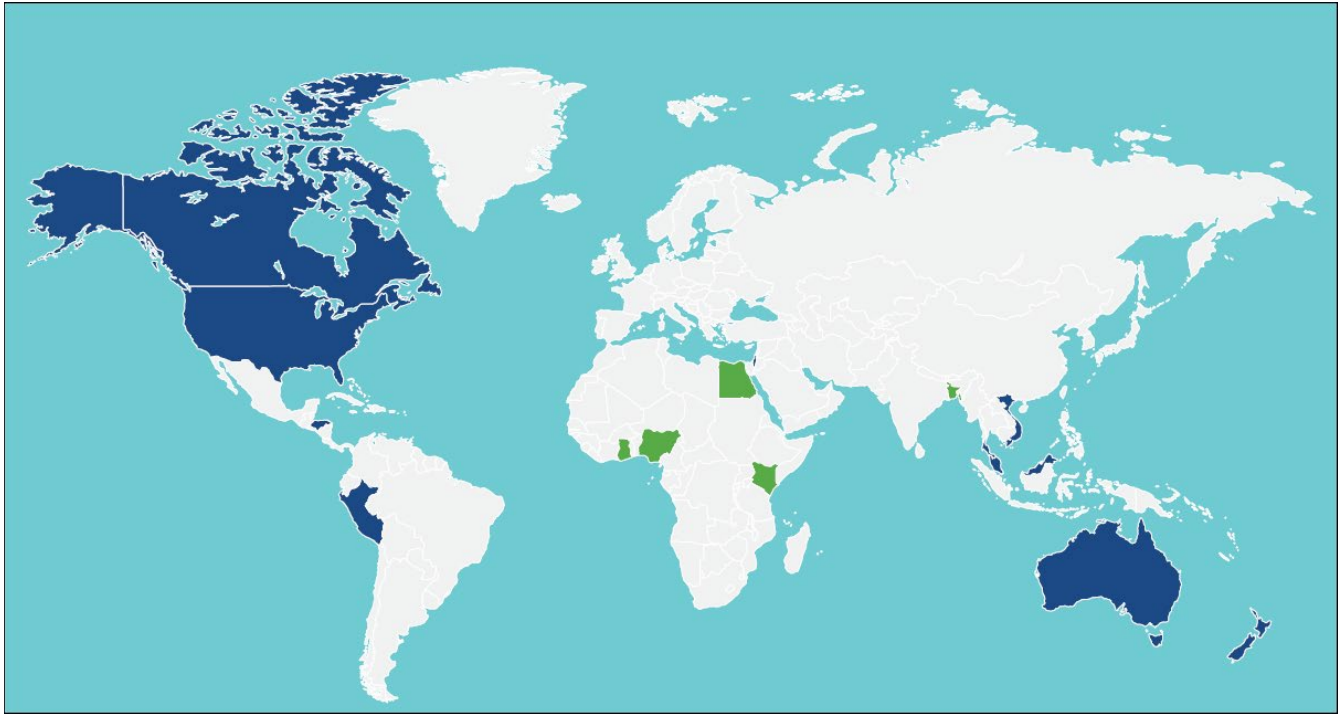
Y. ruckeri; Tenacibaculum maritimum

Shrimps
Penaeus monodon/vannamei



Vibrio sp; *Aeromonas* sp.

Origins of the pathogens



Aquaculture pathogens sequenced in this project

Honduras, Ecuador, Peru, USA, Canada, Israel, Australia, New-Zealand, Vietnam, Thailand, Malaysia

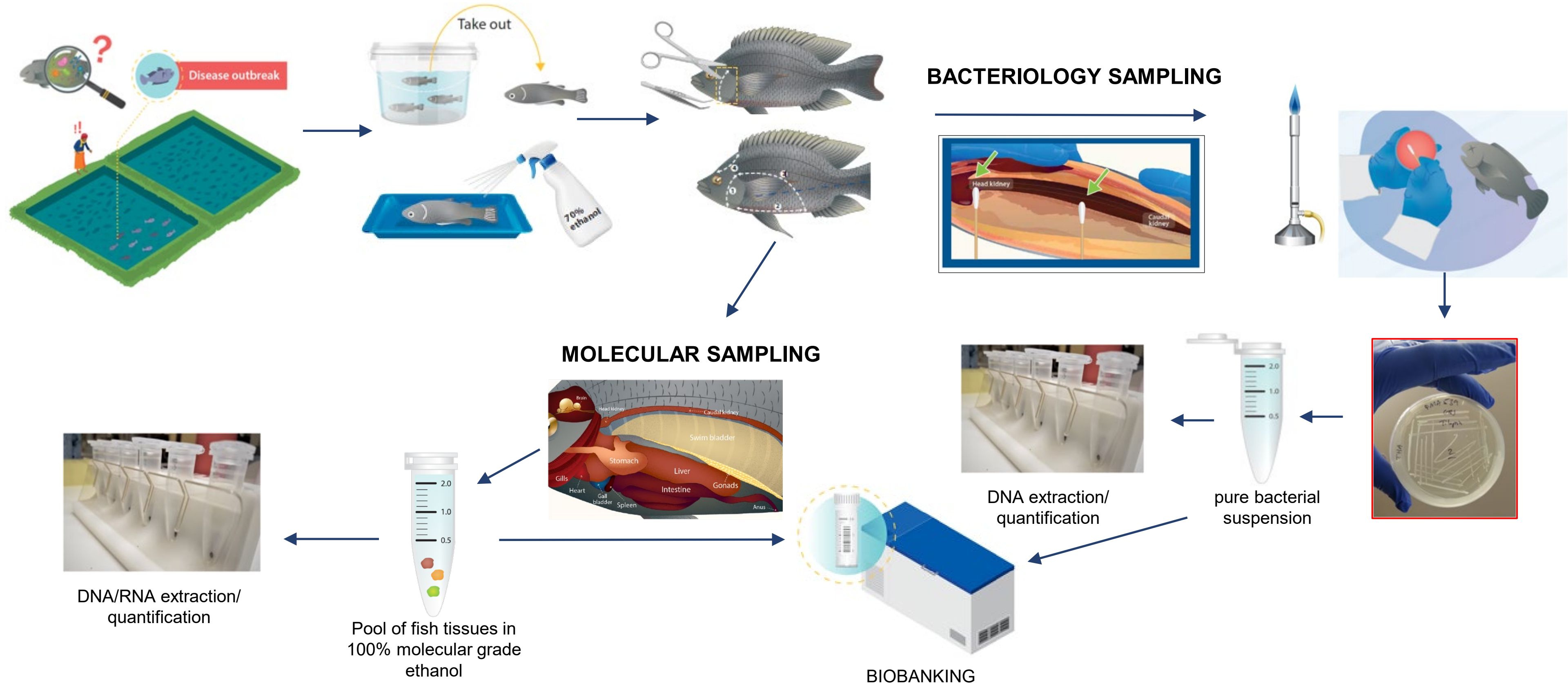
Future works using Nanopore

Bangladesh, Egypt, Ghana, Nigeria and Kenya

Sample collection & DNA extraction

Disease outbreak: samples collection

Sampling from diseased fish

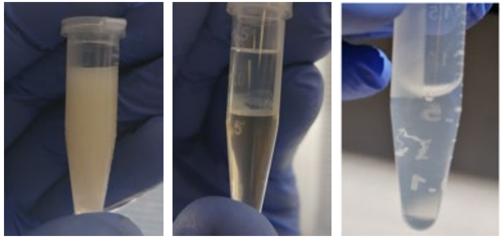


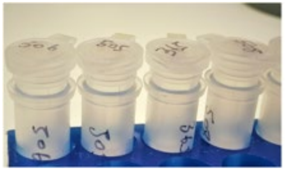
DNA extraction - cells lysis methods

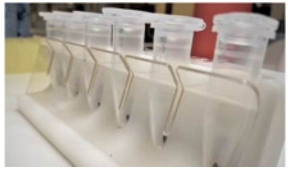
	Advantages	Disadvantages
SDS + Proteinase K	Overall good performer for DNA extraction	May lead to co-precipitation of carbohydrate
CTAB	Sample with high polysaccharide (e.g. capsulated/mucoid microbe)	CTAB is detrimental to environmental
Lysozyme pretreatment +SDS	Suitable for gram positive microbe	May not work across all gram positive microbe
Mechanical disruption (Bead beating)	Samples with tough/thick cell wall	Can lead to fragmentation of DNA (higher smearing)

DNA purification methods

	Advantages	Disadvantages
Phase separation (chloroform extraction)	High DNA yield and integrity Cheap	Requires equipment, generates toxic chemicals (chloroform and phenol), relatively time consuming
Column-based separation	Fast and convenient.	Requires equipment, expensive, often low-yield and molecular weight and require multiple centrifuge steps. Not scalable
Magnetic silica/ carboxylated beads	Scalable, fast and convenient, only magnet needed, high DNA yield and integrity	High cost of commercially produced beads



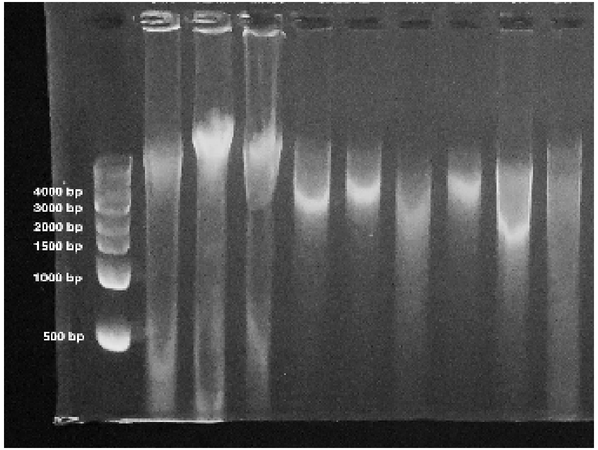




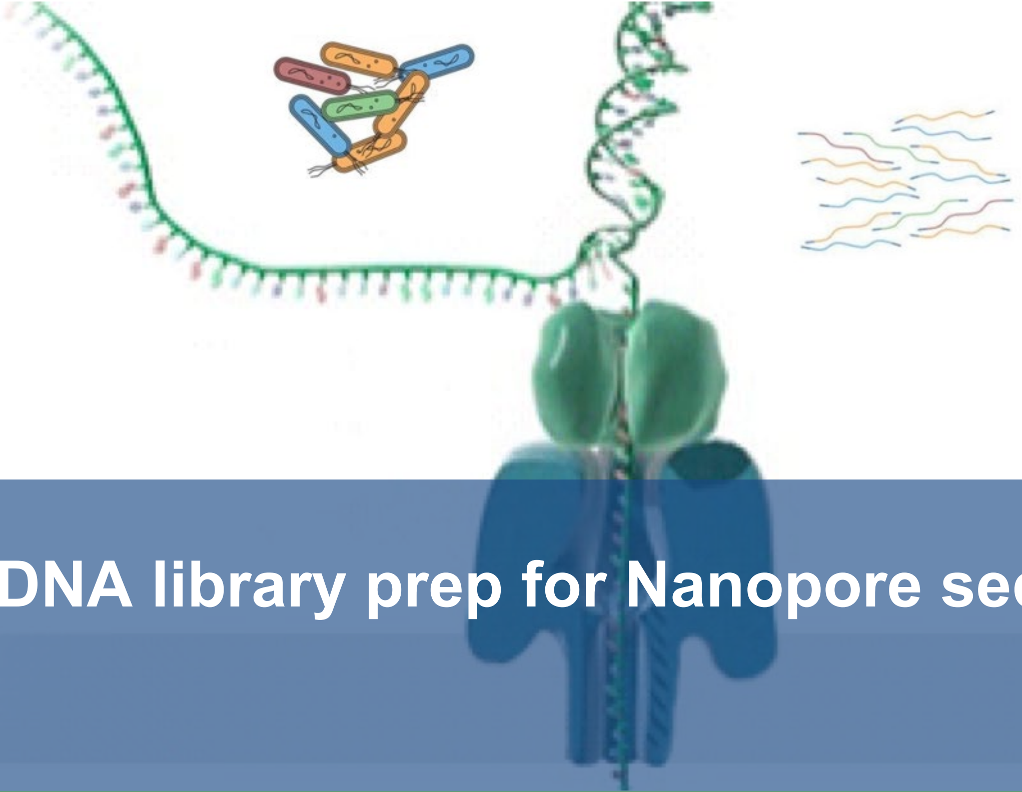
DNA concentration & quality control

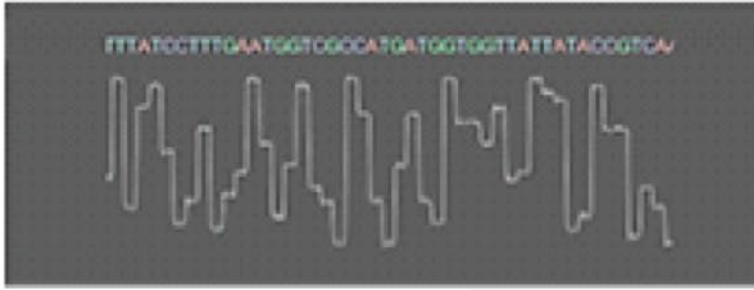
Qubit Fluorometer.
Measure the concentration of dsDNA based on the fluorescence emitted by proprietary dsDNA-specific binding dye.

Nanodrop spectrophotometer to estimate DNA purity based on absorbance measurement.



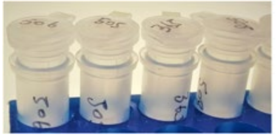
DNA integrity assessed by agarose gel electrophoresis.





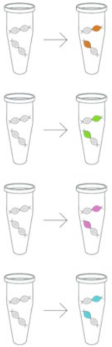
DNA library prep for Nanopore sequencing

DNA samples for Nanopore sequencing

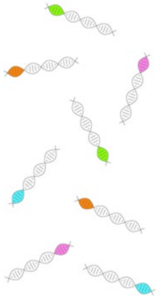


- Bacterial high molecular weight genomic DNA → **whole genome sequencing**
- DNA clean up from gel electrophoresis → **amplicon sequencing**

- **Ligation sequencing kit**
- **Rapid barcoding kit**

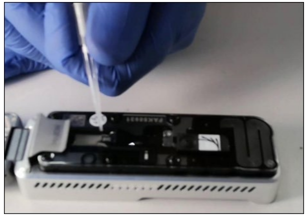


Barcode multiple samples



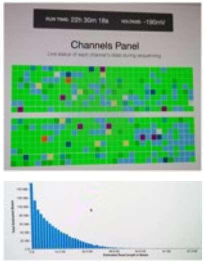
Pool for sequence

Flow cell priming

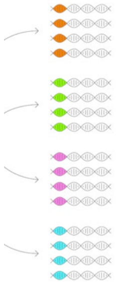


Library loading on flow cell

MinION sequencing

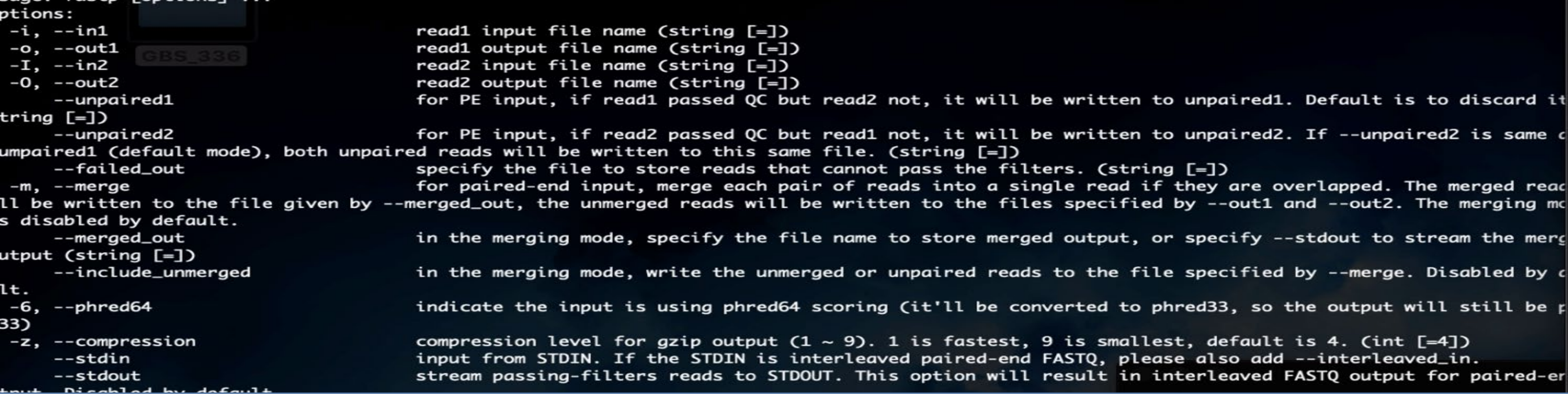


Visualization of sequencing run



Demultiplex barcoded samples

Raw Nanopore reads for Bioinformatics analysis



Bioinformatics: Simplifying Big Data

Linux command line

```
  --threads [N]   Use this many BLAST+ threads [1].
DATABASES
  --setupdb       Format all the BLAST databases.
  --list          List included databases.
  --datadir [X]   Databases folder [/home/gan/miniconda3/envs/abricate/db].
  --db [X]        Database to use [ncbi].
OUTPUT
  --noheader      Suppress column header row.
  --csv           Output CSV instead of TSV.
  --nopath        Strip filename paths from FILE column.
FILTERING
  --minid [n.n]   Minimum DNA %identity [75].
  --mincov [n.n]  Minimum DNA %coverage [0].
MODE
  --summary       Summarize multiple reports into a table.
DOCUMENTATION
  https://github.com/tseemann/abricate
(abricate) gan@Gan:/mnt/c/Ubuntu_Shared/Jerome_WorkShop/keep/Presentation_Rename/Vpara$ abricate *.contigs.fasta
Using nucl database ncbi:  5283 sequences -  2020-Feb-20
#FILE   SEQUENCE        START   END     STRAND  GENE    COVERAGE        COVERAGE_MAP    GAPS    %COVERAGE       %IDENTITY
ATABASE ACCESSION       PRODUCT RESISTANCE
Processing: NF_1.contigs.fasta
Found 3 genes in NF_1.contigs.fasta
NF_1.contigs.fasta      NODE_11_length_192793_cov_91.644690     64786   65250   -       tet(34) 1-465/465       ========/======
/2      99.78   83.48   ncbi    NG_048129.1     oxytetracycline resistance phosphoribosyltransferase domain-containing protein Tet(34)
ETRACYCLINE
NF_1.contigs.fasta      NODE_1_length_601117_cov_80.593360      329283  330134  -       blaCARB-47      1-852/852       ===============
/0      100.00  99.18   ncbi    NG_050564.1     carbenicillin-hydrolyzing class A beta-lactamase CARB-47        BETA-LACTAM
NF_1.contigs.fasta      NODE_4_length_470469_cov_78.618048      189617  191218  +       tet(35) 1-1602/1602     ===============
/0      100.00  99.19   ncbi    NG_063830.1     tetracycline efflux Na+/H+ antiporter family transporter Tet(35)        TETRACYCLINE
Processing: UMP_1.contigs.fasta
Found 3 genes in UMP_1.contigs.fasta
UMP_1.contigs.fasta     NODE_1_length_862548_cov_114.102298     519864  520715  +       blaCARB-33      1-852/852       ===============
/0      100.00  99.53   ncbi    NG_048737.1     carbenicillin-hydrolyzing class A beta-lactamase CARB-33        BETA-LACTAM
UMP_1.contigs.fasta     NODE_4_length_471855_cov_114.677020     280941  282542  -       tet(35) 1-1602/1602     ===============
/0      100.00  98.94   ncbi    NG_063830.1     tetracycline efflux Na+/H+ antiporter family transporter Tet(35)        TETRACYCLINE
UMP_1.contigs.fasta     NODE_9_length_179391_cov_118.384114     64784   65248   -       tet(34) 1-465/465       ========/======
/2      99.78   83.26   ncbi    NG_048129.1     oxytetracycline resistance phosphoribosyltransferase domain-containing protein Tet(34)
ETRACYCLINE
Processing: UMP_3.contigs.fasta
Found 3 genes in UMP_3.contigs.fasta
UMP_3.contigs.fasta     NODE_1_length_876017_cov_54.348261      685103  686704  -       tet(35) 1-1602/1602     ===============
/0      100.00  98.94   ncbi    NG_063830.1     tetracycline efflux Na+/H+ antiporter family transporter Tet(35)        TETRACYCLINE
UMP_3.contigs.fasta     NODE_2_length_862548_cov_54.500130      341834  342685  -       blaCARB-33      1-852/852       ===============
/0      100.00  99.53   ncbi    NG_048737.1     carbenicillin-hydrolyzing class A beta-lactamase CARB-33        BETA-LACTAM
UMP_3.contigs.fasta     NODE_9_length_179145_cov_60.183578      64784   65248   -       tet(34) 1-465/465       ========/======
/2      99.78   83.26   ncbi    NG_048129.1     oxytetracycline resistance phosphoribosyltransferase domain-containing protein Tet(34)
ETRACYCLINE
Processing: UPM_1.contigs.fasta
```

Web-based tools

- Upload **Fasta (or FastQ)** files
- Open-source = Free to use
- Peer-reviewed
- Doesn't rely on your hardware
- Internet Connection (cloud)

Molecular serotyping on selected strains (GBS)

BLAST-based approach
(input = Fasta: assembled draft or complete genome)

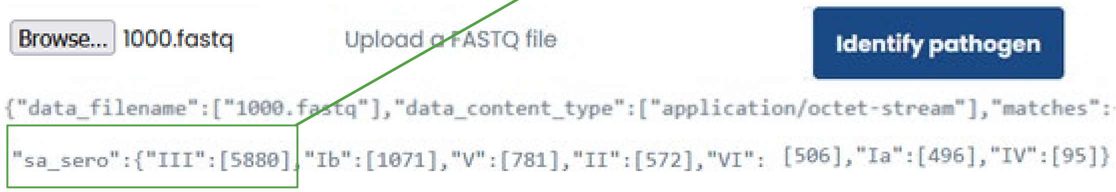
https://github.com/swainechen/GBS-SBG

```
# Name   Serotype
NF_3     GBS-SBG:III-4
# Name   Serotype
UPM_3    GBS-SBG:III-4
```

SAME RESULTS

Diagnostic k-mer-based approach
(input = Fastq > 400 read sequences - Assembly-Free and Real-Time)

In development by WorldFish, UQ and Wilderlab



Application of NGS in aquaculture

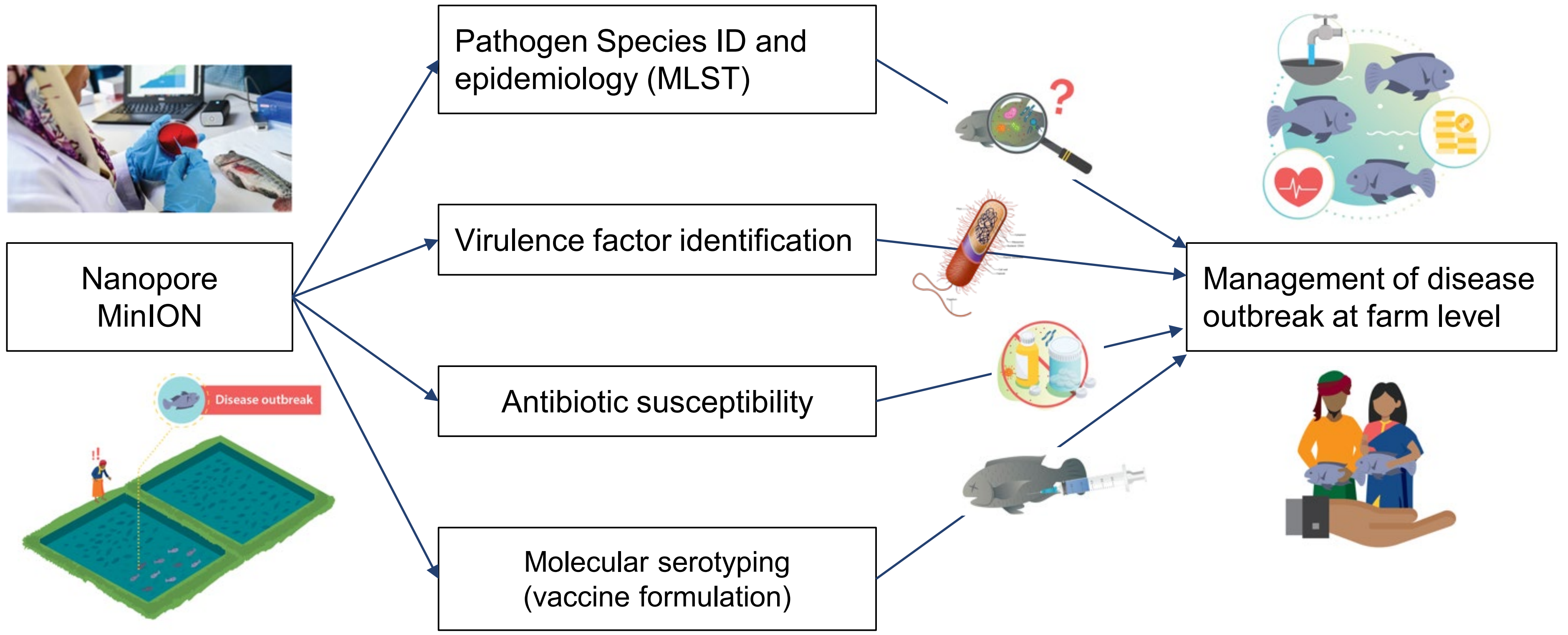
Fish farmers' questions during disease outbreak

What is killing my fish?

How can I treat them?

How can I prevent disease in the next crop?

Genome-based diagnosis of pathogens in aquaculture



Latest publication

J Fish Dis. 2021 Oct;44(10):1491-1502. doi: 10.1111/jfd.13467. Epub 2021 Jun 8.

Rapid genotyping of tilapia lake virus (TiLV) using Nanopore sequencing

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

Affiliations + expand

PMID: 34101853 DOI: 10.1111/jfd.13467

https://onlinelibrary.wiley.com/doi/10.1111/jfd.13467

Thank You

WorldFish

This work was undertaken as part of

CGIAR RESEARCH PROGRAM ON Fish
Led by WorldFish

CGIAR Platform for Big Data in Agriculture

Inspire Challenge

In partnership with

The University of Queensland Australia

WILDERLAB

Mahidol University BIOTEC-NSTDA