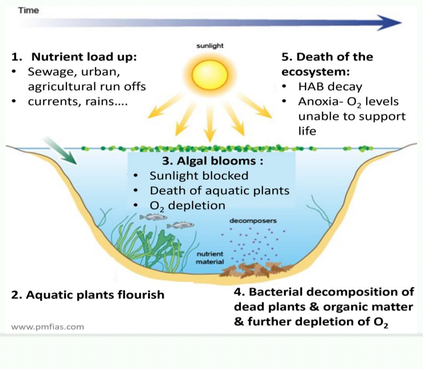
CAN THIRD GENERATION SEQUENCING RAPIDLY DETECT HARMFUL ALGAL BLOOMS(HABs) ?

[1]Christaline George*, [1]Caroline Chénard, [1]Avneet Kaur, [4]Aurore Trottet, [2]Sandric Chee Yew Leong, [5]Daniel Vaulot, [1]Adriana Lopes Dos Santos, [1,3]Federico M Lauro

<*christal002@e.ntu.edu.sg>

INTRODUCTION

Phytoplankton cause algal blooms, some of which are harmful - HABs



Massive amounts of fish kills due to past events of HABs in Singapore

Dinoflagellate blooms

- 2009- 200,000 farmed fishes
- 2014- 500 **tonnes** of fish from 53 farms
- 2015 - 600 **tonnes** of fish in 77 farms/ **loss at $1.3 million/farmer**

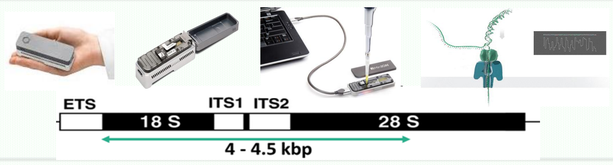
- **HAB mitigation & control**: Early & precise characterization of bloom causing organisms.
- **Challenge**: **Tedious conventional methods** requiring expertise and slow analysis.
- **Solution: Third Generation Sequencing with MinION™** - relatively lower costs, precise & rapid analyses.

OBJECTIVES

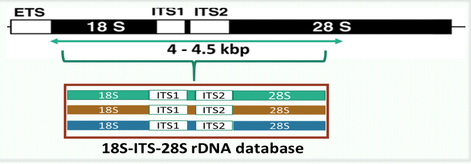
1. Design and optimize:

I. Methodology for MinION sequencing of eukaryotic microalgal rDNA unit (18S - 28S)

II. Bioinformatics pipeline for error correction and analyses of sequence



2. Build a reference database containing these rDNA error corrected sequences



SEQUENCED SAMPLES

Diatoms	Dinoflagellates	Chlorophytes
Coscinodiscus sp.	*Scrippsiella* sp.	*Tetraselmis* sp.
Cyclotella sp.	*Karenia mikimotoi*	*Ostreococcus tauri*
Phaeodactylum tricornutum	*Karlodinium cf veneficum*	*Picochlorum sp.SENEW3*

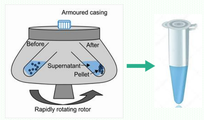
EXPERIMENTAL METHODS

1 Sample collection, filtration and preparation

Environmental samples

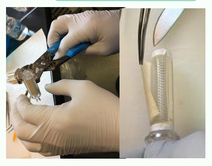
1L filtered through 0.22µm sterivex

Single species

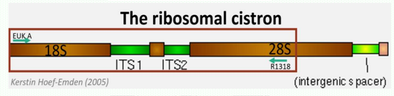


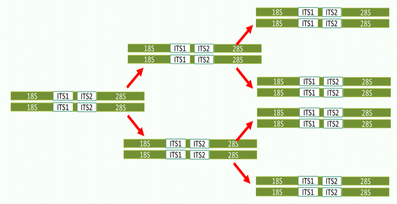
20-30mL of cultures
500g for 10mins

2 DNA extraction

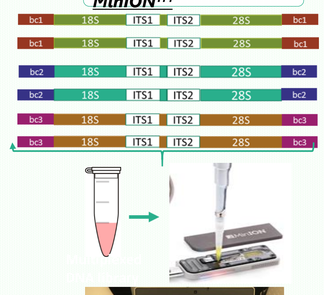


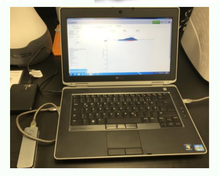
3 PCR to amplify rDNA cistron





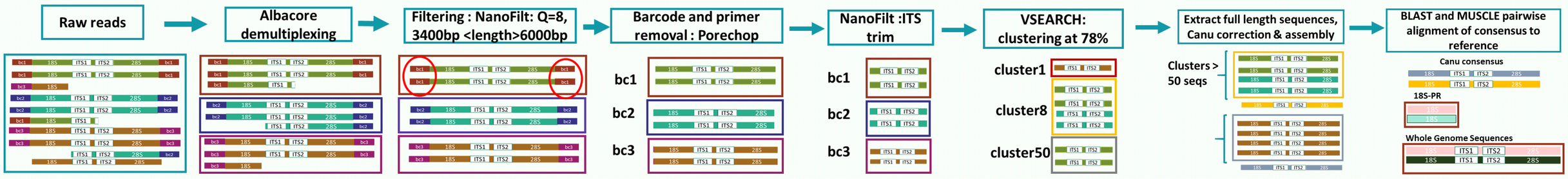
4 Barcoding, multiplexing and Amplicon sequencing using *MinION™*





RESULTS

BIOINFORMATICS PIPELINE –ERROR COORECTION AND READS ANALYSES



Single species sequences analyses

Table 1: Similarity of canu (MInION) sequences compared to PR² references (containing only 18S sequences) based on the pairwise distance matrix using MEGA version 7.2

	Coscinodiscus sp.	*Cyclotella* sp.	*Scripsiella* sp.	*Karlodinium veneficum*	*Karenia mikimotoi*	*Tetraselmi* sp.
Accuracy of consensus compared to PR2	99.50%	98.95%	99.70%	99.50%	99.00%	99.30%

Table 2: Similarity of canu (MInION) sequences compared to NCBI-WGS references (full length rDNA sequences) based on the pairwise distance matrix using MEGA version 7.2

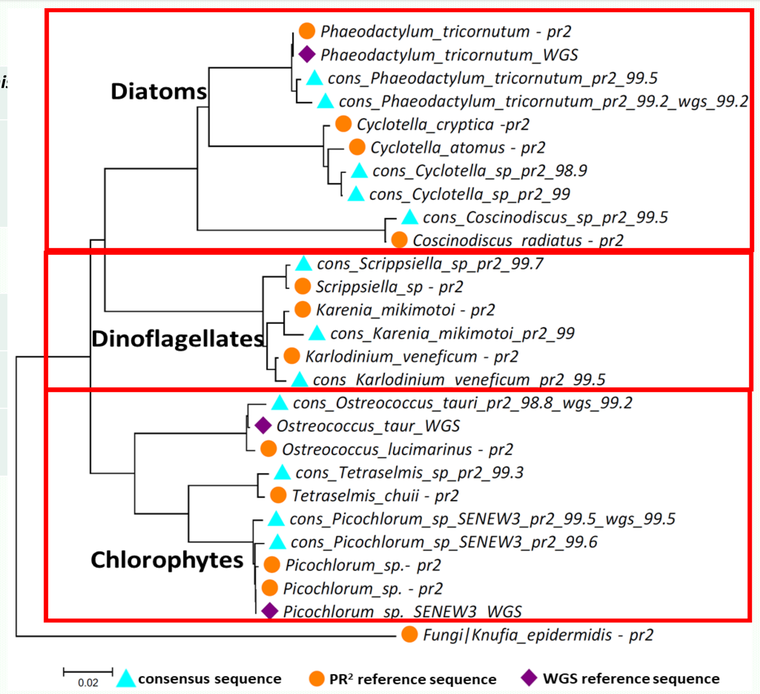
	Picochlorum sp SENEW3	*Ostreococcus tauri*	*Phaeodactylum tricornutum*
Accuracy of consensus compared to PR2	99.55%	98.80%	99.20%
Accuracy of consensus compared to WGS	99.50%	99.20%	99.20%

Table 3: Reduction in error rates and increased Q-scores of MinION

Single species sequences	Sequencing error rate	Q-score
Raw reads	**12 – 22%**	6.6 - 9
Error corrected reads	**≤ 0.7%**	21.5

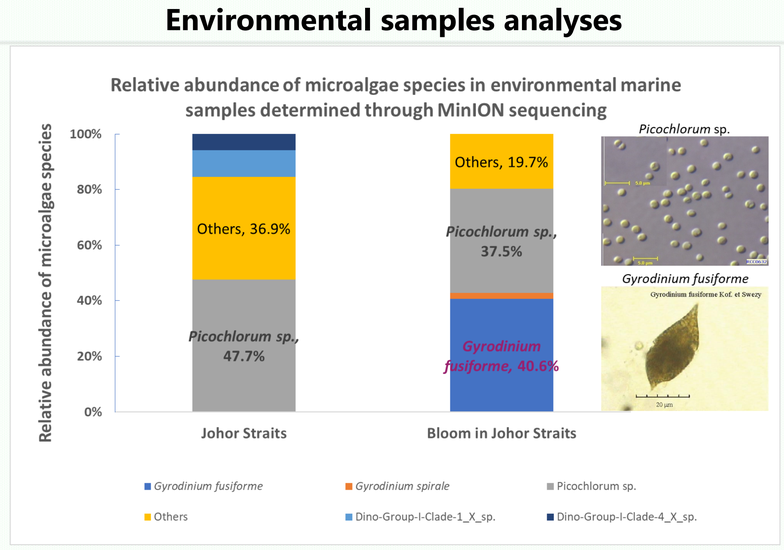
Q (Phred score) = -10 $\log_{10}$ P

Average accuracy of the reads = 99.3%

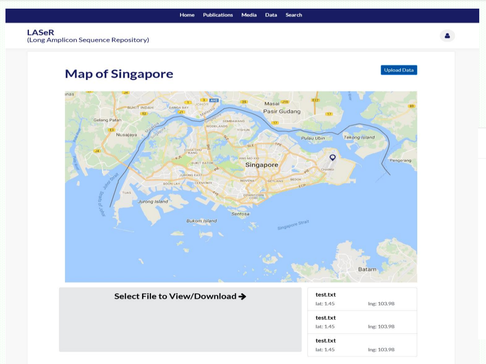


Phylogenetic tree constructed using the Neighbor Joining method of canu consensus sequences clustered with PR2 and WGS reference sequences

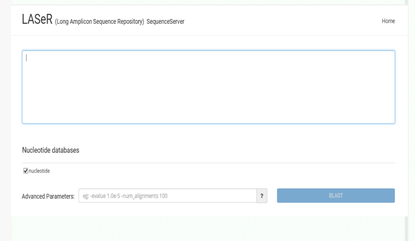
Environmental samples analyses



'LASeR ' :Long Amplicon Sequence Repository- error corrected rDNA sequences



- Deposit references
- Search & download refernces



Microscopy Vs MinION™

Factor	Microscope	MinION™
Analysis time	~5-48hrs. per sample	~<2 hrs. for 96 samples/MinION
Identification level	Phylum, genus and sometimes species level.	Phylum, genus and species level identification
Detection limit	2 cells/mL	<10 cells/mL
Accuracy	Dependent on taxonomic expertise	Taxonomist not required due to the availability of a bioinformatics pipeline

CONCLUSION & FUTURE DIRECTIONS

***MinION™* sequencing technology - a great potential HABs monitoring- RAPID & PRECISE DETECTION**

- Portable sequencer for field applications enabling rapid taxonomic characterization.
- Increase error correction iterations with Canu for read quality improvement for clustering at species and subspecies level.
- Series dilution and qPCR for quantification with MinION.
- Target toxin genes in phytoplankton through the CRISPR-CAS9 targeting mechanism.

REFERENCES

Piles of dead fish at Pasir Ris beach. (2015, March 1). Retrieved July 20, 2017, from http://www.todayonline.com/singapore/piles-dead-fish-pasir-ris-beach.

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

Koren S, Walenz BP, Berlin K, Miller JR, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Research. (2017

ACKNOWLEDGEMENT

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its Marine Science Research and Development Programme Award No. MSRDP-P13.
I would like to thank Edwin Tat Seng, IT Senior Engineer at NTU for all the help with installation of the required software for data analysis and for all the associated troubleshooting procedures.,

1. Asian School of the Environment, Nanyang Technological University, Singapore (NTU). 2. Tropical Marine Science Institute (TMSI), National University of Singapore (NUS) 3. Singapore Centre for Environmental Life Sciences Engineering (SCELSE) 4. DHI Water & Environment(s) Pte Ltd. 5. Station Biologique De Roscoff & CNRS (The French National Centre For Scientific Research) 6 Oxford Nanopore Technologies