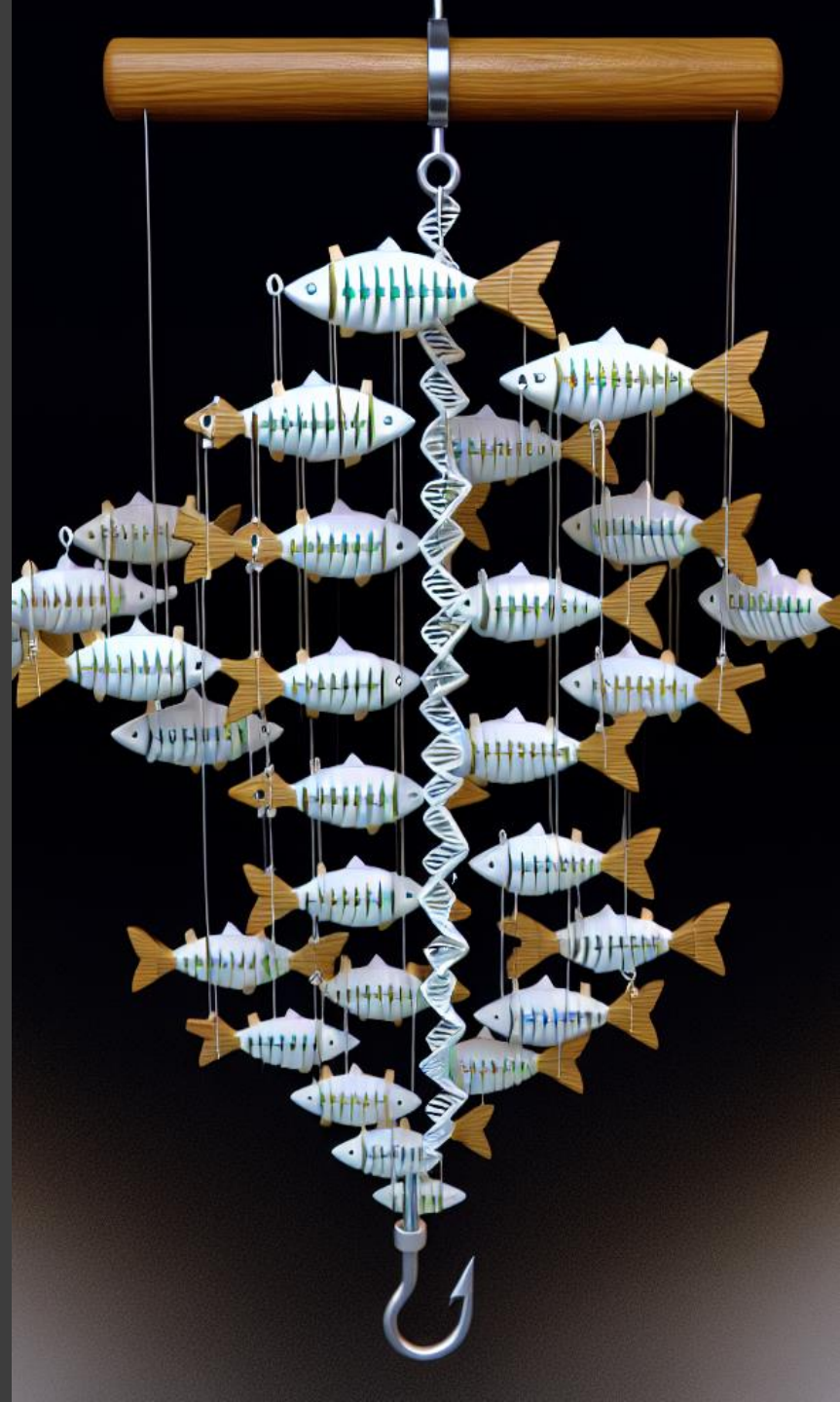
Fishing for mitogenomes with long-read PCR and long-read sequencing: What's the catch?

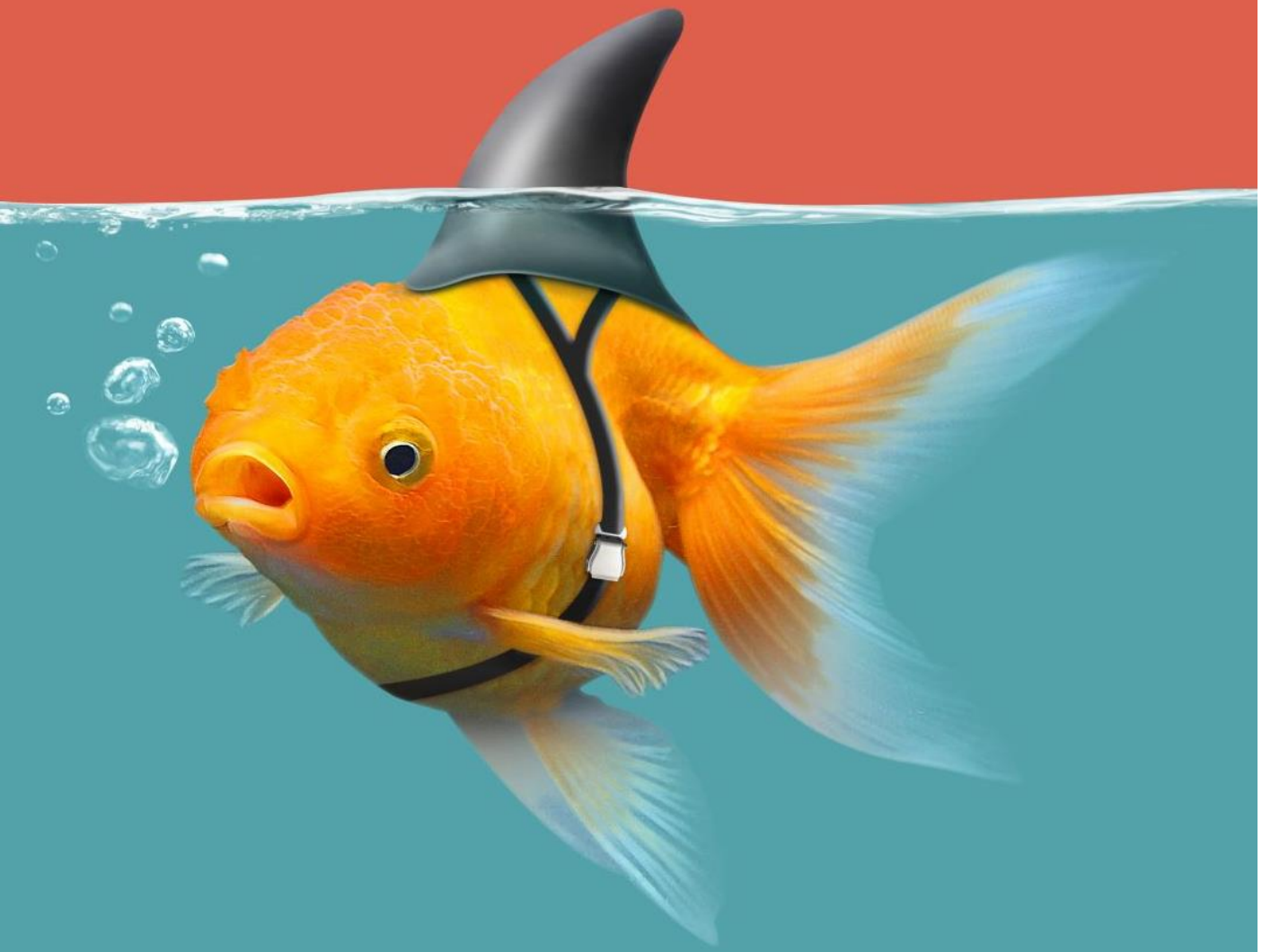
Marc JJ Chang, A Carré de Malberg, WWR Chan, YCA Ip, Z Jaafar & D Huang

Lee Kong Chian Natural History Museum, National University of Singapore

2nd SeDNA Conference, Wellington, New Zealand

20 February 2025





eDNA METABARCODING

- Is a powerful approach for monitoring fish communities
- Requires robust sequence databases for identification

DISCONNECT IN DATABASES?

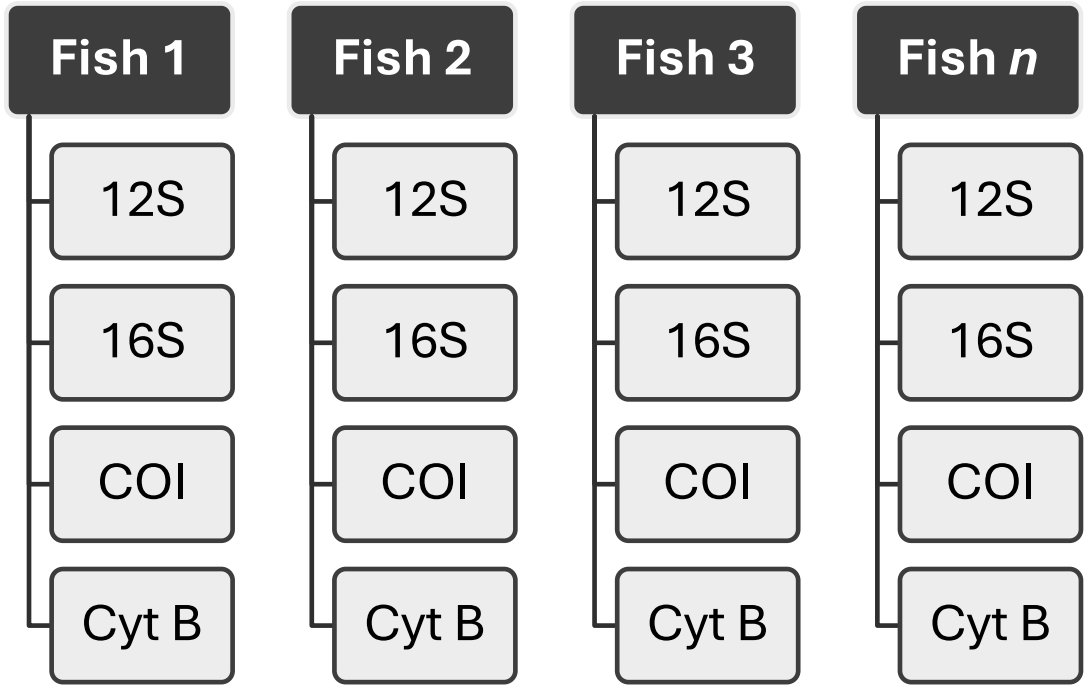
- Most databases built with 1–2 markers
- Fish eDNA studies use 12S rRNA, 16S rRNA, COI, Cyt B
- Shift towards multi-marker eDNA metabarcoding strategies
- Can we better future-proof our reference databases?

Xiong et al., 2022

Table 2
Summary of local databases.

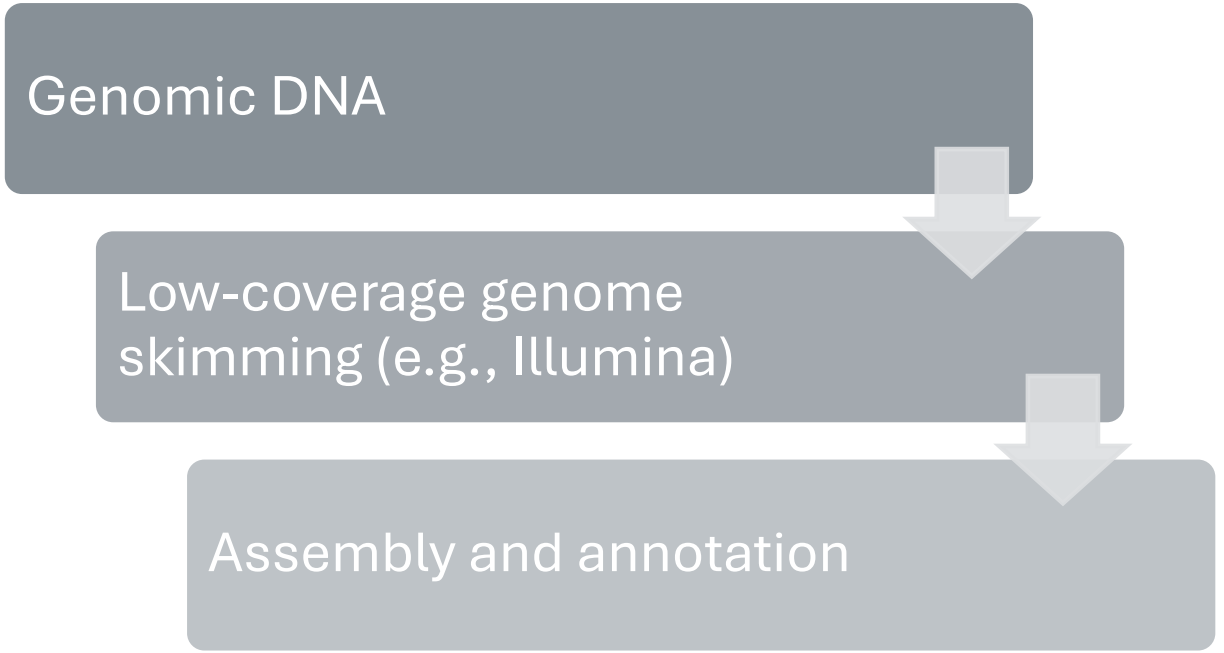
The location of targeted fishes	N of fish species	Markers	Reference
Canadian	190	COI	Hubert et al. (2008)
North America	752	COI	(April et al., 2011)
South Africa	53	COI	Cawthorn et al. (2011)
Argentine Marine and Brackish Waters	119	COI	Mabragana et al. (2011)
Southeastern Nigeria	70	COI	Nwani et al. (2011)
Sao Paulo State, Brazil	135	COI	Ribeiro et al. (2012)
Mediterranean Biodiversity Hotspots	498	COI	Geiger et al. (2014)
North European shelf	93	COI	Knebelsberger et al. (2014)
Mediterranean Sea	218	COI	Landi et al. (2014)
Germany	92	COI	Knebelsberger et al. (2015)
Okinawa Churaumi Aquarium	180	12S	Miya et al. (2015)
Northeastern coast of Brazil	78	COI	Brandao et al. (2016)
Lower Parana River	79	COI	Diaz et al. (2016)
The U.K	67	12S,Cytb	Hanfling et al. (2016)
Europe	86	12S	Valentini et al. (2016)
Java and Bali	159	COI	Dahruddin et al. (2017)
Indo-Myanmar	109	COI	Barman et al. (2018)
The South China Sea	272	COI	Hou et al. (2018)
Northern Western Ghats of India	81	COI	Patil et al. (2018)
Guianese freshwater	231	12S	Cilleros et al. (2019)
French Polynesian shore	540	COI	Delrieu-Trottin et al. (2019)

(1) Generate references sequences for each marker?



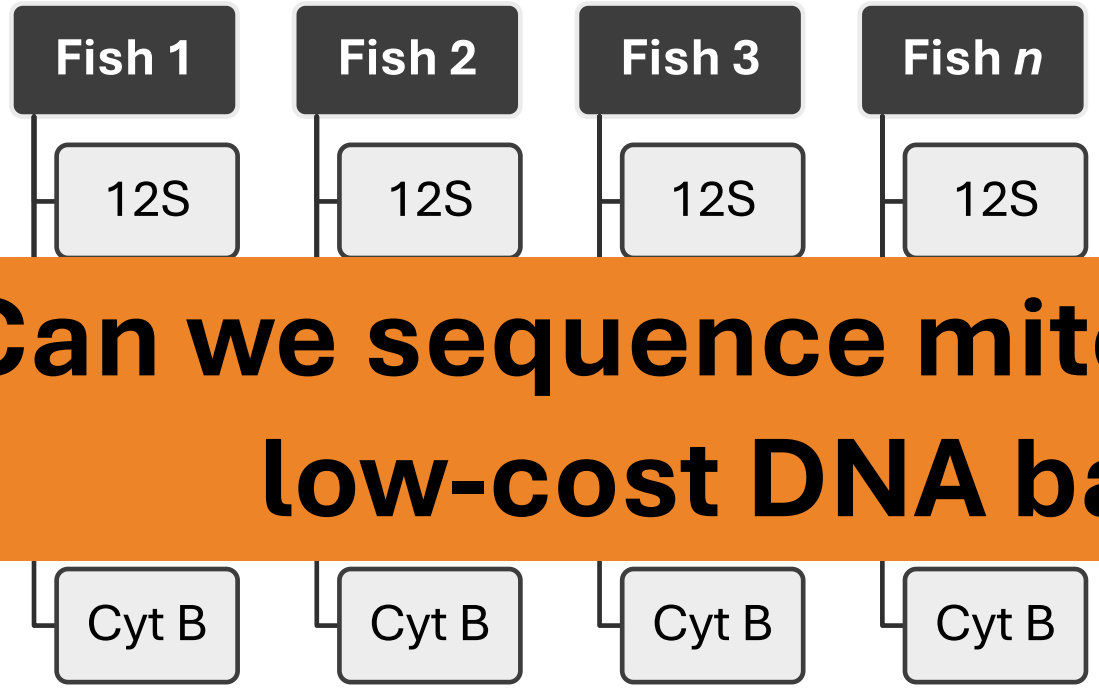
- PRO: Low-cost amplicon sequencing
- CON: A lot of PCRs!
- CON: Does not account for changes in marker choice

(2) Sequence the entire mitogenome?

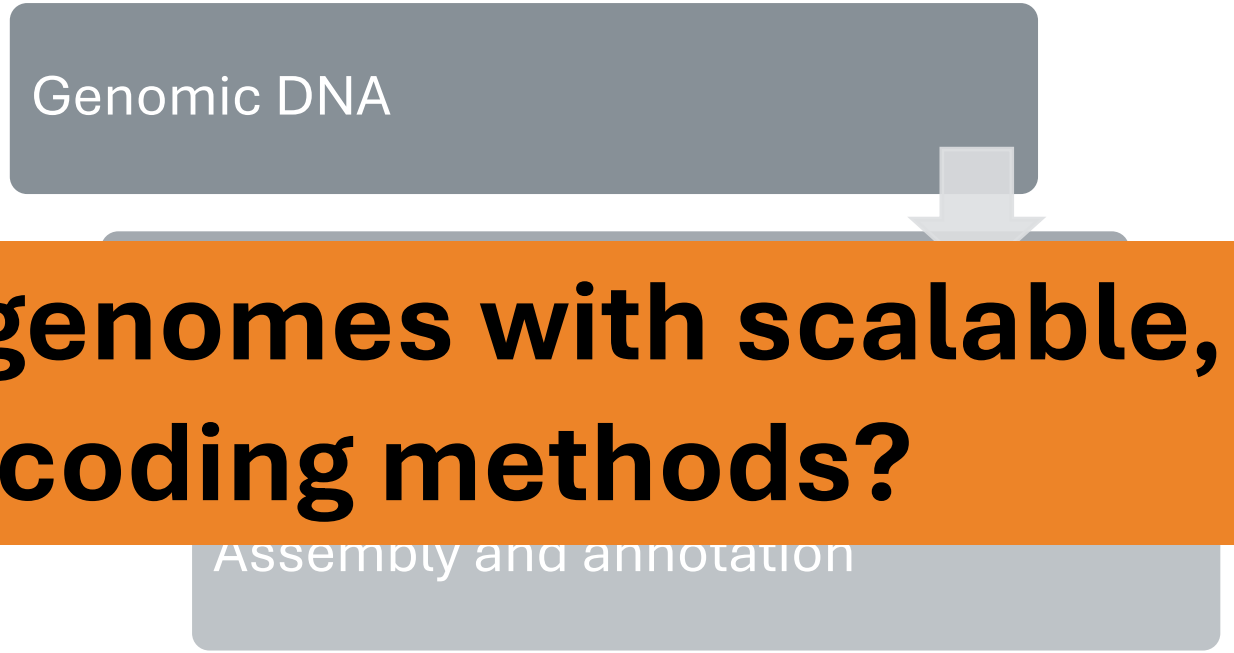


- PRO: Capture all eDNA regions
- CON: Genome skimming is *random*
- CON: Laborious to scale up

(1) Generate references sequences for each marker?



(2) Sequence the entire mitogenome?



Can we sequence mitogenomes with scalable, low-cost DNA barcoding methods?

- **PRO: Low-cost amplicon sequencing**
- CON: A lot of PCRs!
- CON: Does not account for changes in marker choice

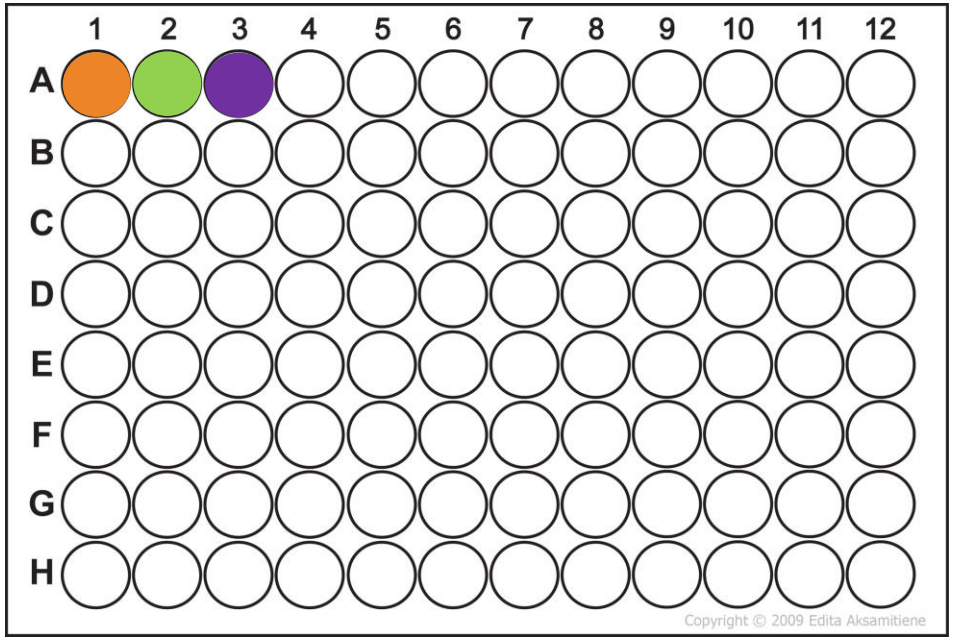
- **PRO: Capture all eDNA regions**
- CON: Genome skimming is *random*
- CON: Laborious to scale up

BARCODING THE MITOGENOME

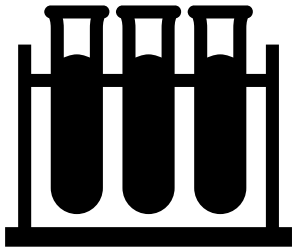
- Long-range PCR (LRPCR) with Actinopterygii16S primers (Deiner et al., 2017)

- Sample multiplexing

- "Close the loop" with 16S DNA barcodes using Palumbi (1996) primers
- Custom 9bp tagged primer strategy (Srivathsan et al., 2021)



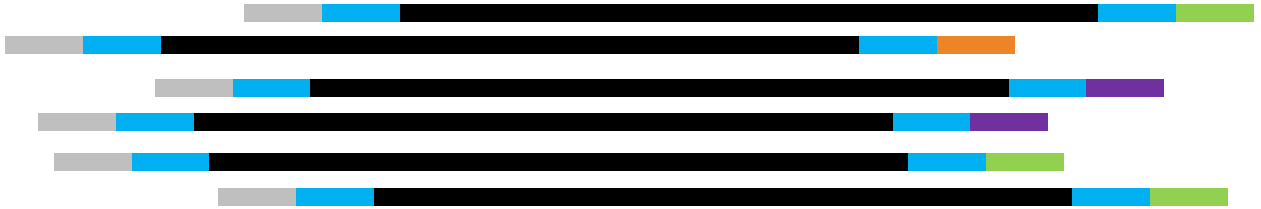
HARNESSING LONG-READ SEQUENCERS



Pool samples into libraries

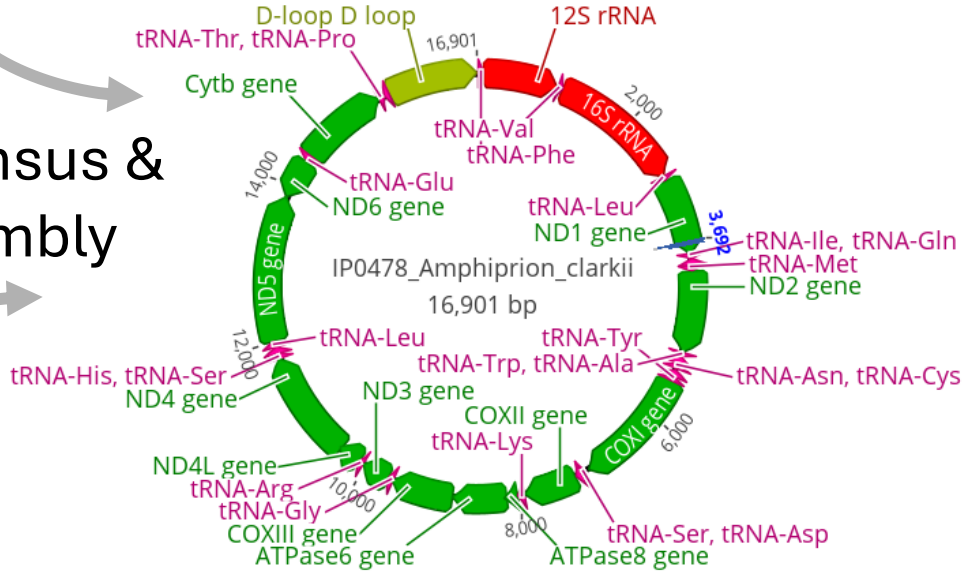
Oxford Nanopore Technologies

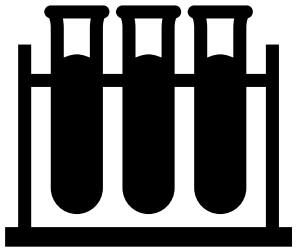
Sequence



- Sequence **long** f
- Reads are nonsea

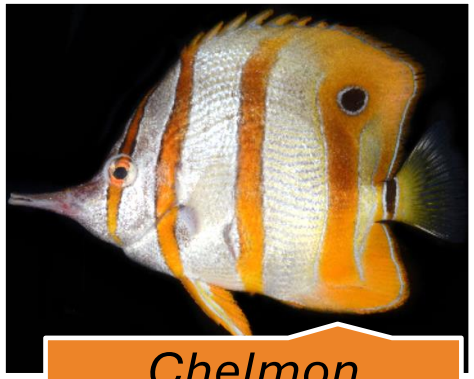
Consensus & Assembly



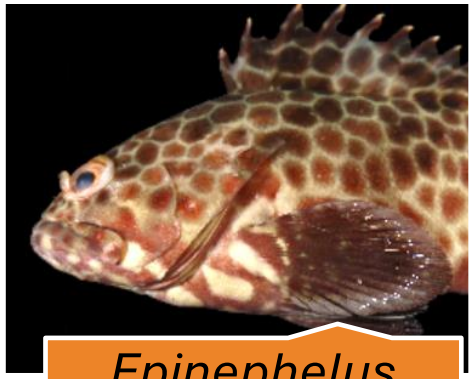


181 MITOGENOMES SEQUENCED SO FAR

Images from Biodiversity of Singapore



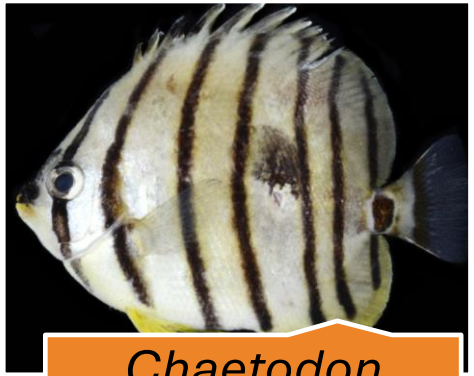
Chelmon rostratus



Epinephelus quoyanus

Gerres oyena

Amphiprion frenatus

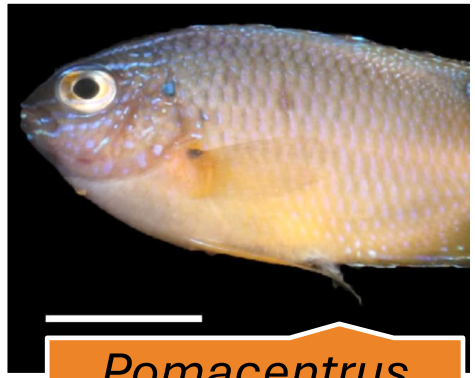


Chaetodon octofasciatus

Dischistodus prosopotaenia

Dactylopus dactylopus

Cryptocentrus melanopus



Pomacentrus tripunctatus

SEQUENCING COSTS

- Nanopore reagents and consumables account for bulk of the sequencing costs

RUN	DESCRIPTION	COST (USD)
1	• 1x library prep reaction • 1x Flongle flow cell • 35 amplicons	115 76.5
2	• 1x library prep reaction • 1x MinION flow cell • 163 amplicons	115 800
3	• 1x library prep reaction • Flushed MinION flow cell • 88 amplicons	115 -

~ 5–6 USD per mitogenome

THE IDEAL MITOBARCODING TIMELINE

MON
- Tissue digestion

TUE
- DNA extraction
- 16S PCR
- LRPCR overnight

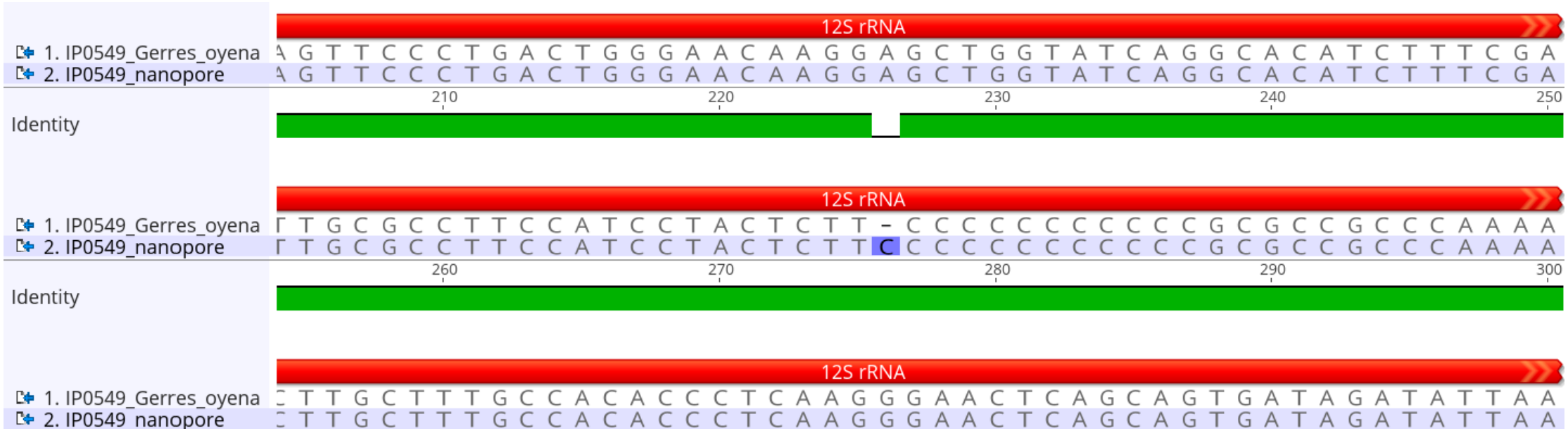
WED
- Gel-check
- PCR clean-up
- Sample pooling

THU
- Nanopore library prep
- Start sequencing for 24 h

FRI
- Base-calling
- Generate consensus
- Assemble the mito-genome

Complete mitogenomes within 1–2 weeks

NANOPORE MITOGENOMES ARE COMPARABLE TO ILLUMINA



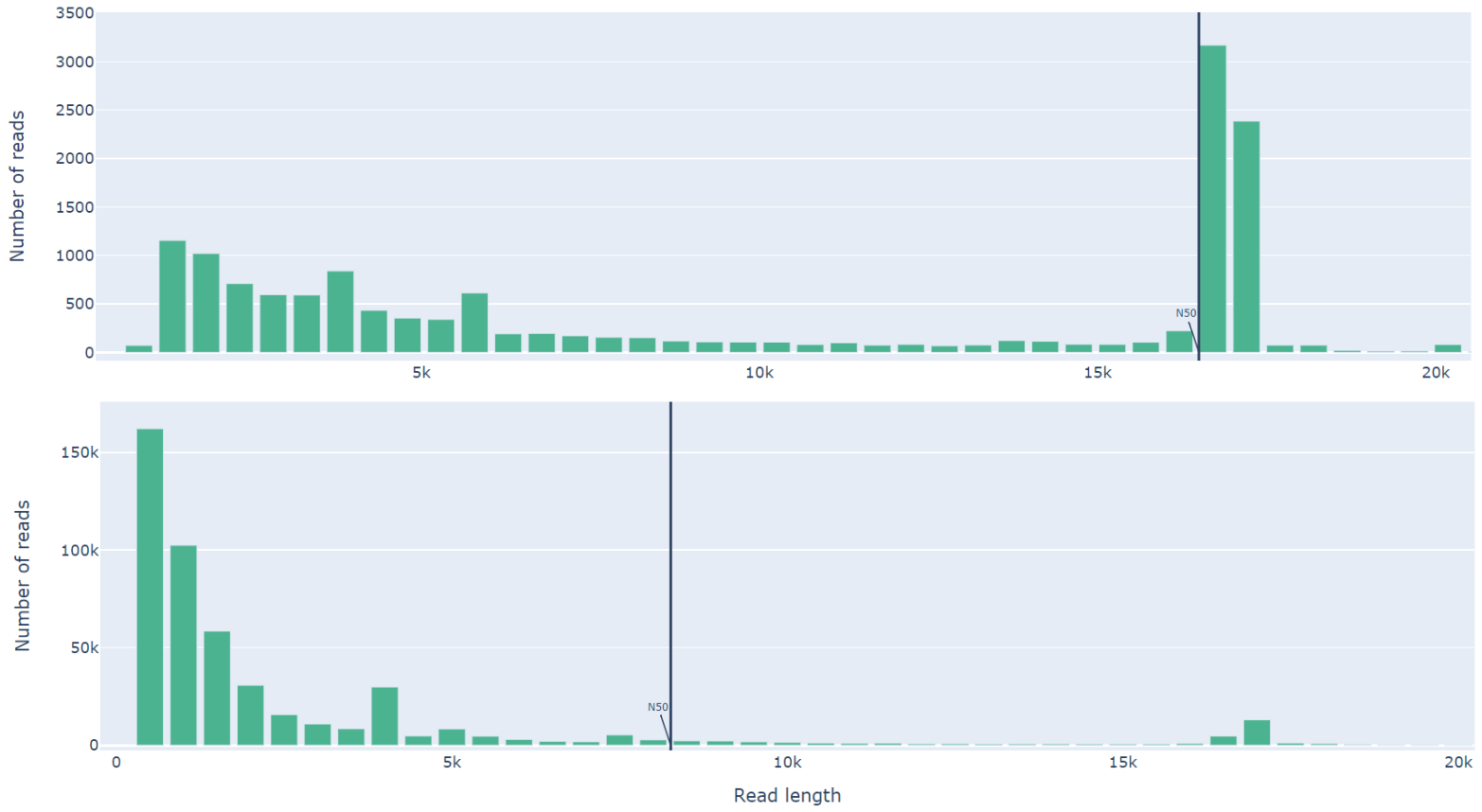
NO substitution errors observed

Most protein-coding genes had NO frameshift errors

WHAT'S THE CATCH?

- Primers do not work for all fish families

- Sensitive to gDNA quality
  - High molecular weight gDNA needed
  - Preferential amplification of short reads



SUMMARY

- We can generate mitogenomes **quickly and cost-effectively** via long-range PCR and long-read sequencing (with caveats)
- Nanopore-sequenced mitogenomes are **just as high quality** as Illumina-based mitogenome assemblies
- Better poised to future-proof reference libraries

ACKNOWLEDGEMENTS

NUS National University of Singapore
Lee Kong Chian Natural History Museum
Dalio Philanthropies
National Research Foundation
Prime Minister's Office
Singapore
The International SeaKeepers Society
Department of Biological Sciences
R·E·S·T
Reef Ecology Study Team
National University of Singapore
National Parks
SG NSCC Singapore
National Supercomputing Centre
Temasek Lifesciences Laboratory

Contact: changjjm@nus.edu.sg

THANK YOU!!