giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

german cooperation

asean

THE JOURNEY

Strengthening Laboratory Capacity on COVID-19 Bio-Genomics for ASEAN Countries: A hands-on Training and Knowledge Sharing Workshop

Foreword from the Director General of Health Malaysia

It is with immense pleasure, I welcome everyone to our Strengthening Laboratory Capacity on COVID-19 Bio-Genomics for ASEAN Countries: A hands-on Training and Knowledge Sharing Workshop.

First and foremost, we are very grateful to the German Agency for International Cooperation (GIZ) for their financial support on this project, ASEAN EOC Network, ASEAN Secretariat and GIZ (Thailand) for coming together and assisting Malaysia in organizing this workshop.

On behalf of Malaysia, I am delighted to acknowledge the participation from ASEAN Member States (AMS) with highest commitment and contributions in making this workshop a success. This is the 3rd initiative that Malaysia have coordinated in laboratory preparedness for the dangerous pathogen under the ASEAN Health Cluster 2 activities for ASEAN region, after UNITEDengue project in 2014 to 2016, which was implemented together with Singapore, and Flavivirus Surveillance project in 2019 to 2021. As for Strengthening Laboratory Capacity on COVID-19 Bio-Genomics for ASEAN Countries, this initiative outlines additional work experience on Whole Genome Sequencing (WGS).

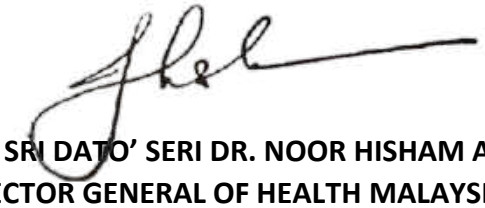
As with other viruses, SARS-CoV-2 virus constantly change their genetic material through mutation as part of evolution and adaptation process. Whilst most mutations may not have a significant impact on pathogenicity, adaptability and infectivity, some mutations may provide an advantage for the viruses to become more virulent. The COVID-19 pandemic has thought us a lesson on the importance of having capacity and capability in whole genome sequencing (WGS) especially with the frequent change of Variant of Concerns (VOCs) and Variant of Interests (VOIs).

The training on Strengthening Laboratory Capacity on COVID-19 Bio-Genomics for ASEAN Countries is introduced to assist all AMS to build up capacity building and enhancing the existing laboratory surveillance related to WGS. At the same time, it will ensure sharing of high-quality, timely and reliable sequencing data of SARS-CoV-2 that are relevant in national and ASEAN context. A series of training was conducted to meet the training needs and the objective of the project. At the end of the project, we would have a set of SARS-CoV-2 genomic surveillance data for ASEAN region for COVID-19 pandemic.

The joint action among AMS in SARS-CoV-2 virus laboratory genomic surveillance activities will provide a common, high quality, coordinated and effective response within ASEAN laboratory network to manage this pandemic and any other incident in the future. Let us all doing the best we can to fight the virus.

Thank you very much and we look forward to a productive and successful partnership with all of you and fruitful outcome of the project.

Thank you.



TAN SRI DATO' SERI DR. NOOR HISHAM ABDULLAH
DIRECTOR GENERAL OF HEALTH MALAYSIA
MINISTRY OF HEALTH MALAYSIA

Foreword from the Deputy Director General of Health (Public Health)

The Joint Statement of The ASEAN Health Ministers on ASEAN COVID-19 Response After One Year (2021), ASEAN member states recognise the importance of strengthening multisectoral cooperation and partnership to enhance laboratory capacity across the region by reinforcing genomic surveillance of SARS-CoV-2 to be able to swiftly detect and share information of potential Variant of Concerns (VOCs) and Variant of Interests (VOIs) for the implementation of public health, social and infection control measures.

The National Public Health Laboratory (NPHL) Malaysian has made great strides and effort in preparation for the Strengthening Laboratory Capacity on SARS-CoV-2 Genomics Training for ASEAN Countries from 7 April to 7 September 2022. This shows the commitment of NPHL Malaysia to ensure the that the objectives as outlined in the Joint Statement are achieved in order to control the current pandemic and to prepare for future public health emergencies.

I would like to extend my heartiest congratulations to the NPHL Director, consultant trainers, secretariats from National Public Health Laboratory team, ASEAN Secretariat, GIZ Thailand, all participants and all the contributors of this training project for the commitment and collective effort in successfully organizing this event.

On that note, I hope to see that the COVID-19 Bio-Genomics Training for ASEAN Countries being fully utilized to accelerate the integration of genome sequencing into the practices of the ASEAN health community to advance the region's preparedness and response to future public health threats.

Thank you.

DATUK DR. NORHAYATI BINTI RUSLI
DEPUTY DIRECTOR GENERAL OF HEALTH (PUBLIC HEALTH)
MINISTRY OF HEALTH MALAYSIA

Preface from the Director of National Public Health Laboratory, Malaysia

The year 2020 was a turning point in history, particularly in global health situation. The COVID-19 pandemic has disrupted our way of life with unprecedented consequences on our daily lives. Although COVID-19 has resulted in unforeseen tragedy, it has provided a unique opportunity for us to learn and enhance modern technology especially in laboratory testing.

Modern technology, namely whole genome sequencing has allowed the world to rapidly identify SARS-CoV-2 virus that has not been seen in viral outbreaks in the past. Continued genome sequencing can be further expanded and supports the monitoring of the disease's spread, pathogenicity and evolution of the virus. As more countries move towards implementing sequencing programmes, there will be more opportunities to understand better all aspects of this novel virus.

This timely collaboration and training between the Ministry of Health Malaysia, GIZ and the ASEAN Secretariat could be the breakthrough strategy that we need. The training goal is to strengthen ASEAN Member States' laboratory capability and capacity and share genome data in order to better understand the outbreak, monitor transmission, and prepare for future pathogen threats. Furthermore, the training programme is intended to address organisational capacity development needs as well as to share experiences with ASEAN Member Countries. I am hoping that this training will help to pave the way towards achieving the said goals.

Here, I would like to extend my gratitude and congratulate the NPHL secretariats for taking the lead and well-coordinate this workshop. I would also like to extend my appreciation to the collaborating partners, namely GIZ and the ASEAN Secretariat for providing the invaluable support and assistance to ensure the success of this workshop. Thank you once again for all the dedication and commitment in making this project a success.

Thank You.

DR. FATANAH BINTI ISMAIL
DIRECTOR NATIONAL PUBLIC HEALTH LABORATORY
MINISTRY OF HEALTH MALAYSIA

TABLE OF CONTENTS

CONSULTANT

Dr. Tan Joon Liang completed his Bachelor Degree in Bioinformatics in the year 2012 from Multimedia University Malaysia. Upon the completion of his thesis on analyzing Next Generation Sequencing (NGS) data with the focus on phylogenomics and evolution, he received his Doctorate Degree (PhD) from University of Malaya in 2015. In the course of three years, Dr. Joon Liang Tan has published 17 research papers, of which 12 are ISI-indexed (Q1-Q2) journals. He is the co-winner for the international award, namely Science and Technology Award 2019 by the Malaysian Toray Science Foundation. Dr. Joon Liang Tan is currently serving as Senior Lecturer for Bioinformatics course offered in Multimedia University Malaysia (Melaka Campus). He is still actively pursuing in NGS-based (genomics, transcriptomics, exomics, metagenomics) bioinformatics research, focusing on data mining, phylogenomics, micro- and macro- evolution and populational studies.

Muhd Khairul Luqman Muhd Sakaff is currently serving as FAS (Field Application Specialist); trained and certified on illumina's library preparation and illumina's sequencing technology. He obtained his Bachelor of Science (Honours) degree in Biomedicine in 2008 from Universiti Sains Malaysia and holds a Master of Science in Microbiology (2011) also from Universiti Sains Malaysia. He has more than 10 years of experiences related to molecular biology. Over this time, he was involved in several project related to genome sequencing of thermophilic bacteria, psychrophilic yeast, and rubber tree. He is always at the forefront in providing technical trainings and supports to local national health laboratories and testing labs since the start of SARS-CoV-2 crisis. He is also actively involved in supporting customers on using NGS (Next Generation Sequencing (NGS) for microbial and infectious diseases applications such as SARS-CoV-2 whole genome sequencing (WGS).

MALAYSIA PROJECT TEAM

Dr. Fatanah Binti Ismail
(fatanah@moh.gov.my)

Yu Kie Chem
(yukie@moh.gov.my)

Dr. Maria Suleiman
(mariasuleiman@moh.gov.my)

Selvanesan Sengol
(selvanesan@moh.gov.my)

Dr. Nor Zahrin Binti Hasran
(nzahrin@moh.gov.my)

Hannah Phoon Yik Phing
(hannah_phoon@moh.gov.my)

Ushananthiny Ravi
(ushananthiny_r@moh.gov.my)

MALAYSIA PROJECT TEAM

Wan Noraini Binti Wan Yussof
(wnoraini66@moh.gov.my)

Zirwatul Adilah Binti Aziz
(adilahwkt57@gmail.com)

Rehan Shuhada Binti Abu Bakar
(rehanshuhada@moh.gov.my)

Nurul Aina Murni Binti Che Yazid
(ianscuy@yahoo.com)

W Nur Afiza Binti W Mohd Ariffin
(nur.afiza@moh.gov.my)

Kamal Hisham Bin Kamarulzaman
(kamal_hisham@moh.gov.my)

Noriah Binti Mohd Yusof
(noriah.my@moh.gov.my)

MALAYSIA PROJECT TEAM

Parvathy Sundara Rajoo
(ssparvathy@moh.gov.my)

Kiroshika Pilai Veel Pilay
(kiroshika_veelpilay@live.com)

Amrish Shah Bin Osman
(amrishshah78@gmail.com)

Salfaryna Binti Alzahari
(salfaryna@moh.gov.my)

Siti Fairuz Binti Mar Ramli
(it.mkak@moh.gov.my)

Nuurun Izaati Binti Zainuddin
(it.mkak@moh.gov.my)

Lia Natasya Munirah Binti Dzolkifli
(it.mkak@moh.gov.my)

Nor Aina Sofia Binti Mohtar
(it.mkak@moh.gov.my)

ASEAN SECRETARIAT

Dr Ferdinal Fernando
(ferdinal.fernando@asean.org)

Micheal Glen
(michael.glen@asean.org)

Randy Andrian
(randy.adrian@asean.org)

GIZ THAILAND

Pouchamarn Wongsanga
(pouchamarn.wongsanga@giz.de)

Oraphan Tatha
(oraphan.tatha@giz.de)

Pakin Somkieatprayool
(pakin.somkieatprayool@giz.de)

1.0 COVID-19 GENOMICS TRAINING: A SYNOPIS

Novel coronavirus, which was named as coronavirus disease 2019 (COVID-19) by the WHO on February 11, 2020, has rapidly increased in epidemic proportions since it first appeared in Wuhan, China, in December 2019. Ensuring timely diagnosis of the COVID-19 infection is the most important step to interrupt the chain of transmission. Amidst the fact that this infection can be highly infectious, there is an indispensable need to strengthen the laboratory capacity and network preparedness plan.

The advanced deployment of laboratory capacity for molecular diagnosis of SARS-CoV-2 made early recognition of COVID-19 cases possible and contributed to the containment of the infection in most countries. However, the battle against COVID-19 is far from over and in response to pandemic speed is indeed critical. We have a steep learning curve as COVID-19 is rapidly evolving against limited established surveillance infrastructure and this is compounded with limited experience in understanding the disease and rapidly changing epidemiology

The on-going COVID-19 pandemic is a great opportunity for the utilization of Whole Genome Sequencing technology for SARS-CoV-2 sequencing for improvement of surveillance testing, development of vaccines, diagnostics, therapeutics and investigation of disease and other tools for response. Laboratory personnel must acquire the skills to perform whole genome sequencing of the SARS-CoV-2 virus and their phylogenetic network analysis of the viral genomes and able to link this information to epidemiological data for public health decision makings. However, before starting a sequencing program, it is important to have a clear understanding of the objectives of sequencing, a strategy for analysis, and a plan for how findings will be used to inform public health responses.

This project presents a hands-on training platform for knowledge sharing and enhanced communication among ASEAN Countries to understand the circulating SARS-CoV-2 variants in the region. Early sharing of SARS-CoV-2 genome sequences is important as it allows comparison of locally acquired virus sequences with the global virus genomic diversity, molecular diagnostic assays to be developed rapidly for regional public health preparedness and intervention. As more countries are moving to implement whole genome sequencing, ASEAN Countries may need to build a strong collaboration and resilient regional sequencing network which can maximize the impact of sequencing for public health intervention, not only for SARS-CoV-2 but also for future emerging pathogens and existing pathogens of public health importance.

The general objective of training is to strengthen the laboratory capability and capacity of the ASEAN member countries to generate SARS-CoV-2 genomic data in real time to understanding outbreak, monitoring transmission and possible future pathogen threats.

While the specific objectives are:

1. To perform whole genome sequencing of the SARS-CoV-2 virus isolated from infected patients.
2. To Analyze complete genomes of SARS-CoV-2isolates.
3. To perform phylogenetic network analysis of the viral genomes to better understand the COVID-19 epidemiology at each member states and its relationship with the world-wide pandemic.
4. Startup COVID-19 surveillance system for monitoring strains/mutations in ASEAN Countries.

At the end of the workshop, participants are expected to be able to perform the following activities in their own laboratories:

1) Genomic RNA (gRNA) extraction from viral isolates.
2) Library preparation for SARS-CoV-2 genome sequencing.
3) Whole genome sequencing of SARS-CoV-2 using the ARTIC multiplex PCR method.
4) Application of bioinformatics approaches to generate and analyse SARS-CoV-2 genomes.
5) Genomic surveillance and mutations monitoring.

2.0 TRAINING PROGRAMME OVERVIEW

WEEK 1: 7 APRIL 2022 (THURSDAY) OFFICIAL PROJECT OPENING		
Malaysian Time	**Agenda**	**Delivered by**
14:00 – 14:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
14:15 – 14:45	Official opening ceremony and Training launch	• Ministry of Health Malaysia • ASEAN Secretariat • GIZ Thailand
14:45 – 14:50	Photo Session (via Zoom)	NPHL Malaysia IT Team
14:50 – 16:30	***BIOSAFETY & BIOSECURITY APPROACHES TO SARS-CoV-2 WGS: From Detection to Best Practices and Risk Assessments*** 1. Principles of Biosafety and Biosecurity 2. Risk Assessment 3. Mitigation measures	 W Nur Afiza W Arifin Parvathy Sundara Rajoo Amrish Shah Osman
16:30 – 17:30	SPECIAL PLENARY SESSION	Invited expert Dr Joshua Quick *UKRI Future Leaders Fellow, University of Birmingham.* Emcee: Hannah Phoon Yik Phing
17:30 – 17:40	Q&A Session	Emcee: Hannah Phoon Yik Phing
17:40	END	

TRAINING PROGRAMME OVERVIEW

WEEK 1: 8 APRIL 2022 (FRIDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 11:15	**Session 1: COVID detection by PCR** a. RNA extraction b. Master Mix c. Results verification d. Sample storage	Rehan Shuhada Abu Bakar & Noriah Mohd Yusof
11:15 – 12:15	**Session 2: Viral isolation** a. Tissue culture - Cell line maintenance b. Virus isolation	Selvanesan Sengol
12:15 – 13:15	**Session 3: Sampling strategy** a. Reconfirmation by PCR b. Sample selection	Yu Kie Chem
13:15 – 13:30	Q&A Session	Emcee: Nurul Aina Murni Che Azid
13:30	END	

WEEK 4: 10 MAY 2022 (TUESDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 4: DNA amplification and quantification (20 Isolates)** **Lectures:** 1. Introduction to NGS (Next Generation Sequencing) 2. Introduction to MiSeq 3. Overview of Library Prep 4. cDNA conversion 5. QIAXcel DNA quantification 6. ARTIC PCR 7. Purification 8. QIAXcel & QUBIT DNA quantification	Muhd Khairul Luqman
13:00 – 13:15	Q&A Session	Emcee: Noriah Mohd Yusof
13:15	END	

TRAINING PROGRAMME OVERVIEW

WEEK 5: 17 MAY 2022 (TUESDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 5: Library preparation** 1. Create sample sheet (LRM) 2. Tagmentation 3. Post Tagmentation Clean up 4. Amplification 5. PCR Clean up 6. Library QC 7. Normalization and pooling 8. Pool library quantification and QC	Muhd Khairul Luqman
13:00 – 13:15	Q&A Session	Emcee: Rehan Shuhada Abu Bakar
13:15	END	

WEEK 7: 24 MAY 2022 (TUESDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	Agenda	Delivered by
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	Session 6: Insert pool libraries into MiSeq system 1. Pool library denaturation and dilution. 2. Begin sequencing and run monitoring. 3. Raw data collection and QC of FASTQ	Muhd Khairul Luqman
13:00 – 13:15	Q&A Session	Emcee: Kamal Hisham Kamarul Zaman

WEEK 9: 13 JUNE 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	Agenda	Delivered by
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	Session 7: Genome analysis Lectures and Tutorials	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Amrish Shah Osman
13:15	END	

TRAINING PROGRAMME OVERVIEW

WEEK 10: 20 JUNE 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 8: Genome analysis** **Lectures and Tutorials**	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Parvathy Sundara Rajoo
13:15	END	

WEEK 11: 27 JUNE 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 9: Genome analysis** **Lectures and Tutorials**	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Kiroshika Pillai
13:15	END	

WEEK 12: 4 JULY 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 10: Genome analysis** **Lectures and Tutorials**	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Zirwatul Adilah Aziz
13:15	END	

WEEK 13: 18 JULY 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 11: Genome analysis** **Lectures and Tutorials**	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Yu Kie Chem
13:15	END	

TRAINING PROGRAMME OVERVIEW

WEEK 14: 25 JULY 2022 (MONDAY) LABORATORY LECTURES AND DEMONSTRATIONS		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 13:00	**Session 12: Genome analysis** **Lectures and Tutorials**	Dr Tan Joon Liang
13:00 – 13:15	Q&A Session	Emcee: Selvanesan Sengol
13:15	END	

19 SEPTEMBER 2022 (MONDAY) OFFICIAL CLOSING CEREMONY		
Malaysian Time	**Agenda**	**Delivered by**
10:00 – 10:15	Login to Zoom platform and Registration	NPHL Malaysia IT Team
10:15 – 10:30	Welcome remarks	Director, NPHL Malaysia
10:30 – 10:50	Remarks by GIZ, Thailand	Project Director, GIZ Thailand
10:50 – 11.10	Montage Presentation: Activities in the Project	Secretariat, NPHL Malaysia
11.10 – 11.30	Closing remarks by Lead Country	Deputy Director General of Health (Public Health), MOH Malaysia
11.30 – 11.35	Photo session (via Zoom)	Secretariat, NPHL Malaysia
11.35 – 12.00	Certificate Awards Ceremony	Director, NPHL Malaysia
12.00 – 12.15	Key announcements	Moderator
12.15	**End Of Closing Ceremony**	–

3.0 PARTICIPANTS

3.1 MALAYSIA

"Before this workshop, I had a hard time implementing SARS-CoV-2 workflow in my laboratory. Since then, my outputs have never been better and the data obtained are much higher in quality."

KAMAL HISHAM BIN KAMARUL ZAMAN (NPHL)
Science Officer
(kamal_hisham@moh.gov.my)

REHAN SHUHADA BINTI ABU BAKAR (NPHL)
Science Officer
(rehanshuhada@moh.gov.my)

NORIAH BINTI MOHD YUSOF (NPHL)
Science Officer
(noriah.my@moh.gov.my)

LIM CHEAU JU (PBHLJB)
Science Officer
(viromkajb@gmail.com)

DR. MUHD HASHIM CHEW (PHLI)
Science Officer
(hep-2@hotmail.com)

SYAHIDA BINTI OMAR (PHLKB)
Science Officer
(syahidaomar@moh.gov.my)

MAZNIN NISAH BINTI WAJIL (PHLKK)
Science Officer
(mazninnisahwajil@gmail.com)

3.2 BRUNEI DARUSSALAM

MOHAMMAD AIZZUDDIN HJ MIRASIN

Science Officer
(izzati.azhan@moh.gov.bn)

"Excellent and very informative workshop"

DR NOR AZIAN HJ HAFNEH

Chief Scientific Officer
(norazian.hafneh@moh.gov.bn)

DR MUHD HAZIQ FIKRY HJ ABDUL MOMIN

Science Officer
(haziq.momin@moh.gov.bn)

NUR AMIRAH IBARAHIM

Science Officer
(nuramirah.ibarahim@moh.gov.bn)

IZZATI AZHAR

Science Officer
(izzati.azhar@moh.gov.bn)

3.3 CAMBODIA

CHIN SAVUTH

Head of Virology Unit
(savuth_chin@niph.org.kh)

"Thank you for giving the opportunity to join the workshop"

CHHE VISAL

Science Officer
(visal21ais@gmail.com)

MANH SENGLY

Science Officer
(sengly8893@gmail.com)

MRS. PRUM SITHA

Science Officer
(sithaprum87@gmail.com)

MRS. SOKHA DUL

Science Officer
(sithaprum87@gmail.com)

MRS. VY PHALLY

Science Officer
(phally068@niph.org.kh)

3.4 INDONESIA

HARTANTI DIAN IKAWATI
Researcher
(hartantidian83@gmail.com)

"Fruitfulldiscussion with all participants"

HANA APSARI PAWESTRI
Researcher
(hpawestri@gmail.com)

ARIE ARDIANSYAH NUGRAHA
Lab Technician
(rieardian@yahoo.com)

KARTIKA DEWI PUSPA
Researcher
(kartika.apt@gmail.com)

3.5 VIETNAM

National Institute of Hygiene and Epidemiology

UNG THI HONG TRANG
(Utht@nihe.org.vn)

HOANG VU MAI PHUONG
(hvmp@nihe.org.vn)

VU HIEU
(Hieu.vu.1196@gmail.com)

TRAN THI HIEN
(hnhientran@gmail.com)

NGUYEN VU SON
(Nvs@nihe.org.vn)

Pasteur Institute, Ho Chi Minh

CAO MINH THANG
(tminhcao@gmail.com)

DAO HUY MANH
(daohuymanh@gmail.com)

PHAM THU HANG
(Ptthang12@gmail.com)

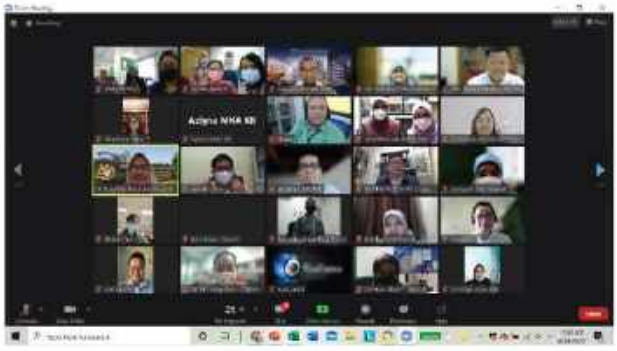
NGUYEN TRUNG HIEU
(nguyentrunghieu204@gmail.com)

NGUYEN THU NGOC
(ngocpas@gmail.com)

4.0 OPENING CEREMONY

"Through the Strengthening Laboratory Capacity on COVID – 19 Bio Genomic for ASEAN Countries: A Hands-On Training and Knowledge Sharing Workshop for the period of 20 weeks (April – September 2022), we hope to empower all the ASEAN Member States with the knowledge and skills of WGS, to enchance communication and to enable efficient information sharing in understanding the circulating SARS-CoV-2 variants in the region."

Dr. Hani Mat Hussin, the Deputy Director of Disease Control Division, Moh.



Plenary Session
The ARTIC Sequencing Protocol for SARS-CoV-2
Dr. Joshua Quick

Dr Joshua Quick has vast experience and knowledge as the UKRI Future Leaders Fellow at the University of Birmingham developing novel methods for rapid antimicrobial resistance (AMR) prediction.

He serves as a molecular biologist specialising in next-generation sequencing and has worked on both bacterial genomics and viral surveillance for outbreaks. During the West African Ebola virus epidemic, he travelled to Guinea in West Africa and established the first mobile laboratory using nanopore sequencing to perform viral surveillance.

The widely-used amplicon-based sequencing method for low-titre clinical samples which has been widely adopted against zika, yellow fever, dengue, chikungunya and ebola viruses was a brainchild of Dr Quick and his team of scientists. Currently, his focus is on bacterial single-cell technology as an approach to tackle the growing threat of resistance by rapidly identifying resistance genes in complex sample types such as clinical samples.

"I think it's an important point, because the pandemic comes in waves of new variant, emerging creating a large number of cases and displaces the previous variant and then subsided, very different demand for sequencing at different times from almost huge demands where it completely saturating every capacity that we have down to almost nothing. You need completely different approach when designing such protocol. You want to design protocol which got capacity of 50,000 sample per week, you need extensive investment of automation, cloud computing infrastructure but then all of that samples disappear to a few hundred you might consider to design a protocol that works well for a lot smaller sample size, manual processing, fast turnaround, it's a different game but then you're going have to pay more per cost per genome." Dr Joshua Quick

Take away message:

You want to design a protocol which has got a capacity of 50,000 samples per week, you will need extensive investment of automation, cloud computing infrastructure but then all of that samples may disappear to a few hundred and then you might consider designing a protocol that works well for a lot smaller sample size, manual processing, and fast turnaround.

5.0 VIRTUAL SESSIONS

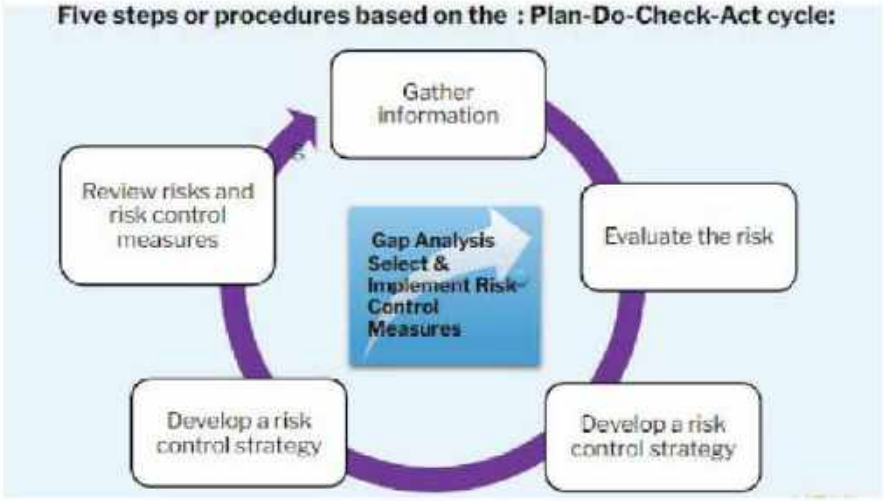
The fourteen weeks of interactive virtual sessions for sharing of knowledge and technical expertise provided the organizer and participants with a new online learning experience. The virtual sessions made the opportunity for learning possible in the COVID-restriction era, regardless of geographical location.

5.1 LECTURE SERIES

a) BIOSAFETY & BIOSECURITY SESSIONS

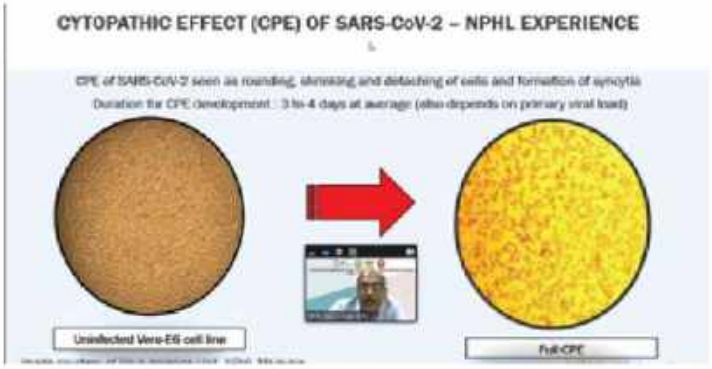
Assessment of the risks allows safety measures to be balanced with the actual risk of working with biological agents.

2.1 RISK ASSESSMENT PROCESS



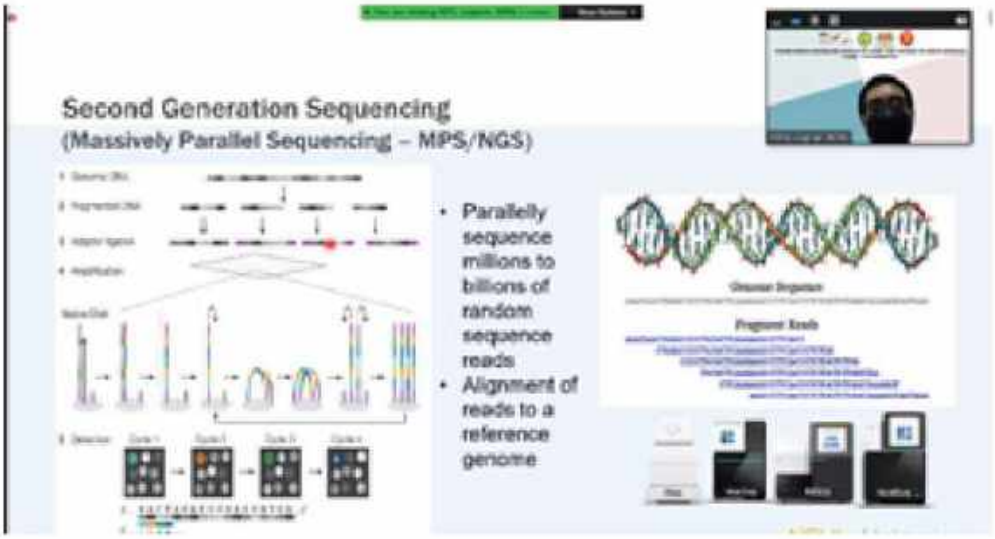
Various safety procedures must be followed to ensure that laboratory personnel, community and environment are all protected.

b) COVID-19 Detection



c) Whole-Genome Sequencing (WGS)
Technical Session :
Lecture and video demonstration

WGS is a comprehensive method for analyzing entire genomes.
Each step is crucial to achieve a good sequence.



d) Bioinformatic Analysis Sessions

The application of tools of computation and analysis to interpretation of sequence

Expectation from participants

KHM_Chhe Visal NIPH: I expect that I will deeply understanding about bioinformatics.

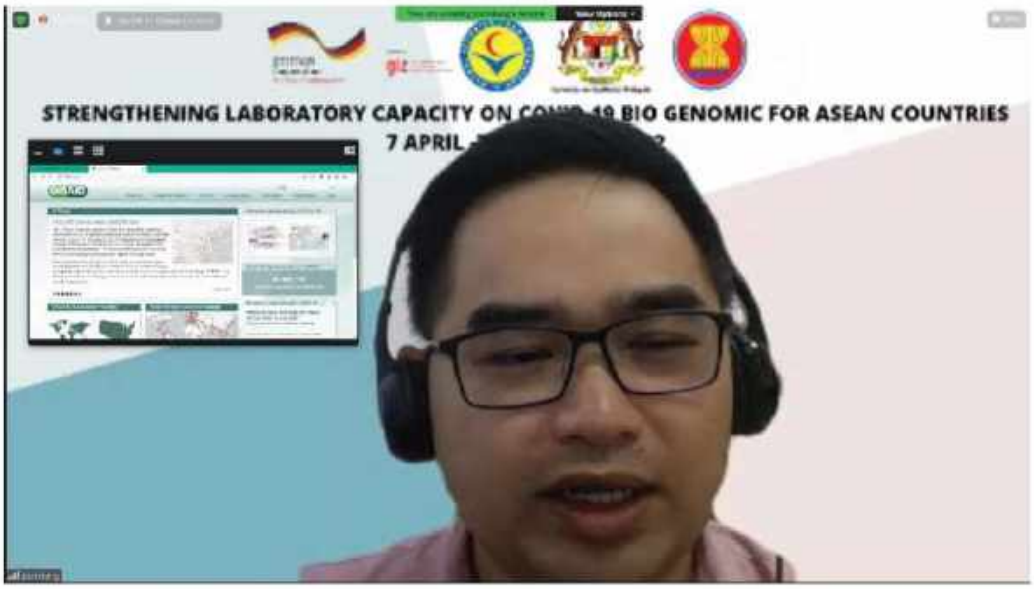
BRN, Nor Azian Hafneh DLS: Able to exercise using the recommended bioinformatic tools for quality result

IDN_Hartanti Dian Ikawati NRLI: Mining database to run bioinformatics analysis

KHM-Savuth NIPH: Able to perform genome sequencing analysis

MYS, W Nur Afiza NPHL: Able to read and interprete and analyse the sequence

10:21:25 From KHM_Sokha Dul_NIPH to Everyone: able to understand the Genomics Analyses.



5.2 ATTENDANCE PERFORMANCE EVALUATION

(Evaluation analyzed based on attendance record)

Malaysia

Cambodia

Indonesia

Brunei

Thailand

Vietnam

5.3 TROUBLESHOOTING AND CONSULTATION PLATFORM

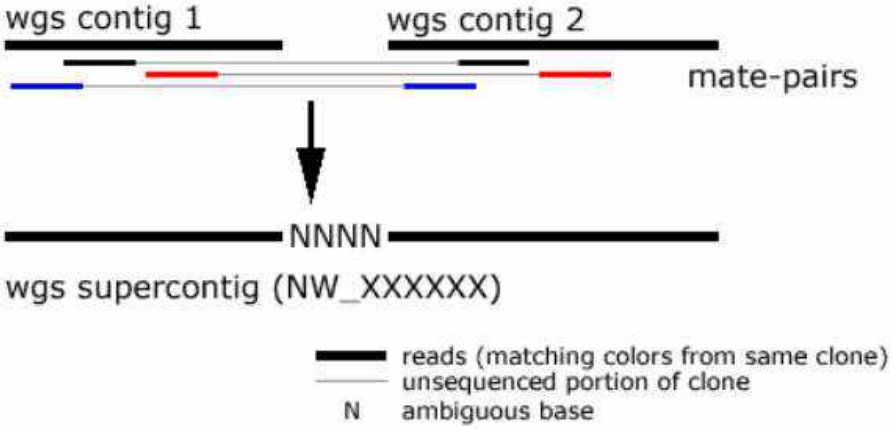
This document is a compilation of the questions and answers exchanged, both verbally and in written format during virtual session series.

a) WHOLE GENOME SEQUENCING (WGS)

MUHD KHAIRUL LUQMAN

i) DNA amplification and quantification (Questions & Answers on 10th May 2022)

1. CAM, Savuth-NIPH:
   We see a lot of 'N' in sequences, what happened?
   Answer:
   - **Sometimes, user may see 'N' assigned to represent a nucleotide position in the assembled/consensus sequence instead of usual A, T, C or G.**
   - **The 'N's is used to:**
     - **represent to sequence ambiguity of the nucleotide position in situation when user/software is not able to a precise nucleotide base call of a nucleotide position.**
     - **represent captured gaps, where the nucleotide sequence within the gap is not known but gap is covered by sequencing reads. User may sometimes see this during genome assembly as below image.**

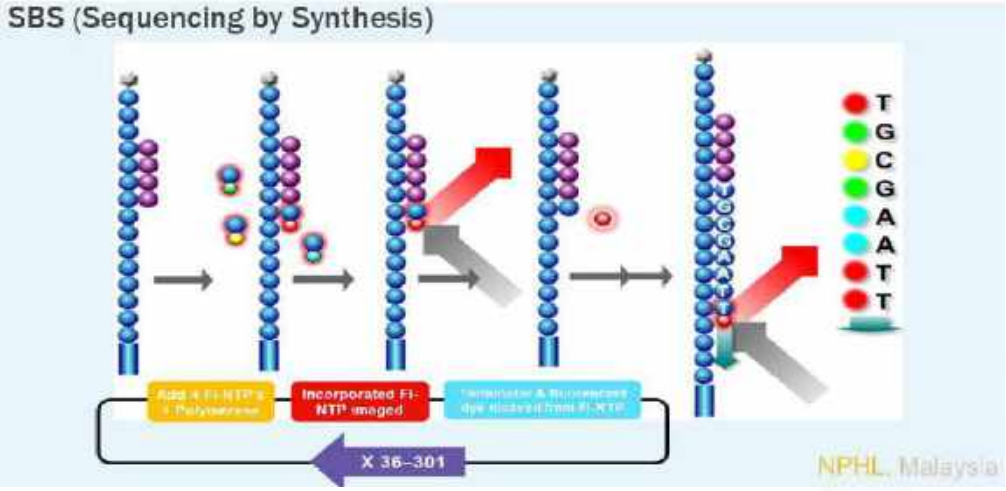


- **Reference:**
  - https://www.ncbi.nlm.nih.gov/assembly/basics/ (19th August 2022)

2. CAM,Prum Sitha_NIPH:
   What is the difference between SBS (Sequencing by Synthesis) and Sequencing by Enrichment?

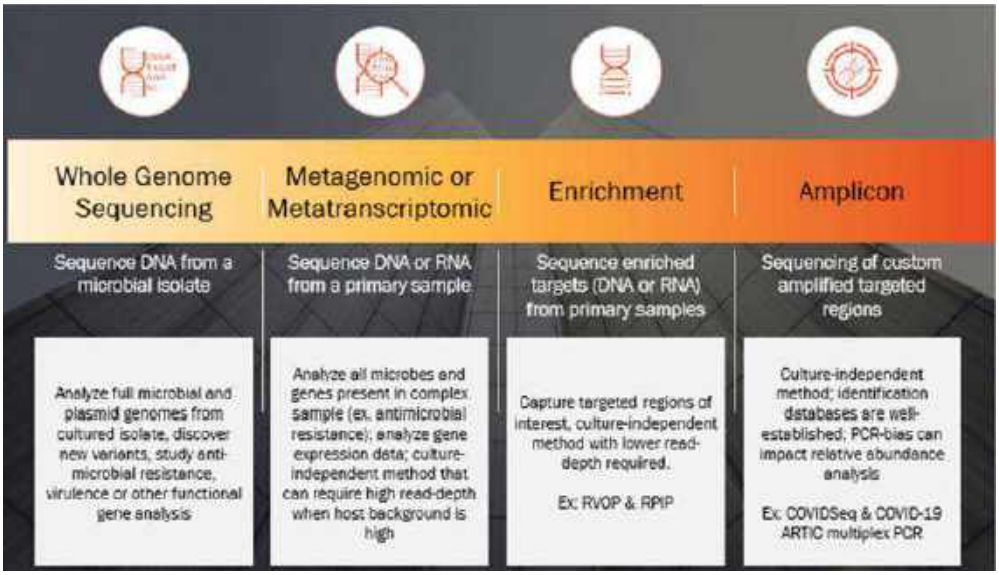
   Answer:

   - **SBS refers to the illumina's Sequencing by Synthesis chemistry, which is a type of sequencing chemistry used by illumina sequencer that allow the sequencer to perform sequencing of pooled libraries as below image.**



   - **On the other hand, sequencing by enrichment refers to a type of library preparation method that included enrichment step as part of the library preparation workflow. This enrichment step allows the user to fish/capture targeted/defined regions of interest library preparation and allow user to sequence and analyze these captured regions. User may refer to the image below on the library preparation types for further information. User may also**

contact illumina channel partner or illumina to get further info in this type of library prep.

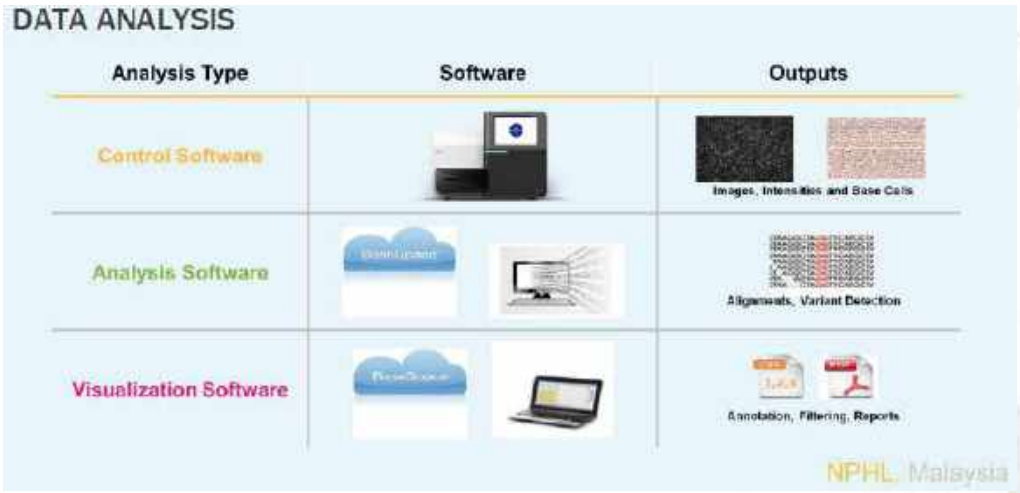


- **Reference:**
  - **LIBRARY PREPARATION & SEQUENCING 4.1 Introduction to NGS (Next Generation Sequencing) pdf.**

3. THA Nuttida Thailand NIC:
   Can you recommend another analysis software and visualization software?

   Answer:

   - **Ilumina offers numerous applications for user to do analysis and visualization on the illumina BaseSpace Sequence Hub web site (https://basespace.illumina.com) as seen below. Since there is a lot of application on BaseSpace, I highly recommend all participant to have a look to see if any of the BaseSpace's applications is relevant to participant use case.**



- **Aside from illumina's BaseSpace, there are other alternative commercial and open source bioinformatic applications that may fit the user use case.**
- **Sometime user may need to look in specialized bioinformatic pipeline set up since some project may require customized bioinformatic analysis pipeline.**
- **What applications suitable for use will depend on each user use case as each project/lab may have different processing and analysis requirement.**
- **Moreover, Dr. Tan who will be the instructor for the upcoming bioinformatic section of this training programme will be guiding all participant on example of applications suitable for sequencing data processing and bioinformatic analysis.**
- **Reference:**
    - **https://basespace.illumina.com (19th August 2022)**
    - **LIBRARY PREPARATION & SEQUENCING 4.1 Introduction to NGS (Next Generation Sequencing) pdf.**

4. IDN_Hartanti Dian Ikawati_NRLI:

   What will happened if we put two different unique indexes in the same sample?

   Answer:

   - **In this scenario, the samples will be represented by 2 libraries that have its own unique index combination.**
   - **After sequencing is completed, the data/reads will be de-multiplexed back to each individual sample based on the assigned index. If same sample have been prepared twice with different index, user will see that the data for this sample is being present in two different fastq files set.**

- **Thus, user can combine back these 2 fastq files set for analysis later to get full amount of data for the affected sample.**

5. Prum Sitha, CAM NIPH:
   What is the best cycle for SARS-CoV-2 sequencing?

   Answer:

   - **For sequencing SARS-CoV-2 genome, I have no exact recommendation. However, user may check the cycle setting tested by illumina for COVIDSeq kit (user can refer to the published illumina technical note) and see if it is suitable for user use case.**
     - **User can either use 2 x 76. The sequencing run will complete earlier but the read length produce is shorter. Suitable for user who has tight turnaround time.**
     - **Or user may use 2 x 151 cycle. The sequencing run will take longer but the read length produced will be longer. Suitable for user who need a better genome assembly result.**
   - **If user want to use other sequencing cycle setting aside from above setting, user will need to test them first to see how the sequencing result data affect the analysis.**
   - **Reference:**
     - **Illumina COVIDSeq Assay (96 samples) product page https://www.illumina.com/products/by-type/clinical-research-products/covidseq-assay.html (19th August 2022)**
     - **Sequencing guidelines for COVID-19 surveillance using the Illumina COVIDSeqTM Test (RUO Version) - https://sapac.support.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-covidseq-ruo-read-length-depth-tech-note-m-gl-00088/illumina-covidseq-ruo-read-length-depth-tech-note-m-gl-00088.pdf (19th August 2022)**

6. Savuth, CAM NIPH:
   For native samples, we do sequencing based on PCR under Ct 25 and do expect to get good sequences. Unfortunately, some samples with Ct value or 25 or lower did not get good sequence assembly outcome. Why is this happening?

   Answer:

   - **Although user may use SARS-CoV-2 samples with Ct value under 25 to maximise the chances of having good input samples for the library preparation, this PCR only check the Ct value by amplifying a small part of the viral genome.**

- **This qPCR assessment may miss out possible degradation in other part of the genome, which user can only know after sequencing and analysis is done. Thus, due to this, good Ct value do not always correlate with good sequencing outcome.**
- **Therefore, sometimes user who uses SARS-CoV-2 samples with good Ct value may still not able to get good genome assembly outcome since Ct value result does not tell user about the genome intactness of the SARS-CoV-2 samples.**

7. Hartanti Dian, Indonesia NRLI:
   Mr. Luqman, I found that the beads I use is clumping and form a slimy structure that cannot be separated with magnetic stand. Why is this happening? It only happens for one sample, while others seem fine.

   Answer:

   - **The usual suspect for the clumping seen during library preparation workflow is that some the beads may become over-dry during the workflow.**
   - **This is due to the beads drying out after being exposed for too long to the air. If the beads over-dry, the user will have harder time to resuspend these beads back into the mixture.**

8. Hang, Vietnam PIHCM:
   Which method do you use for ARTIC PCR QC step?

   Answer:

   - **For ARTIC PCR QC step, user may check the concentration (using qubit) and the PCR product fragment (using agarose gel or fragment trace analysis like Bioanalyzer 2100, LabChip etc).**

9. Arie, Indonesia NRLI:
   Should user includes positive control and negative control in every library prep and sequencing?

   Answer:

   - **The decision to include or exclude positive control and negative control need to be decided by the user since for surveillance purpose, user is already working with known positive samples.**
   - **During early round of library prep and sequencing, some labs include positive control and negative control during library prep and sequencing to validate their entire workflow.**
   - **Once the workflow is already validated, some of labs does decide to exclude positive control and negative control in future every library prep and sequencing. So that they can maximise the number of samples per each batch of library prep and sequencing run.**

- **However, some lab still includes positive and negative control for every round of library prep and sequencing due to their lab SOP/reporting requirement.**

ii) Library preparation (Questions & Answers on 17th May 2022)

1. CAM, Sitha Prum_NIPH:
   Could you please mention target pathogen for Nextera DNA Flex Prep Kit?

   Answer:

   - **For all participants information, illumina has recently renamed the 'Nextera DNA Flex Prep' kit to 'illumina DNA Prep' kit.**
   - **Illumina DNA Prep kit does not have any specific target pathogen. This kit can accept input DNA/cDNA/amplicon from many organism (ex: human, plant, fungi, bacteria, viruses etc).**
   - **In addition, the kit may also accept blood and saliva input (this may require user to get additional kit and do additional library preparation steps).**
   - **If interested in illumina DNA Prep kit, user may contact illumina channel partner or illumina for further information.**
   - **Reference:**
     - **Illumina DNA Prep product page - https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-dna-flex.html (19th August 2022)**

2. AM, Sitha Prum_NIPH:
   Is the illumina DNA Prep kit suitable for SARS-CoV-2 library preparation?

   Answer:

   - **Yes, user may use illumina DNA Prep kit as part of the workflow to prepare SARS-CoV-2 library. Example I can offer is the ARTIC COVID-19 WGS library prep method that was covered during this training programme.**
   - **Reference:**
     - **LIBRARY PREPARATION & SEQUENCING 4.3 Overview of Library Preparation pdf**
     - **Illumina DNA Prep product page - https://www.illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-dna-flex.html (19th August 2022)**

3. IDN_Hartanti Dian Ikawati_NRLI:
   Have you ever experience failure in generating the fastq file?

   Answer:

   -

- **Yes, I have seen cases where the illumina sequencer system encountered issue of not able to initiate the fastq files generation.**

4. IDN_Hartanti Dian Ikawati_NRLI:
   And what cause this fastq file generation issue?

   Answer:

   - **Sometime this issue may happen due to possible root cause like communication error. Sometimes this may happen due to other possible root cause.**
   - **Once issue is reported by user, illumina channel partner or illumina will help user to troubleshoot and try to identify the possible root cause.**
   - **One example possible root cause that may happen is software communication issue which prevent communication with MCS (MiSeq Control Software) and LRM (Local Run Manager). This prevented LRM to proceed with fastq file generation after sequencing is completed by MCS. One solution to known to resolve this issue is to power cycle the MiSeq system. Once power cycled, usually the software communication issue is resolved and LRM can proceed to requeue the fastq files generation again.**

5. IDN_Hartanti Dian Ikawati_NRLI:
   Can we resolve this issue without doing another library prep?

   Answer:

   - **As mentioned before, illumina channel partner or illumina will need to investigate and identify the probable root cause to know why the fastq files generation process face issue.**
   - **If for example the possible root cause is identified to due to communication error, power cycling the sequencer may be able to resolve this issue. If the sequencer can proceed to requeue the fastq files generation again, the user may not need to repeat the library preparation and sequencing again.**
   - **If the probable root cause lies elsewhere, then illumina channel partner or illumina will need to identify the possible root cause before determining and advise if user need to re-prep the library again.**

6. IDN_Hartanti Dian Ikawati_NRLI:
   Is it necessary to use LoBind Eppendorf tube consumable in library prep workflow?

   Answer:

   - **Usage of LoBind tube consumable minimizes the loss of the prepared library (as library/DNA may stick to the tube wall of non-LoBind tube). Thus, our denature and dilution process will be much more precise, allowing user to**

load the optimal library concentration into the sequencer to maximize the chances to get good sequencing outcome.

- **Thus, the illumina DNA Prep kit does recommend user to use LoBind tube or equivalent type of LoBind consumables to heighten the chances of user getting the best library prep outcome.**
- **Reference:**
  - **Illumina DNA Prep User Guide: https://support.illumina.com/content/dam/illuminasupport/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prepreference-guide-1000000025416-10.pdf (19th August 2022)**

7. VNM, Hang _ PIHCM:
   May I ask a question. What happen if the concentration of sample is 100-500ng and run PCR index >5 cycles?

   Answer:

   - **Some of the library fragment may have higher chances of being amplified more than other fragments. This is what is known as PCR bias.**
   - **Thus, it is preferable for user to use minimize PCR cycle as needed during Amplification of Tagmented DNA steps. This is to ensure that user get an even library representation of the genome sample for sequencing purpose.**
   - **Thus, user may refer to the number of PCR cycle recommended by illumina) based on the initial input concentration as below image.**

Total DNA Input (ng)	Number of PCR Cycles (X)
1–9	12
10–24	8
25–49	6
50–99	5
100–500	5
Blood/Saliva	5

   - **Reference:**
     - **LIBRARY PREPARATION & SEQUENCING 5.4 Amplification of Tagmented DNA pdf**
     - **Illumina DNA Prep User Guide: https://support.illumina.com/content/dam/illuminasupport/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prepreference-guide-1000000025416-10.pdf (19th August 2022)**

8. CAM, Chhe Visal_NIPH:
   After normalizing each library, should we measure the concentration of DNA again?

   Answer:

- **Usually, user will measure the DNA concentration of the following to ensure normalization and pooling is done correctly:**
  - **Stock libraries**
  - **Normalized pooled library**
  - **Pooled library ready for denature and dilute process (either 4nM or 2nM)**
- **Measuring each normalized library before pooling will depend on user preference. Some users will skip on checking on individual normalized library to save on the qubit reagent. Some users will check each normalized library before pooling these libraries since they want to minimize user handling error (to get even and precise distribution of all libraries in the pooled library).**

9. CAM, Savuth NIPH:
   Can we pool samples more than three?

   Answer:

   - **Yes, user may pool more than 3 library per pool. However please ensure that:**
     - **Each library in the pool has unique index pair ID**
     - **Users have enough unique index pair combination for all samples in the pool.**
     - **The number of libraries in the pool does not exceed the number of recommended libraries per sequencing run. This is to ensure that user will still have sufficient data per library/sample after sequencing run data demultiplexing.**
   - **Reference:**
     - **LIBRARY PREPARATION & SEQUENCING 5.4 Amplification of Tagmented DNA pdf**
     - **Illumina DNA Prep User Guide: https://support.illumina.com/content/dam/illuminasupport/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prepreference-guide-1000000025416-10.pdf (19th August 2022)**

10. MY Kamal _NPHL:
    On creating sample sheet using LRM, what if operator's MiSeq does not have LRM software?

    Answer:

    - **In this situation, the user MiSeq may be using MiSeq system that still have older version of MiSeq control software installed. In this screnario, user may create the sample sheet for the MiSeq either by:**
      - **using IEM (illumina Experiment Manager) software**

- manually create the sample sheet using sample sheet template

- **Reference:**
  - IEM product page - https://sapac.support.illumina.com/sequencing/sequencing_software/experiment_manager/downloads.html (19th August 2022)

11. MY Kamal_NPHL:
    Regarding library clean-up steps. If there are beads seen within the pipette tip when operator want to discard the supernatant, what can operator do in this situation?

    Answer:

    - Accidentally throwing away the beads while discarding the supernatant will leads to loss of sample/library.
    - In this situation, user should put the supernatant with the beads back into the tube/plate and incubate again on magnetic stand. Once the supernatant become clear, then user may try again to discard the supernatant again while keeping the beads this time.
    - Reference:
      - Illumina DNA Prep User Guide: https://support.illumina.com/content/dam/illuminasupport/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prepreference-guide-1000000025416-10.pdf (19th August 2022)

12. MY Hannah_NPHL:
    What will happen if we wrongly select the library kit in LRM?

    Answer:

    - If user accidentally prepared the sample sheet is prepared with the wrong library selection, this will may lead to:
      - wrong index setup
      - wrong sequencing cycle set up
      - wrong reads set up
      - in worst case scenario, user may not get the sequencing data they want since the wrong sequencing set up may not output the sequencing data that user wants.
    - If user have not yet start sequencing, user can still re-check the sample sheet setup and edit the sample sheet to correct for any mistake found.
    - If user have already start/finish sequencing, then user will need to check if the sequencing output data is still usable for their need.

- If the mistake is only within index assignment, then user may edit the sample sheet and requeue the fastq files generation.
- If the mistake is within other configuration (ex: sequencing cycle or index cycle), depending on the situation, in worst case, user may need to do a new sequencing run using corrected sample sheet.

- If this happens, please contact illumina channel partner or illumina to get further advice.

13. MY Hannah NPHL:
    Are there any difference in the version LRM if we use different illumina sequencer such as MiniSeq or iSeq instead of MiSeq?

    Answer:

    - There is no major difference in the LRM version between illumina sequencer type (ex: iSeq 100, MiniSeq, MiSeq etc). The LRM interface and function is standardized across all illumina sequencer that uses LRM.
    - Therefore, user can easily use the LRM in other illumina sequencer due to the standardized LRM interface.
    - Reference:
        - Local Run Manager for intuitive data analysis – https://sapac.illumina.com/products/by-type/informatics-products/local-run-manager.html (19th August 2022)

iii) Insert pool libraries into MiSeq system (Questions & Answers on 24th May 2022)

1. IDN_Hartanti Dian Ikawati_NRLI:

    We never check the NaOH pH before, did you do some optimization regarding this matter to come to this best practise?

    Answer:

    - The best practice of checking the pH diluted 0.2 N NaOH pH (ex: using pH strip or pH meter) was recommended since this allow user a way to check to ensure that the diluted 0.2 N NaOH pH is pH 13 and above.
    - This additional QC is to help user to minimize the 0.2 N NaOH dilution preparation error.
    - This information will also help investigator to narrow down possible root cause if illumina channel partner or illumina support need to investigate any sequencing issue.

2. CAM, Savuth_NIPH:

    In your experiences, what is chemical substances to treat the liquid waste?

    - .

- **If referring to the sequencer liquid waste treatment and disposal, user is recommended to dispose it as chemical waste since this contain formamide**
- **If referring to library prep liquid waste treatment and disposal, user may treat them with decontamination chemical (ex: Decon 90, bleach etc) before disposing them.**
- **However, the best way to do this is for the user consult their respective Lab Safety Officer to ensure that user is using suitable liquid waste treatment procedure. Since waste treatment and disposal procedure should follows the local law and organization regulation**
- Lab Safety Officer may download the SDS (Safety Data Sheet) for illumina product from the following illumina web site link: https://sapac.support.illumina.com/sds.html

3. Quiz from Mr Luqman:
   Q: Why do we denature and dilute library

   Answer:

   - Denaturing and diluting the pooled library for sequencing is done since the original pooled library is in concentrated dsDNA (double stranded DNA fragment) form.
   - Thus, user will need to initially denature the dsDNA into ssDNA (single stranded DNA fragment) form using diluted NaOH, as only library in ssDNA form can bind to the sequencer flow cell and form cluster.
   - Then, user will dilute the ssDNA pooled library to down to the sequencing loading concentration (ex: pM level) as recommended in illumina library prep or illumina sequencer user guide. This is done to ensure that the ssDNA library loaded for sequencing is at the optimal diluted concentration to get the best sequencing outcome. Underloading and overloading the pooled library for sequencing will affect the sequencing data quality/outcome.
   - Reference:
     - LIBRARY PREPARATION & SEQUENCING 6.1 Pool library denaturation and dilution pdf
     - LIBRARY PREPARATION & SEQUENCING 6.3 Sequencing run data collection and QC of FASTQ pdf

4. Quiz from Mr Luqman:
   Q: What do we put in MiSeq to start the sequence

   Answer:

   - User will need to load MiSeq sequencing reagent cartridge (already loaded with denatured diluted pooled library), MiSeq PR2 reagent bottle and MiSeq flow cell.
   - Reference:

- **LIBRARY PREPARATION & SEQUENCING 6.2 Begin sequencing and run monitoring pdf**

5. MY Rehan Shuhada_NPHL:

   Aside from thawing sequencing reagent in water bath, what other thawing method can operator use?

   Answer:

   - **User may thaw the MiSeq sequencing reagent overnight in 2°C to 8°C storage. Then store in 4°C. MiSeq sequencing reagent thawed this way is stable up to 1 week for sequencing run use.**
   - **Reference:**
     - **MiSeq System Guide for Windows 10: https://support.illumina.com/content/dam/illuminasupport/documents/documentation/system_documentation/miseq/miseq-systemguide-15027617-06.pdf (19th August 2022)**

6. MY Rehan Shuhada_NPHL:

   Can operator re-freeze thawed sequencing reagent if not used?

   Answer:

   - **No, please do not freeze back the thawed sequencing reagent. illumina recommend user to not re-freeze thawed sequencing reagent. Instead, user should store this thawed sequencing reagent at 4°C until user is able to start sequencing run.**
   - **In the past, illumina has observed several cases where using sequencing reagent that has been freeze-thawed may leads to suboptimal/bad sequencing outcome.**

7. MY Rehan Shuhada_NPHL:

   What if the flow cell is cracked? Can operator still use them for sequencing?

   Answer:

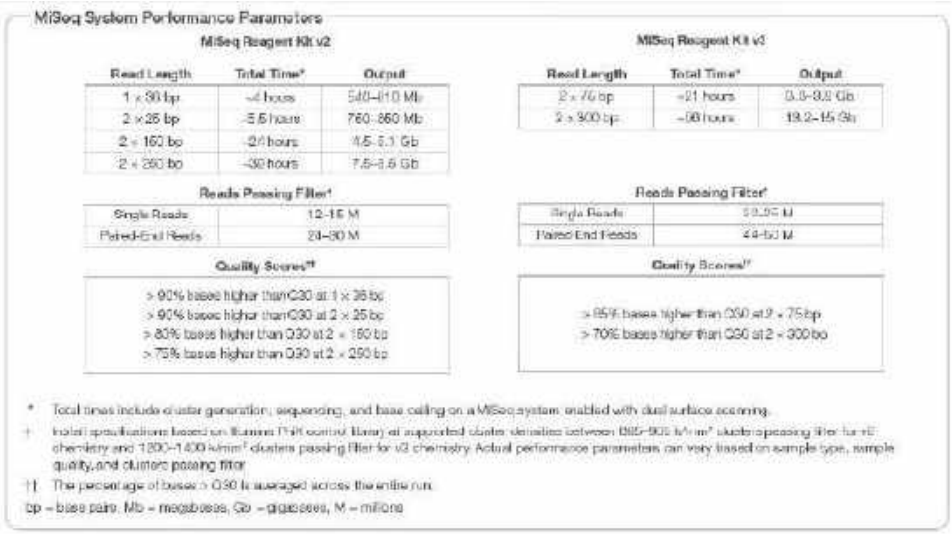
   In this scenario, user should put the cracked flow cell aside. Then contact illumina channel partner or illumina support to report the cracked flow cell issue so that qualified support staff can contact user to further investigate this issue.

8. Quiz from Mr Luqman:

   To determine sequencing run quality, should operator look at the sequencing run's data yield and Q30 scores?

Answer:

- **Yes. Sequencing run's data yield and Q30 scores will be the run metrics that user may use check if their sequencing run outcome is good or not. User should then compare their run metric to the respective illumina sequencing spec published on the illumina web site.**
- **For example, illumina MiSeq sequencing spec is as below image.**



MiSeq System Performance Parameters

MiSeq Reagent Kit v2

Read Length	Total Time*	Output
1 × 36 bp	~4 hours	540–610 Mb
2 × 25 bp	~5.5 hours	760–860 Mb
2 × 150 bp	~24 hours	4.5–5.1 Gb
2 × 250 bp	~39 hours	7.5–8.5 Gb

Reads Passing Filter†

Single Reads	12–15 M
Paired-End Reads	24–30 M

Quality Scores††

> 90% bases higher than Q30 at 1 × 36 bp
> 90% bases higher than Q30 at 2 × 25 bp
> 80% bases higher than Q30 at 2 × 150 bp
> 75% bases higher than Q30 at 2 × 250 bp

MiSeq Reagent Kit v3

Read Length	Total Time*	Output
2 × 75 bp	~21 hours	9.5–9.9 Gb
2 × 300 bp	~56 hours	18.2–15 Gb

Reads Passing Filter†

Single Reads	22–25 M
Paired-End Reads	44–50 M

Quality Scores††

> 85% bases higher than Q30 at 2 × 75 bp
> 70% bases higher than Q30 at 2 × 300 bp

* Total times include cluster generation, sequencing, and base calling on a MiSeq system enabled with dual surface scanning.

† Initial specifications based on Illumina PhiX control library at supported cluster densities between 085–900 k/mm² clusters passing filter for v2 chemistry and 1200–1400 k/mm² clusters passing filter for v3 chemistry. Actual performance parameters can vary based on sample type, sample quality, and clusters passing filter.

†† The percentage of bases > Q30 is averaged across the entire run.

bp = base pairs, Mb = megabases, Gb = gigabases, M = millions

- **If user sequencing run metric is above the illumina spec, then the overall sequencing run outcome is good, and user may proceed to collect the data and do downstream analysis.**
- **If user sequencing run metric is below the illumina spec, user should contact illumina channel partner or illumina support to report this so that qualified support staff can contact user to further investigate this issue.**
- **Reference:**
  - **LIBRARY PREPARATION & SEQUENCING 6.3 Sequencing run data collection and QC of FASTQ pdf**
  - **Example: illumina MiSeq spec: https://www.illumina.com/systems/sequencing-platforms/miseq/specifications.html (19th August 2022)**

9. MY Rehan Shuhada_NPHL:
   Q: If my sequencing run output is less than what I expected, how do I move forward?

   Answer:

   In situation where user face issue with their sequencing run output, please contact illumina channel partner or illumina support so that qualified support staff can contact user to further investigate this issue.

b) BIOINFORMATIC ANALYSIS

Dr. Tan joon liang

1. MY W Nur Afiza_NPHL:
   Difference between related genome and neighbouring genome in the audacity instant?

   Answer: Both should be the same.

2. IDN_Hartanti Dian Ikawati_NRLI:
   How about reference sequence for the variants?

   Answer: If you intend to use the variants information from any website or publication, you have to check reference genome used. Eg: Wuhan-HU-1 or WIV04.

3. MY Rehan Shuhada_NPHL:
   Dr Tan, which references should we used to construct phylogenetic tree? WIV04 (from GISAID) or NC_045512 (reference from NCBI)

   Answer: It depends on what you wish to infer from the phylogenetic tree. Eg: If you wish to study about geographical regions, then download collections from multiple countries or continents; if you wish to study about lineages, you need to download representation from different lineages.

4. MY Rehan Shuhada_NPHL:
   Do we have to have publication first before submitting sequences to NCBI? Or can submit sequences without any publication?

   Answer: It is depending on the journal's policy. Based on my experience, we have to submit sequences prior submitting journal. This is because editorial and reviewers supposed to check for the data availability.

5. MY Rehan Shuhada_NPHL:
   Dr, what should we choose to blast new nucleotide in NCBI? Highly similar (megablast) or somewhat similar sequences (blastn)? Can Dr explain more about this. Thank you.

Answer: Megablast is the most common algorithm used for web BLAST. The algorithm will return closest homolog, as compared to other two algorithms.

6. MY Hannah Phoon_NPHL:
   Dr Tan. can you explain what are the key differences between the different NCBI, GISAID, DDJB, ENA? When should we use the different sites? Are there any pros and cons of using certain sites?

   Answer: All the stated databases are established so there is no issue on the usage. Those databases are similar but not identical. Although they are sharing some identical data, but they are still having different additional dataset and tools. So it is depending on what you need to use.

c) COMPILATION OF EXERCISE

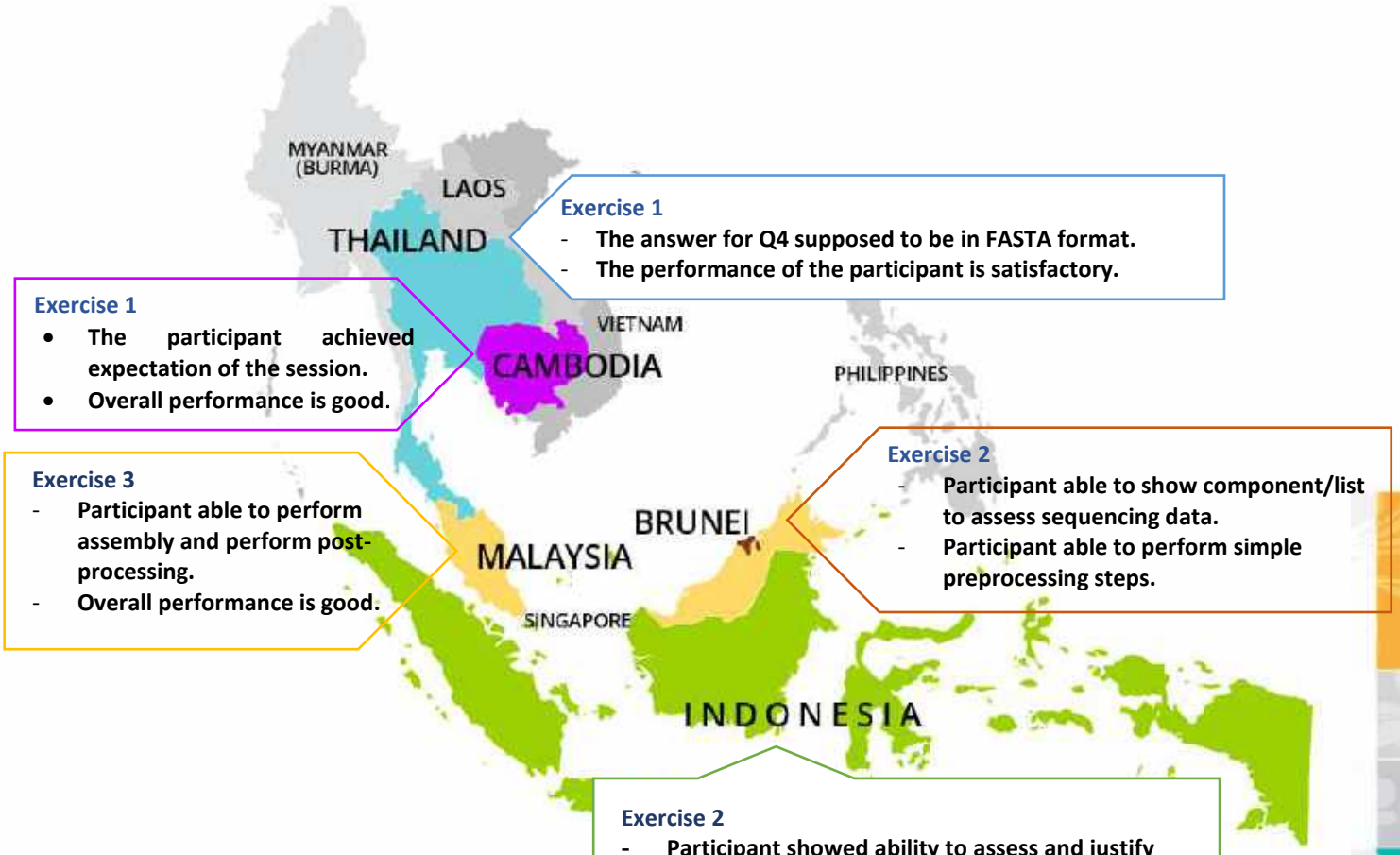
To ensure better understanding of participant during the on-line sessions, series of exercise on bioinformatic analysis such as understanding SARS-CoV-2 genome and database, pre-process sequencing steps and sequence assembly was conducted. Each participant was required to participate in this exercise. Each exercise consists of 3 to 4 questions.

These exercises were grouped into three categories: 1) SARS-CoV-2 databases & genome; 2) Pre-process sequencing steps and 3) Sequence assembly. Five (5) ASEAN countries (Brunei, Cambodia, Indonesia, Malaysia and Thailand) have contributed in these exercises, and Malaysia, Indonesia, Brunei and Cambodia were successfully completed all the exercise given.

The overall performance of the participants was good, achieved the objective and the expectation of the training sessions. However, few participants didn't answer correctly on Exercise 2 and 3. More effort are required to better understand the basic concept of bio-informatic and how it's being applied.

Refer to the Appendix 1 for compilation of the questions and answers of exercises.

Dr Tan's feedback on exercises conducted.



EXERCISE 1

COUNTRY: INDONESIA

1. Provide any TWO links for SARS-CoV-2 databases.

Answer:

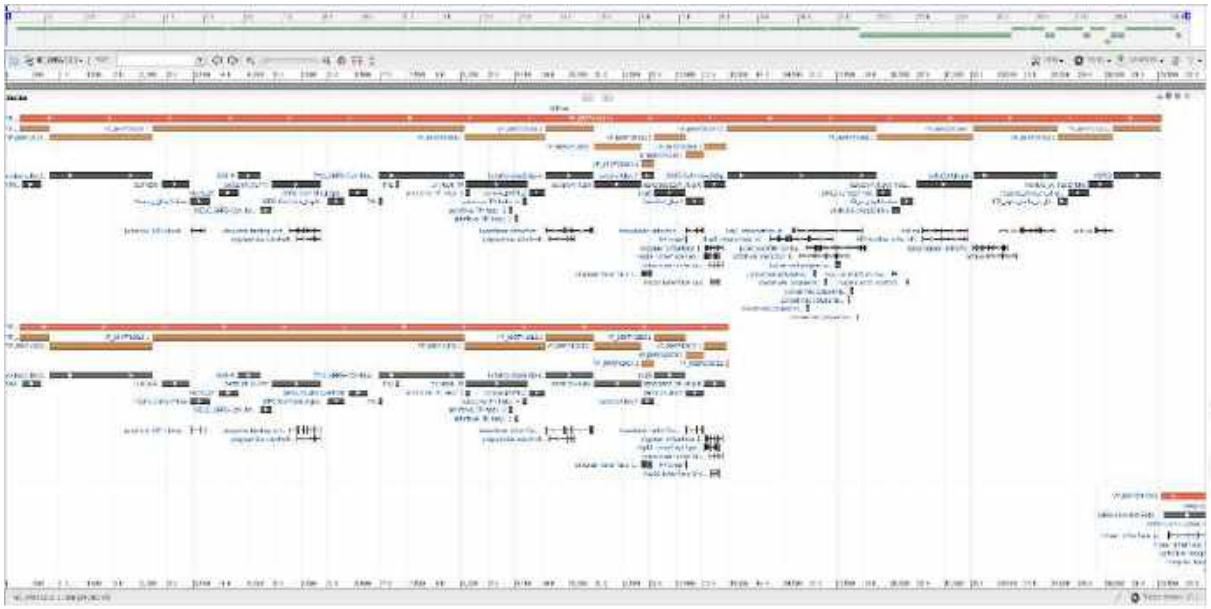
1. https://ncbi.nlm.nih.gov/sars-cov-2/
2. https://www.epicov.org/epi3/frontend#1b0800

2. List all the genes in SARS-CoV-2 genome. Provide the link of database of which being referred to.

Answer: from GISAID (https://gisaid.org/)

3. Screenshot graphical representation of SARS-CoV-2 genome.

Answer:



4. Provide the first 10 lines of genome sequence of SARS-CoV-2 in Fasta format.

Answer:

>NC_045512.2 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGT
TCTCTAAACGAACTTTAAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACTCAC
GCAGTATAATTAATAACTAATTACTGTCGTTGACAGGACACGAGTAACTCGTCTATCTTCT
GCAGGCTGCTTACGGTTTCGTCCGTGTTGCAGCCGATCATCAGCACATCTAGGTTTCGTCC
GGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTCCCTGGTTTCAACGAGAAAACACAC
GTCCAACTCAGTTTGCCTGTTTTACAGGTTCGCGACGTGCTCGTACGTGGCTTTGGAGACT
CCGTGGAGGAGGTCTTATCAGAGGCACGTCAACATCTTAAAGATGGCACTTGTGGCTTAG
TAGAAGTTGAAAAAGGCGTTTTGCCTCAACTTGAACAGCCCTATGTGTTCATCAAACGTTC
GGATGCTCGAACTGCACCTCATGGTCATGTTATGGTTGAGCTGGTAGCAGAACTCGAAG
GCATTCAGTACGGTCGTAGTGGTGAGACACTTGGTGTCCTTGTCCCTCATGTGGGCGAAA
TACCAGTGGCTTACCGCAAGGTTCTTCTTCGTAAGAACGGTAATAAAGGAGCTGGTGGCC
ATAGTTACGGCGCCGATCTAAAGTCATTTGACTTA

Dr Tan's Feedback:

- The participant achieved objectives of the session.
- The performance of the participant is good.

EXERCISE 2

COUNTRY: MALAYSIA

1. What is the difference between trimming and filtering?

Answer:

- **Trimming**: Removes the reads at one or both ends so that only a region of high-quality bases are left
- **Filtering**: Removes the whole sequencing reads that will exclude in genome assembly

2. Provide basic statistics (eg: number of reads, lengths, adapter content) of raw sequencing reads for any ONE of your sequenced sample.

Answer:

> raw sequencing retrieved from SRA toolkit.

Measures	SRR14678991_1.fastq	SRR14678991_2.fastq
Total sequences	84,538	84,538
Sequence length	251	251
Number of reads	21,219,038	21,219,038
Adapter content	TruSeq3	TruSeq3

3. What criteria would you use to incorporated for the preprocessing steps? Why?

Answer:

TruSeq3-PE.fa:30:10:2 HEADCROP:30 MINLEN:67

- Remove adapter TruSeq3
- Cut the 30 nucleotide bases on the 5' end. Has big gap of sequence content between the complimentary nucleotide bases at the front end
- Drop the read if it is below a 67

4. Provide basic statistics (eg: number of reads, lengths, adapter content) of preprocessed sequencing reads for any ONE of your sequenced sample.

Answer:

Measures	SRR14678991_1.fastq	SRR14678991_2.fastq
Total sequences	71,803	71,803
Sequence length	221	221
Number of reads	15,868,463	15,868,463
Adapter content	0	0

Dr Tan's Feedback:

- Participant showed ability to apply the concepts and execution of preprocessing
- Overall performance is good.

EXERCISE 3

COUNTRY: BRUNEI

1. State the summary of your assembly (ies). Eg: number of contigs and genome size.

   Answer:
   4 contigs
   >k119_2 len= 333
   >k119_0 len= 6200
   >k119_1 len = 866
   >k119_3 len = 22843

2. State any TWO possible contributing factors for assembly generating more than one contigs.

   Answer:
   Large genome sizes
   High repetitive regions

3. Suggest any ONE method to test for contamination.

   Answer:
   Run the contig on BLAST and check the species of the hits. Also check for percentage identity and query coverage. Have a look at the alignment as well.

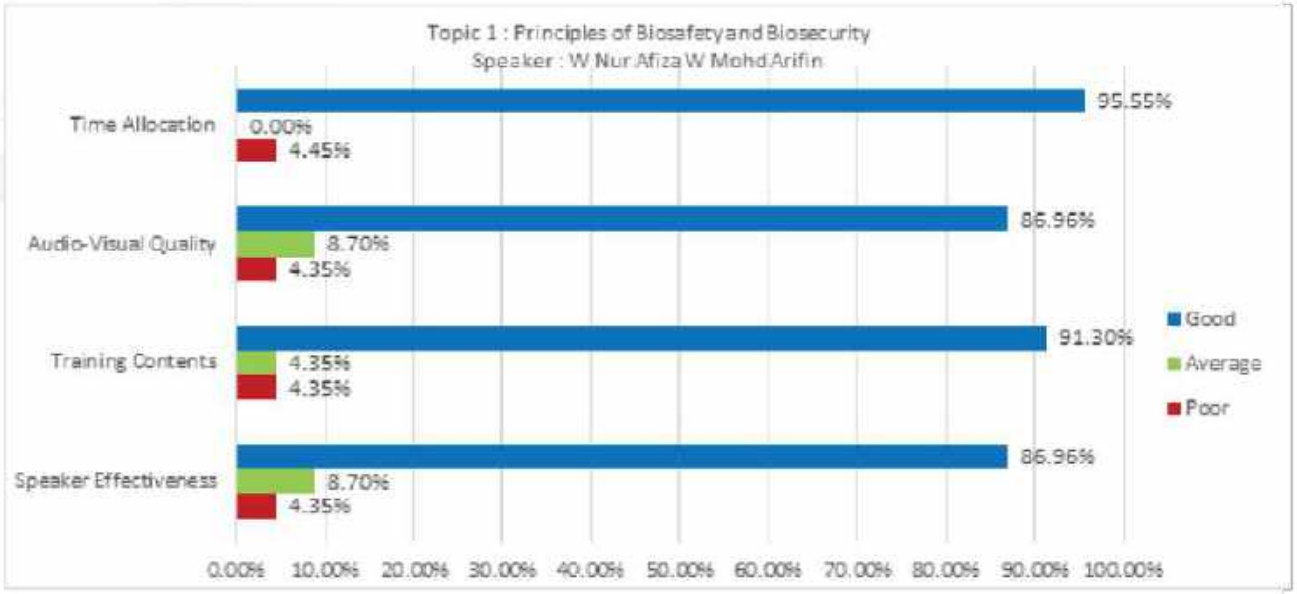
Dr Tan's Feedback:

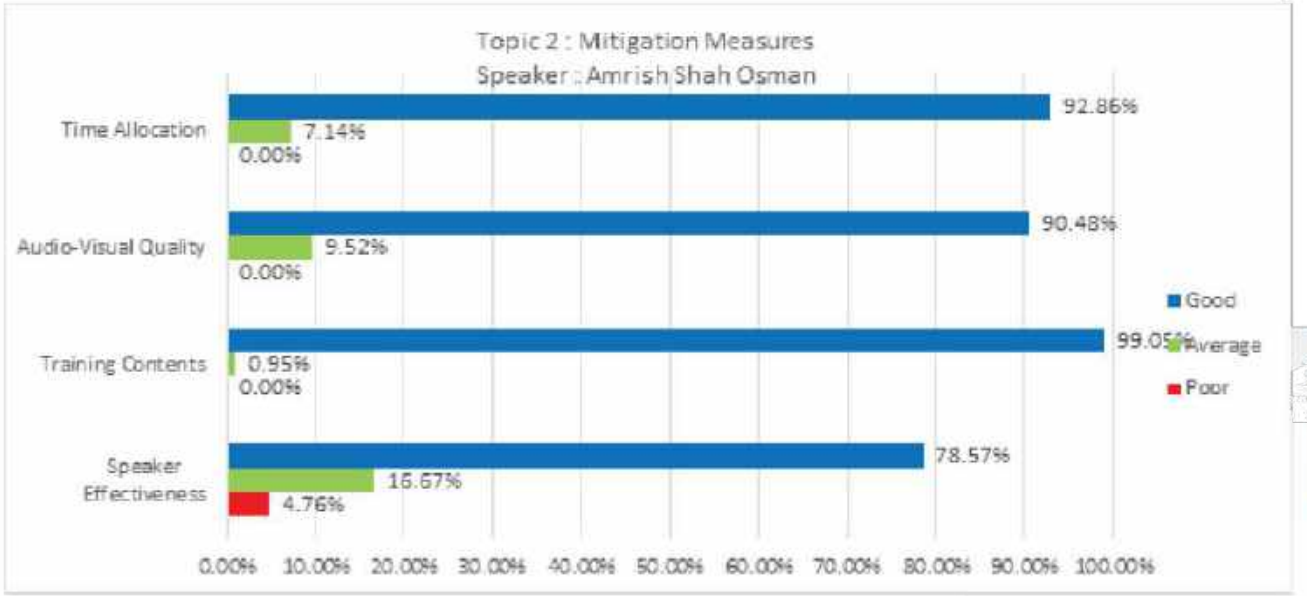
- Participant showed ability to perform assembly and perform post-processing.
- Overall performance is good.

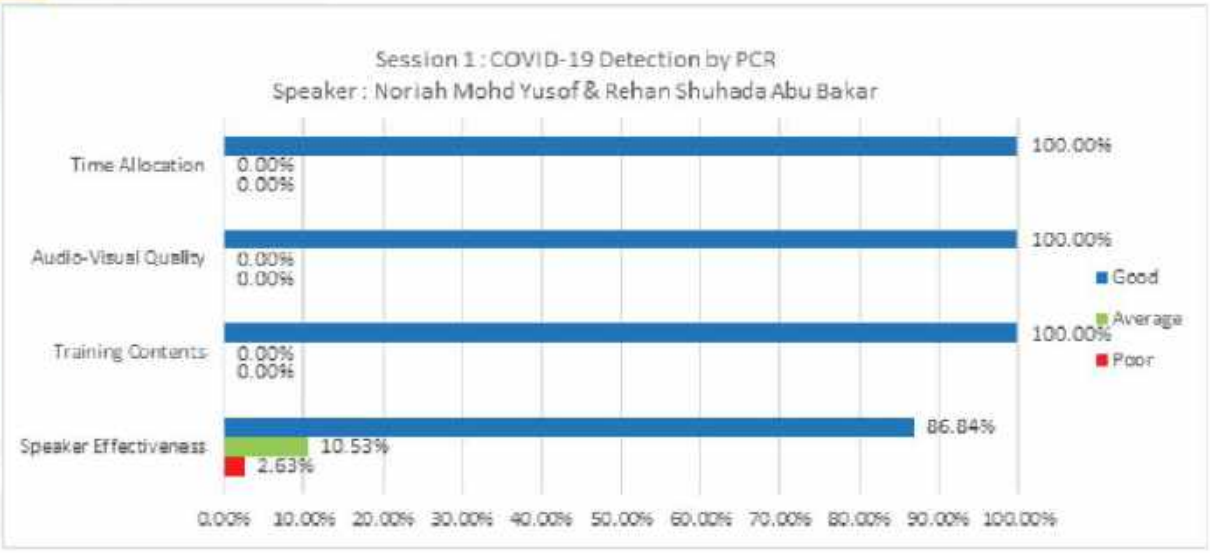
5.4 EVALUATION REPORT

There are 12 session throughout this training covered about Principle of Biosafety and Biosecurity, COVID-19 Detection, isolation and sampling strategy, whole genome sequencing laboratory process and SARS-CoV-2 genome analysis. Training was evaluated on each session and speaker.

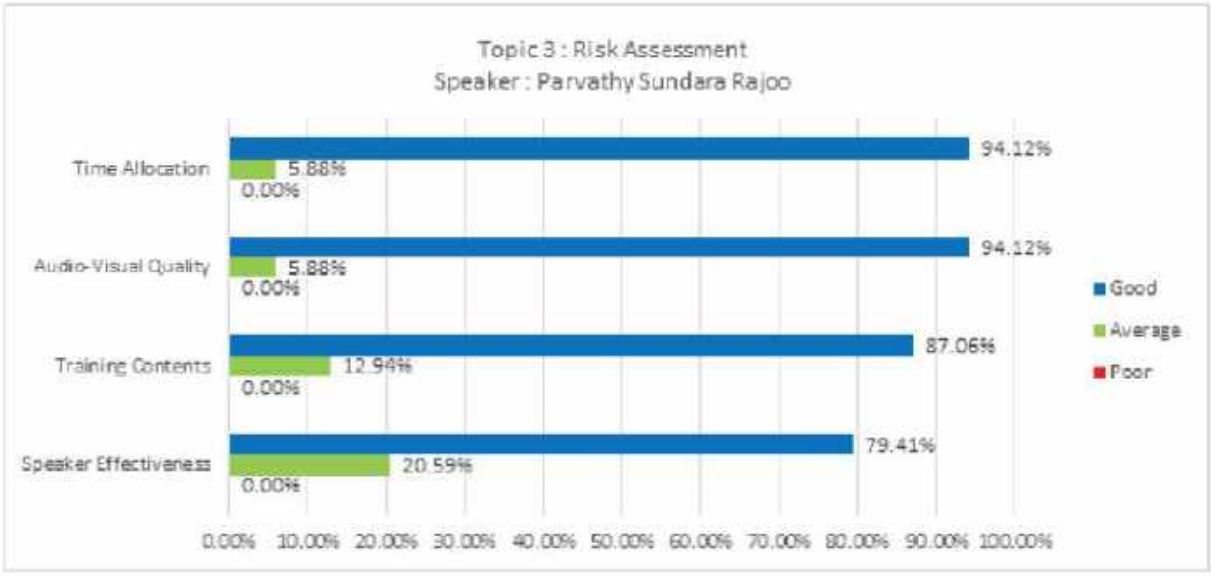
Biosafety & Biosecurity

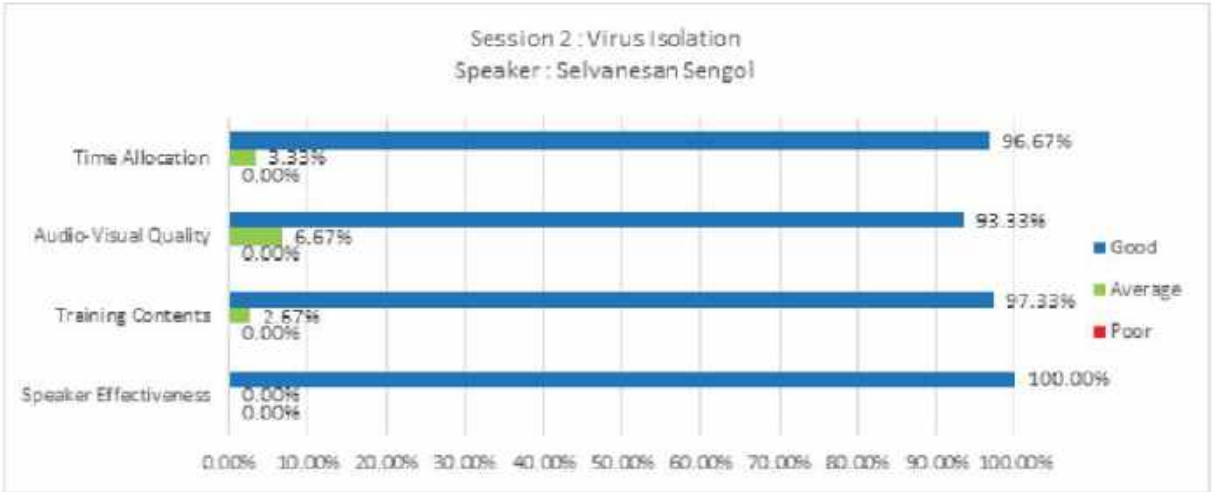




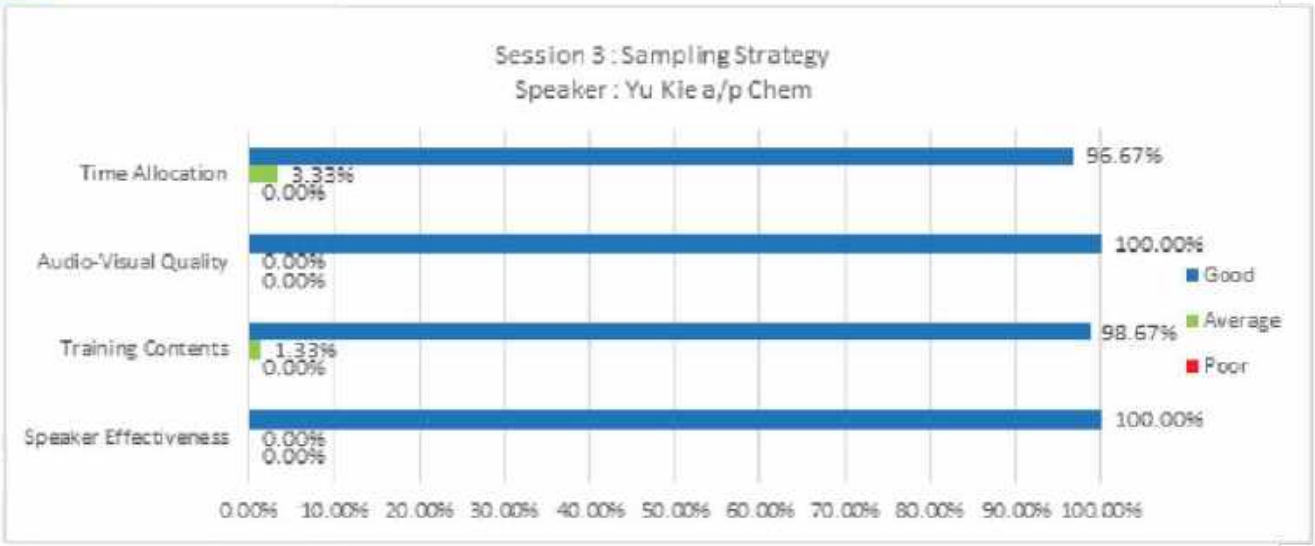


COVID-19 Detection and Isolation

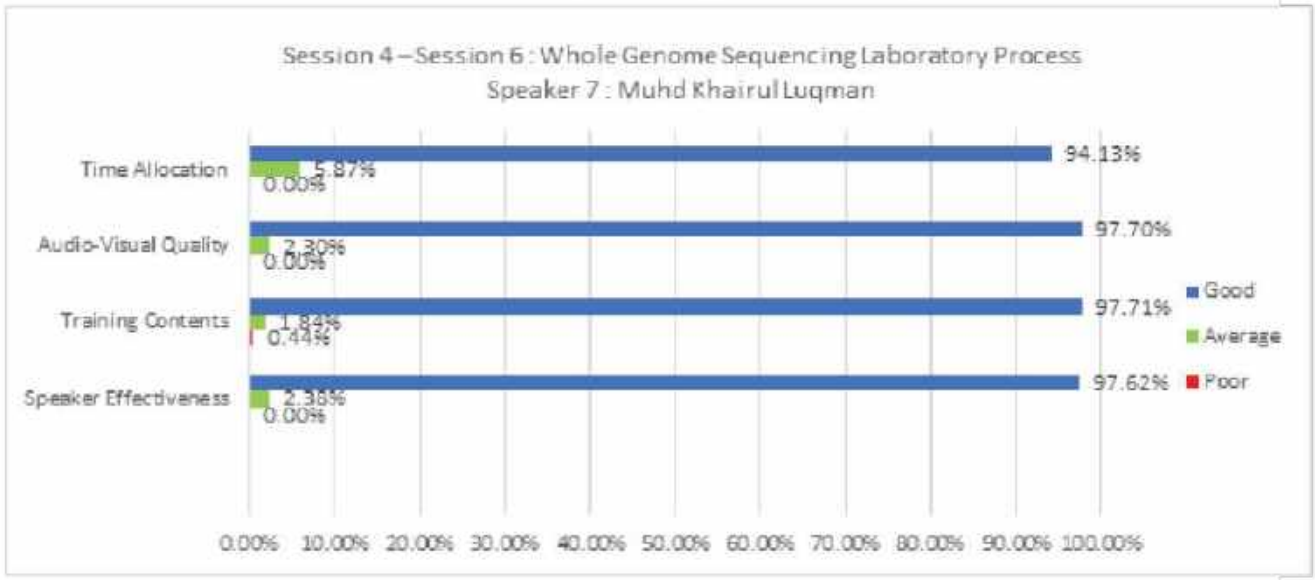


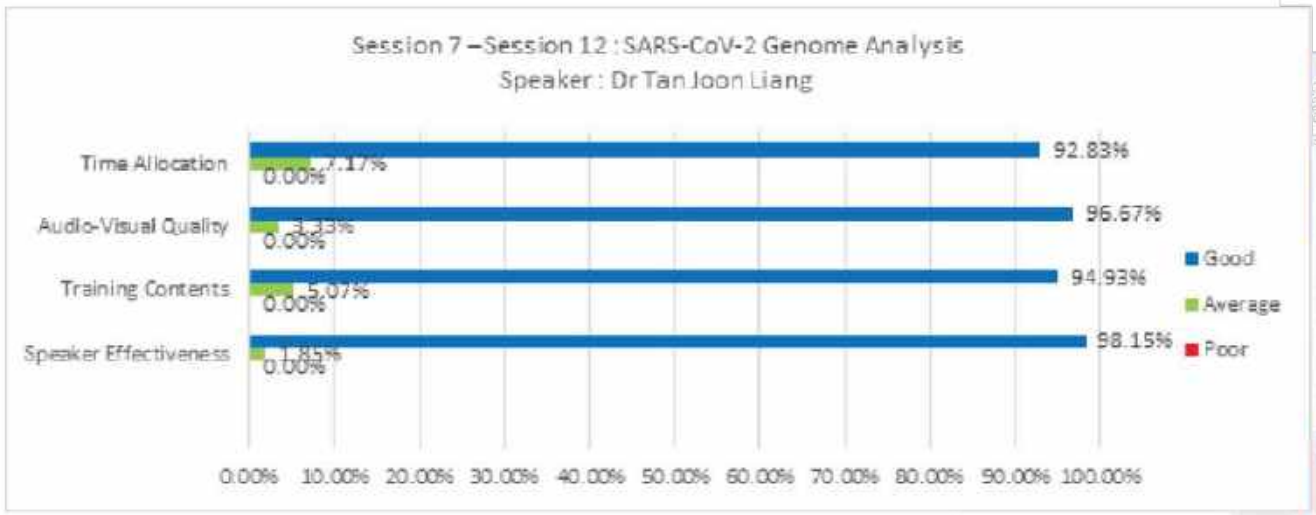


COVID-19 WGS Samples Collection



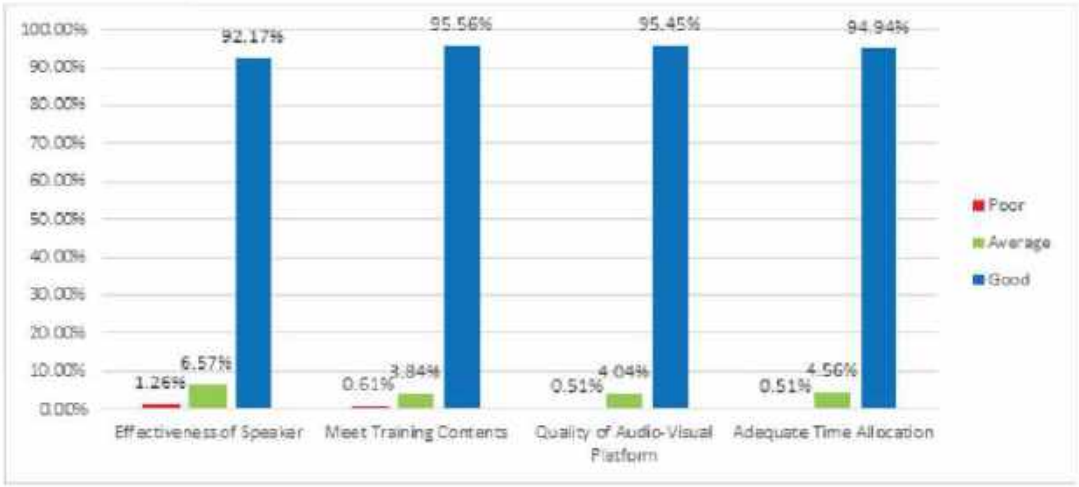
Whole Genome Sequencing Laboratory Process





5.5 CONCLUSION

Overall, the evaluation result of this training were excellent where good feedback received was 92.17% for speaker effectiveness, 95.56% for training contents, 95.45% for audio visual quality and 94.94% for the time allocation component. Details of evaluation score is describe in the graph below:



Lesson Learned

1. The attendeed gained sufficient knowledge and practice to teach and share the techniques with colleagues at their respective laboratories.
2. The troubleshooting and consultation sessions allowed participants with different requirements to pose their queries regarding methods, equipment and analyses.
3. Networking among AMS laboratories technical personnel for future collaboration.

Course Evaluation: Comments and Feedback

The majority of participants commented on this training as follow:

1. Presentation is easy to understand.
2. Full of information.
3. The training improve participants knowledge.
4. Video sharing is attractive.
5. Clear and concise explanation.
6. Provide fundamenta bioinformatics knowledge.

Suggestion For Improvement

1. Conducting on-site training could be more effective.
2. Provide more illustration during the training to better understand the topics presented.
3. Evaluation form should be provided early of the training.
4. Add video or flow how to interpret performance parameter.
5. More practical sessions are needed in order to understand and proficient in genome analysis.
6. The organizer should notify the participant in advance if any preparations, such as registering an account in any tools system used for training.
7. It is necessary to provide video recordings in bioinformatics topics because what is written in the manual may differ form what the trainer explains during the sessions. The video will help with the troubleshooting process.

6.0 HANDS – ON TRAINING ACTIVITIES

Each country is allocated some period of time throughout the training to perform the sequencing laboratory processing and procedure. They are also have been given privilege for any troubleshooting and consultations from the expert if they encounter any issues during their hands-on session at their respective laboratory.

MALAYSIA, BRUNEI, CAMBODIA, INDONESIA

6.1 ACTIVITY BY COUNTRY AT RESPECTIVE LABORATORY

- MALAYSIA

Malaysia's participants consist of 7 members from national and regional Public Health Laboratory. They have conducted their hands-on training session at Molecular Unit, Virology Section in National Public Health Laboratory, NPHL using Illumina sequencing platform. A total of 24 SARS-CoV-2 successfully sequenced.

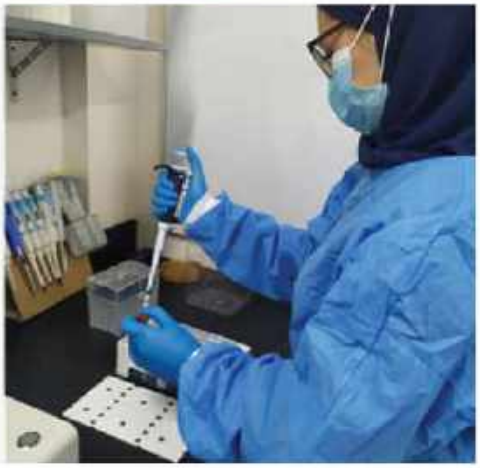





- BRUNEI

Brunei's participants consist of 5 members from Microbial Genomics Services, Clinical Molecular Diagnostic Laboratory for Infectious Disease, Brunei Darussalam. They have conducted their hands-on training session at their respective laboratory using MGI sequencing platform. During the session, they encountered several challenges during result analysis such as abundance of N's in the middle of the consensus sequence and some issues with the IGVD auto-analyzing. Through consultations by the expert, they manage to obtain good quality sequences.






- CAMBODIA

A total of 6 participants from National Institute of Public Health, Cambodia have conducted their hands-on training session at their respective laboratory in Phnom Penh using Illumina sequencing platform. A total of 23 SARS-CoV-2 specimens have been processed and successfully sequenced.

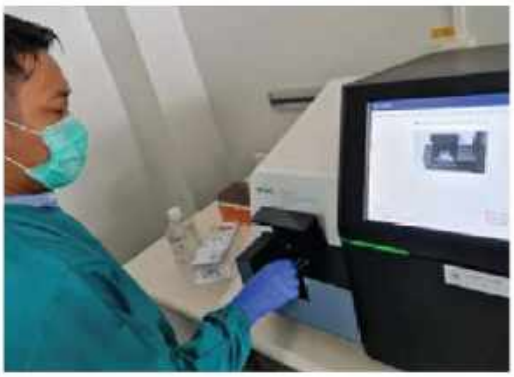





- INDONESIA

Indonesia's participants consist of 4 members from National Reference Laboratory Indonesia have conducted their hands-on training session at their respective laboratory using Illumina sequencing platform. They have successfully performed and analyze the SARS-CoV-2 sequence at the end of the training.



6.2 SARS-COV-2 GENOME ANALYSIS

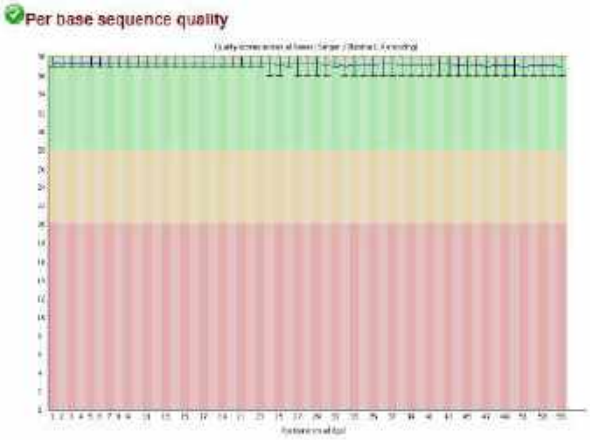
Analysis of SARS-CoV-2 genome which comprise sequence quality analysis, contamination checking, assembly and variant calling.

Sample

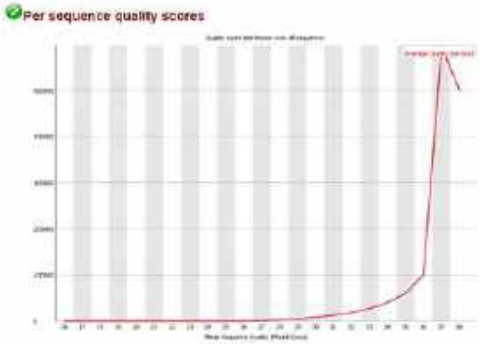
Raw Quality

Sequence quality for the first few bases are lower in quality, and with read length of 76bp

Pre-processed Quality

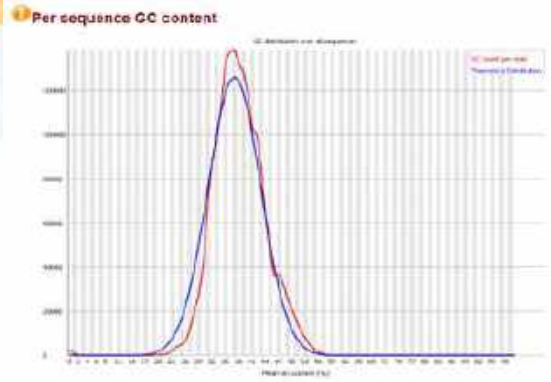


After pre-processing step, average sequence quality relative to read position has increased

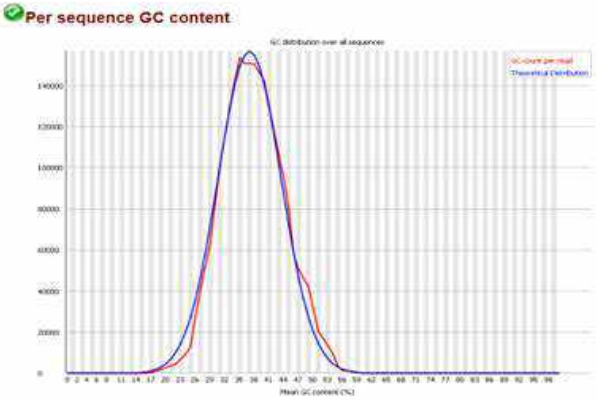


Low quality per base sequence content during the initial read and end of the read

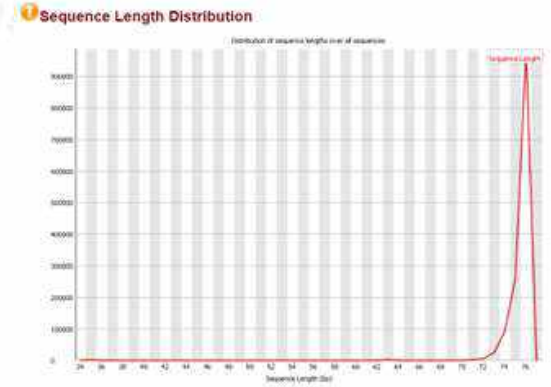
Pre-processing has removed low quality base pairs in the beginning and end of each read



Per sequence GC content in the read were not

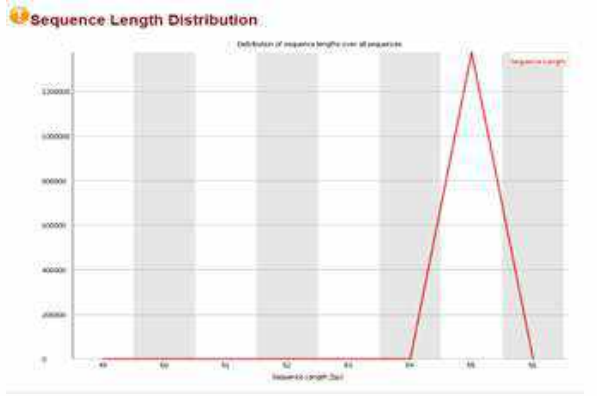


Pre-processing step has helped bring the GC content close to the t



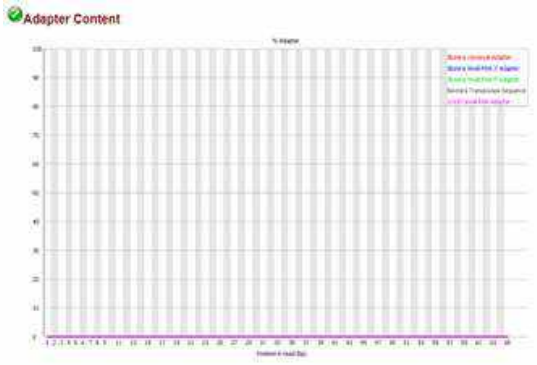
evenly distributed

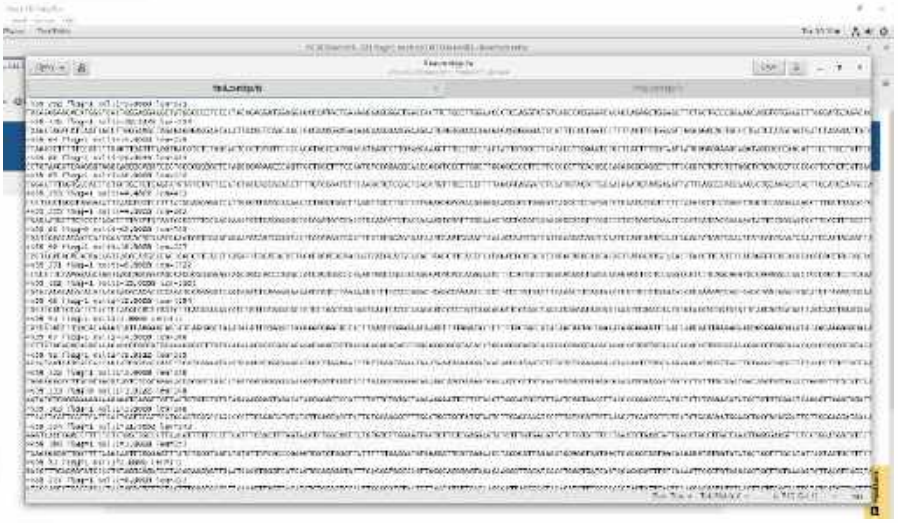
Sequence length distribution were varied



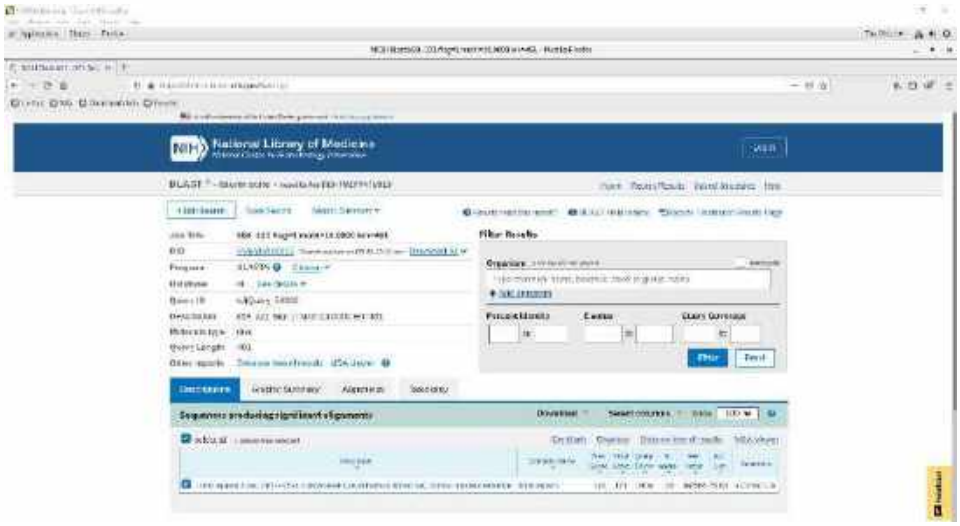
heoretical distribution

Sequence length distribution were trimmed to only one length

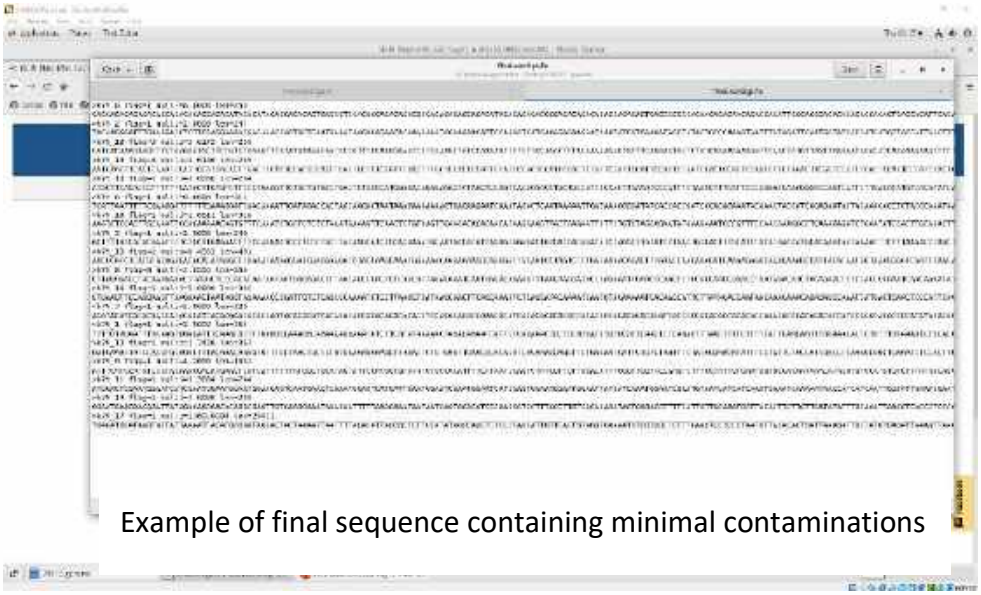




Some sequences have a lot of contamination



Based on NCBI database, the sequences were found to be human contamination



Example of final sequence containing minimal contaminations

7.0 CLOSING CEREMONY

'None of this achievement would have taken place without the great support, guidance and constant motivation by ASEAN EOC Network, ASEAN Secretariat and by German government, GIZ Thailand Team. And, not forgetting the lead country team and consultants from Malaysia who has put in all the effort from planning, execution and successful completion of this project. Congratulations once again!'- Datuk Dr. Norhayati Binti Rusli, Deputy Director of Health (Public Health)

Ladies and gentlemen, there is a saying in Bahasa Melayu "*setiap pertemuan pasti ada perpisahan*", which means "when there is a meeting then there will be goodbye". Hope to meet all of you again in another knowledge sharing session. Till then, Thank you.'- Dr. Fatanah Binti Ismail, Director of NPHL

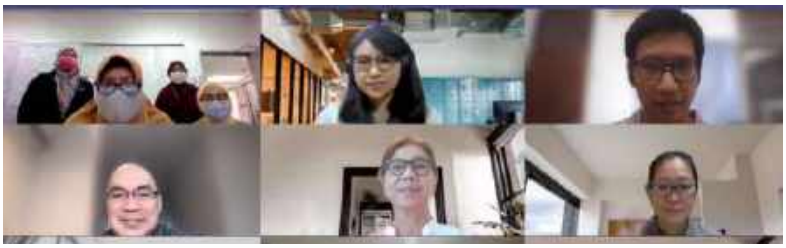
We received evaluation with a good feedback and comment from AMS participant that we take note and make use of this for further collaboration and potential activity that will be implemented under the new project.'- Pouchaman Wongsanga, GIZ Thailand.

Congratulations!!!! Excellent teamwork and collaborative effort made the training successful, effective, and impactful.

8.0 BEHIND THE SCENE - Planning and Execution

"Laboratory personnel must acquire the skills to perform the whole genome sequencing of the SARS CoV2 virus, their phylogenetic analysis and able to link this information together with an epidemiological data. By doing this, we hope it would guide the public health in decision making and responses at a near-real time.

I hope that this project in Strengthening Laboratory Capacity on SARS-CoV-2 Genomics Training for ASEAN Countries has achieved their target and has enhanced communication among AMS to understand the circulating SARS-CoV-2 variants in the region."- **Former NPHL Director, Dr Noorliza Binti Mohamad Noordin**



Coordination meeting with ASEAN Secretariat and GIZ Thailand, 26 Jan 2022.

Pre-training Workshop with AMS, 24 Feb 2022

"Thanks for opening our eyes to new stages of opportunity and strength. We will forever be grateful for your guidance."-WGS NPHL Project Team

9.0 SPECIAL ACKNOWLEDGEMENT

DATUK DR NORHAYATI BINTI RUSLI

(Deputy Director General of Public Health, Ministry of Health Malaysia)

DATUK DR CHONG CHEE KHEONG

(Former Deputy Director General of Public Health, Ministry of Health Malaysia)

DR HAJAH NOORLIZA BINTI MOHAMAD NOORDIN

(Former Director of Malaysia National Public Health Laboratory)

DR HANI BINTI MAT HUSSIN

(Deputy Director of Disease Control Division, Ministry of Helath Malaysia)

DR. NOVIA KUSWARA

(Project Coordinator)

10.0 DEEPEST GRATITUDE ACKNOWLEDGEMENT

01 GIZ THAILAND

02 ASEAN SECRETARIAT

03 DIRECTOR NPHL

04 CONSULTANT/EXPERT AND TRAINEES

05 NPHL SECRETARIAT AND IT TEAM

06 ASEAN REPRESENTATIVE BRUNEI

07 ASEAN REPRESENTATIVE CAMBODIA

08 ASEAN REPRESENTATIVE INDONESIA

09 ASEAN REPRESENTATIVE MALAYSA

10 ASEAN REPRESENTATIVE THAILAND

11 ASEAN REPRESENTATIVE VIETNAM

THANK YOU

FOR BEING APART OF

THE JOURNEY

STRENGTHENING LABORATORY

CAPACITY ON COVID – 19

BIO-GENOMICS FOR ASEAN COUNTRIES:

A HANDS-ON TRAINING AND KNOWLEDGE SHARING


STRENGTHENING LABORATORY CAPACITY ON COVID-19 BIO
GENOMIC FOR ASIAN COUNTRY(S) : A HANDS ON TRAINING AND
KNOWLEDGE SHARING WORKSHOP