



Next Generation Sequencing (NGS) for Influenza Viruses and RSV

Yi-Mo Deng

Training workshop in whole genome sequencing and bioinformatic analysis of
respiratory viral pathogens, Pune, 3-7 June 2024



WHO Collaborating Centre
for Reference and
Research on Influenza
VIDRL



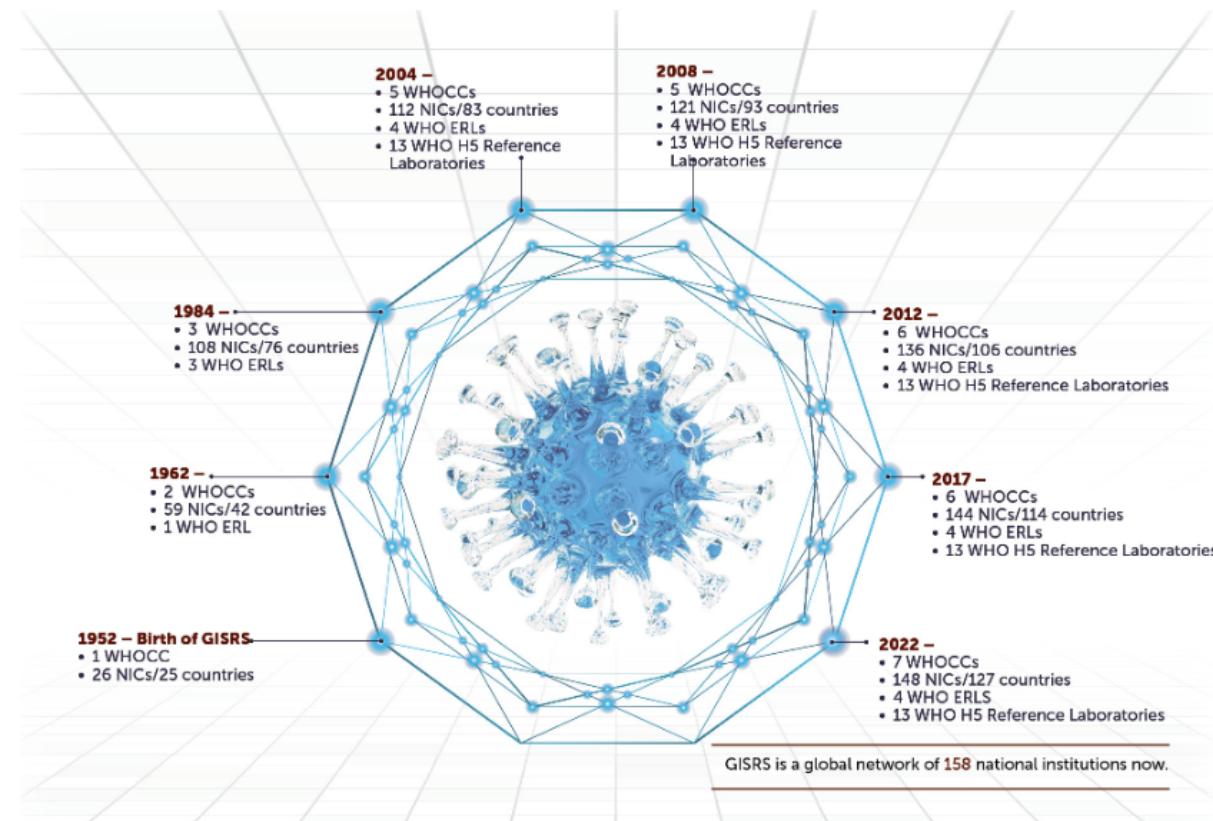
Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS



- GISRS: Global Influenza Surveillance and Response System
- Created in 1952, initially only for influenza



- Expanded to RSV, SARS-CoV-2, and other respiratory viruses

Key activities: Annual Influenza vaccines recommendations for Northern and Southern Hemispheres

- WHO recommended vaccines for 2024-25 northern hemisphere influenza season:
 - Egg-based vaccines:
 - An A/Victoria/4897/2022 (H1N1)pdm09-like virus
 - An A/Thailand/8/2022 (H3N2)-like virus
 - A B/Austria/1359417/2021 (B/Victoria lineage)-like virus
 - A B/Phuket/3073/2013 (B/Yamagata lineage)-like virus
 - Cell culture- or recombinant-based vaccines:
 - An A/Wisconsin/67/2022 (H1N1)pdm09-like virus
 - An A/Massachusetts/18/2022 (H3N2)-like virus
 - A B/Austria/1359417/2021 (B/Victoria lineage)-like virus
 - A B/Phuket/3073/2013 (B/Yamagata lineage)-like virus

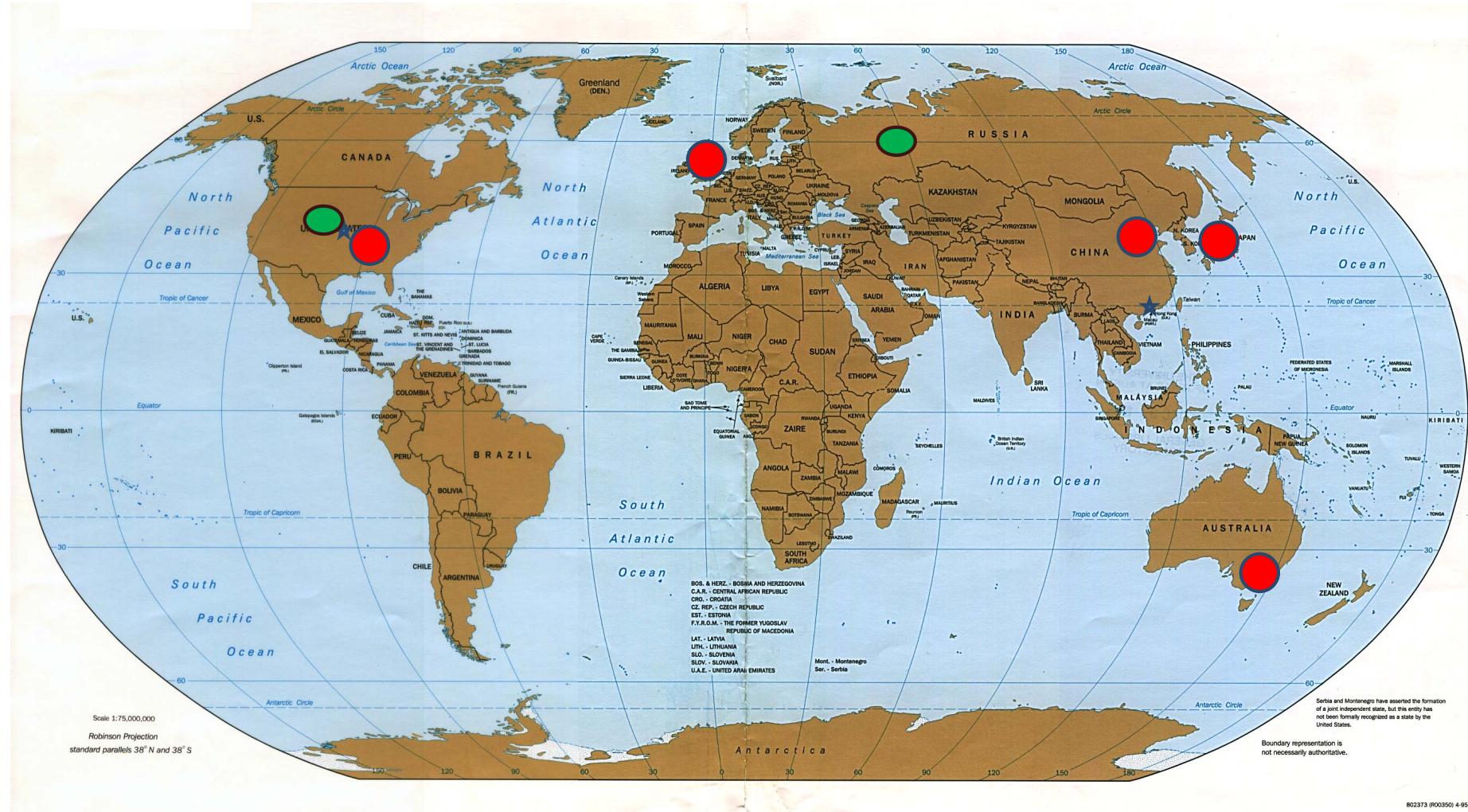
Key activities: respond to potential pandemics

- Closely monitor influenza viruses from both human and animal
- Risk assessment of novel influenza viruses from avian and other mammals
- Capacity building for integrated respiratory virus surveillance, testing and sequencing
- Prepare for and respond to current and future outbreaks with epidemic and pandemic potential

Outline

- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

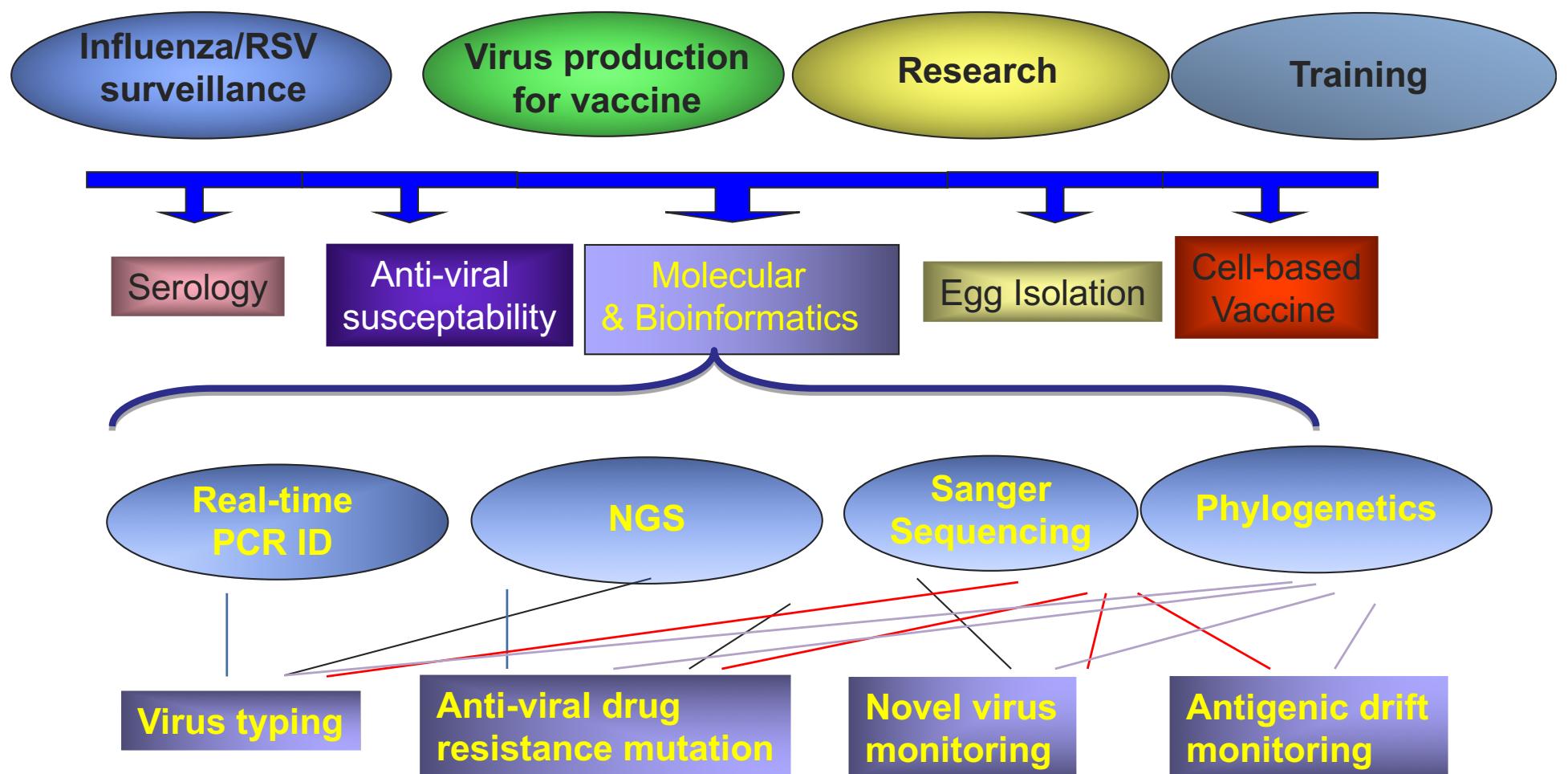
WHO CC for Reference and Research on Influenza



WHO CC Melbourne at the Peter Doherty Institute for Infection and Immunity



Major activities for WHOCC Melbourne



Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
 - Influenza viruses
 - Influenza virus NGS
 - RSV
 - RSV WGS
 - Challenges
 - WHO Guidance for NGS

NGS, revolution of sequencing technology



Massive parallel high-throughput sequencing, deep sequencing,
non-target specific sequencing possible

- Sequence by synthesis or direct single molecule sequencing
- Various platforms of NGS developed over the years
 - Illumina**, Pacific Biosciences, Roche 454, Ion Torrent,
 - Third Generation Sequencing: **Oxford Nanopore Technology (ONT)**

NGS workflow



1. Library preparation

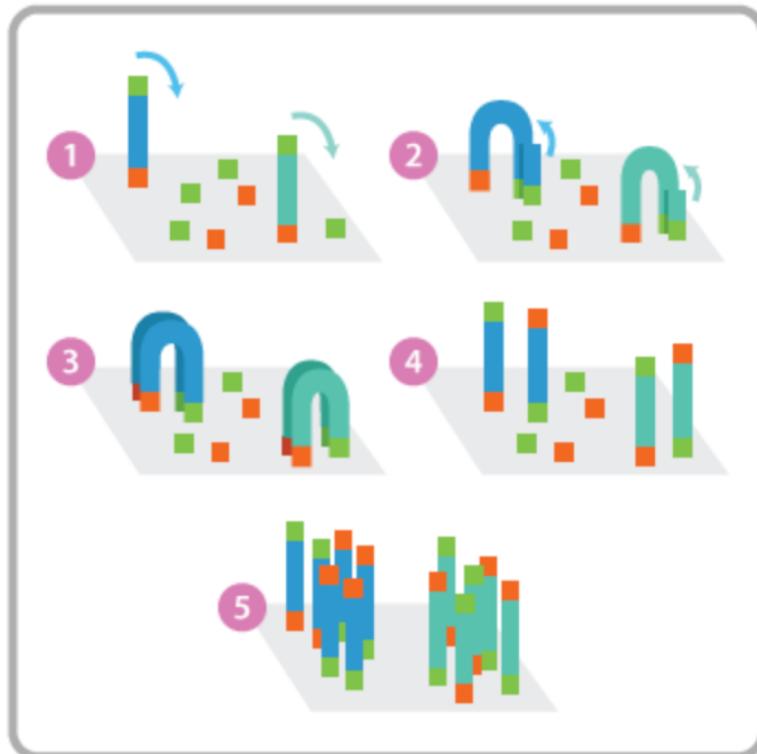
- amplicon based: generate multi segments PCR products
- Fragmentation and adaptor ligation (bar-coded)



NGS workflow

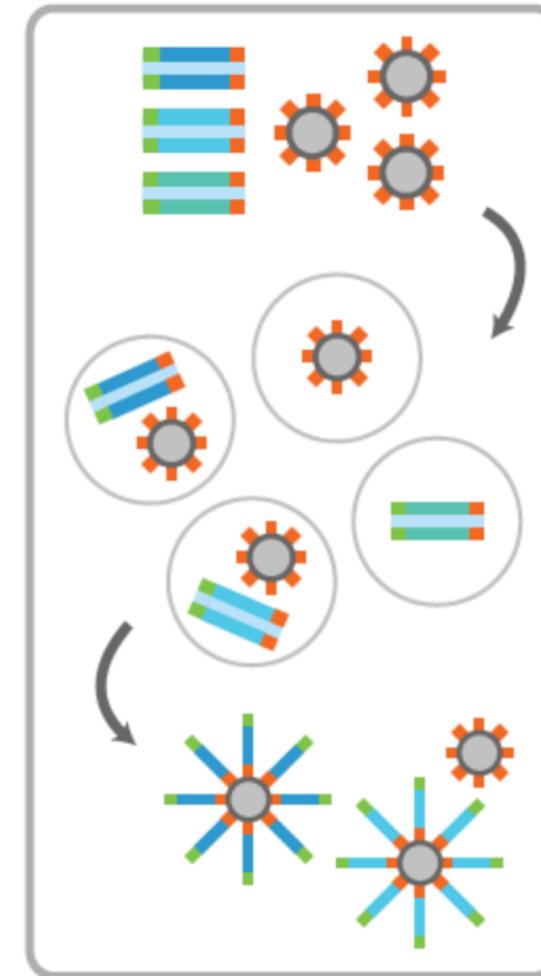
2. Clonal amplification

Bridge PCR



Illumina

Emulsion PCR



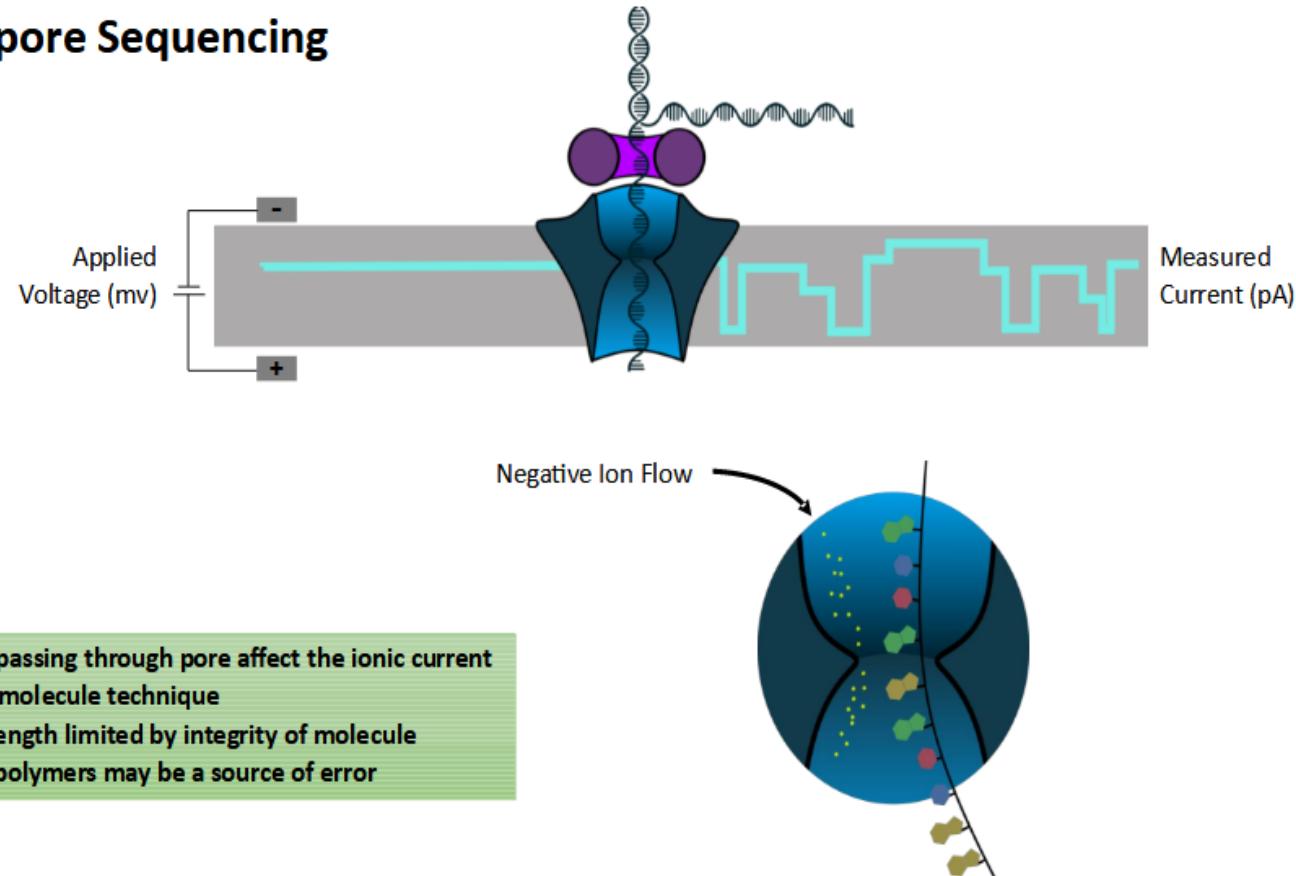
Ion Torrent

Third Generation Sequencing - ONT



No clonal amplification - Unique for ONT

Nanopore Sequencing





3. Sequencing millions of the same short reads within each clone

- Different NGS platforms using different sequencing methods
- Each has its own advantages and limitations

Signal processing, base calling and data assembly

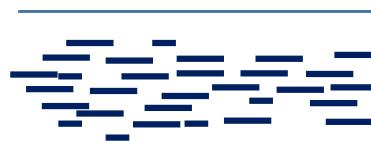


4. Data processing and analysis

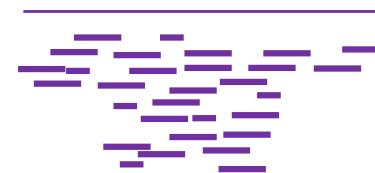
Barcode 001



Barcode 002



Barcode 003

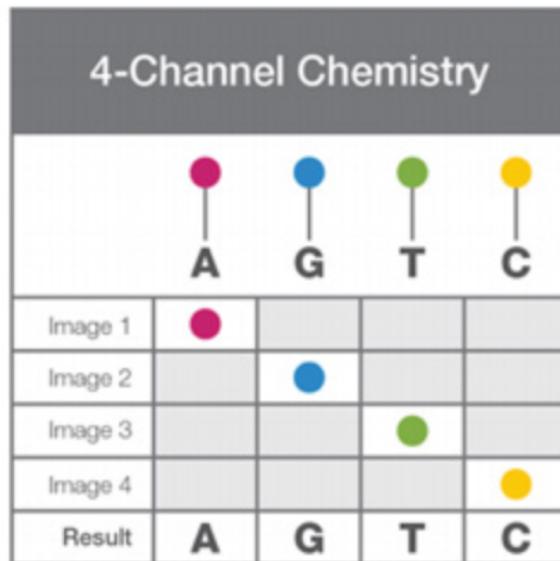


- Signal converting and barcode decoding
- De novo analysis
- mapping on reference sequence
- bioinformatics pipelines

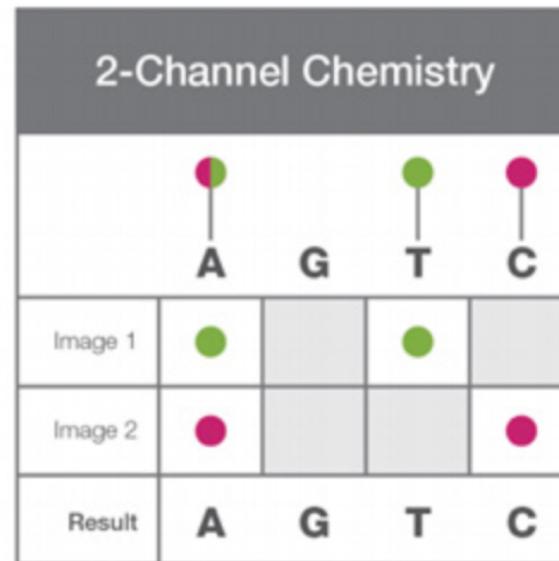
Illumina Platforms

Evolution of Illumina Sequencing Chemistry

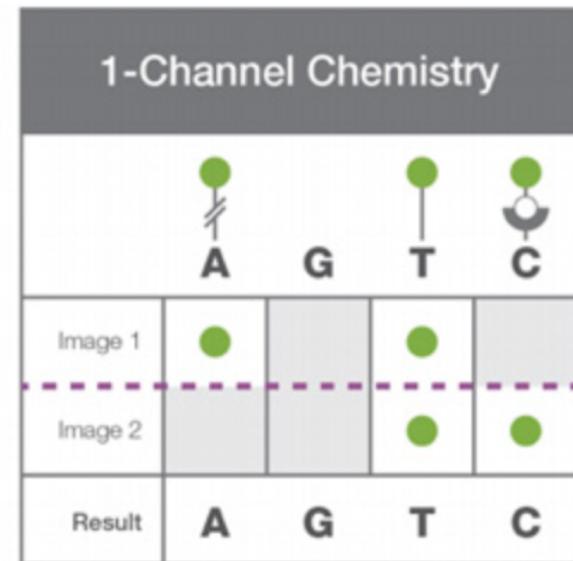
HiSeq/NextSeq/MiSeq



MiniSeq

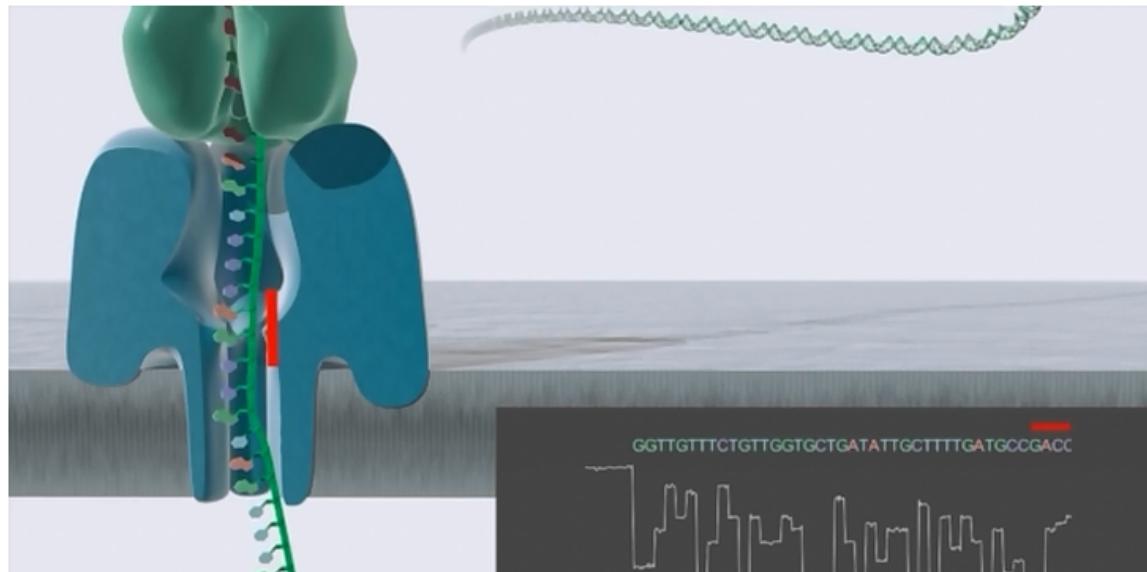


iSeq



Third Generation Sequencing – Oxford Nanopore platforms

- Nanopore-based electronic systems for analysis of single molecules
- No DNA amplification
- Long read length
- Real-time sequencing
- Capable of sequencing RNA directly
- Quicker and simpler library preparation
- Portable sequencer



Different Nanopore Devices



Flongle



MinION Mk1B



MinION Mk1C



GridION

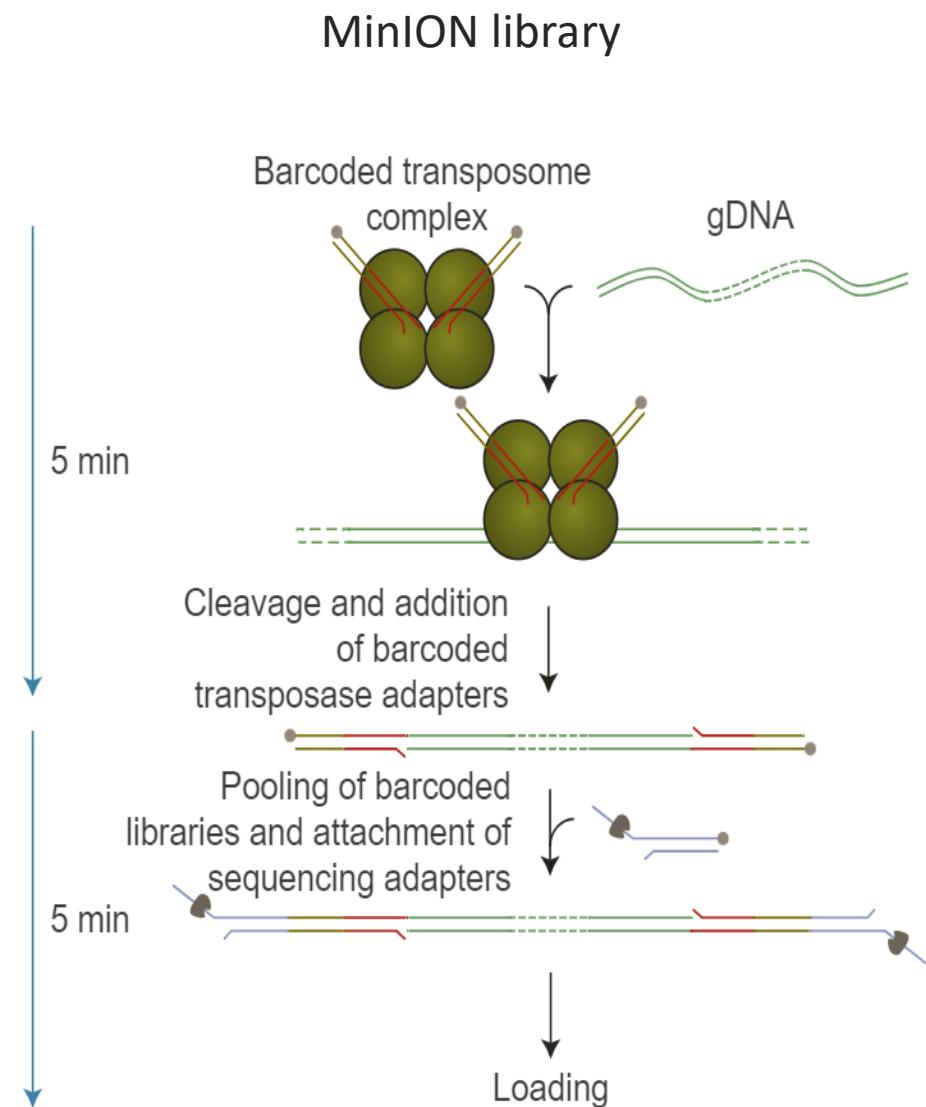


PromethION



VolTRAX

Illumina vs MinION



Illumina vs MinION



	Illumina	MinION
Workflow time	2-3 days	1 day
Required quantity for Input material	Low	High
Sequence accuracy	Very high	High
Length of sequence	Short reads	Short or long reads
Data generation	Not real-time	Real-time
Sequencing time	Long	Quick
Price	Higher	Lower
Sequence capacity	Very high depend on platform	Lower

Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

Influenza Viruses

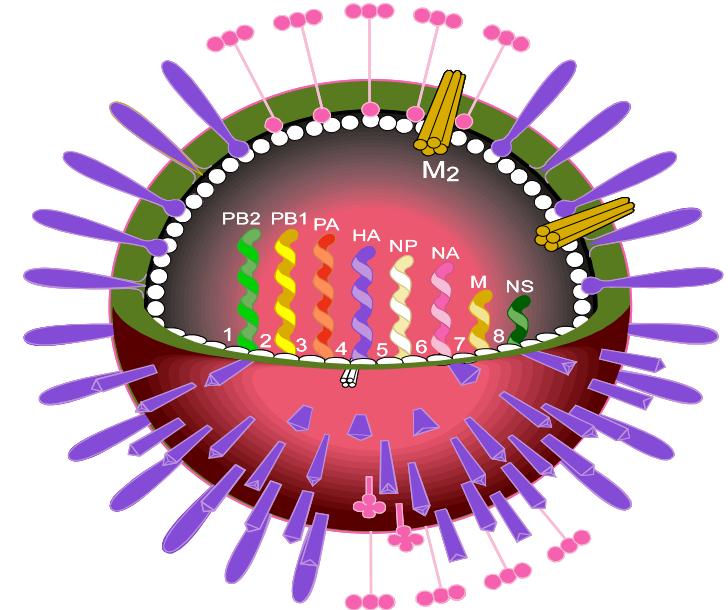


- Influenza viruses cause seasonal flu epidemics every year
- Some can be severe, or even lead to death
- Four types of influenza viruses
 - Influenza A:
 - cause both seasonal flu and pandemic flu in humans;
 - also infect other animals, like avian, pig, horse, etc
 - Influenza B:
 - cause seasonal flu in humans
 - Influenza C:
 - only cause mild symptom in humans, not epidemic
 - Influenza D:
 - only known to affect cattle with spillover to other animals
 - not known to infect human

Characteristics of influenza virus

- Influenza A and Influenza B

- Segmented negative sense RNA genome
- Highly variable surface glycoproteins
 - Haemagglutinin (HA)
 - Neuraminidase (NA)
- Influenza A subclassified by HA and NA
 - 18 HA and 11 NA types
 - Many subtypes exist in avian, cross-species infection
- Influenza B – two lineages
 - Yamagata lineage and Victoria lineage
- High mutation rate - Antigenic drift
- Potential for reassortment between influenza viruses - Antigenic shift, pandemic threat
 - H1N1pdm09



How influenza viruses evade host immune system



- Antigenic drift



- happens with both A & B viruses
- effects HA and NA genes/proteins
- fairly gradual change – reason for updating vaccines
- due to mutations (lack of proof reading RNA polymerase)
- causes continued annual outbreaks and epidemics

- Antigenic shift

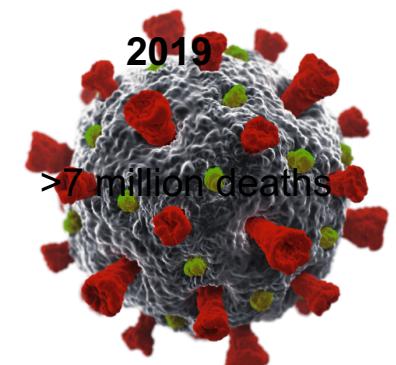
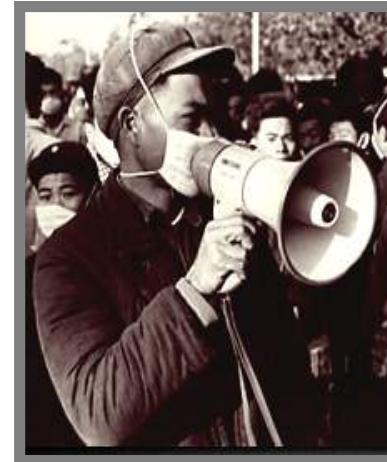


- only occurs in influenza A viruses
- sudden change
- due to reassortment of viruses
 - eg several avian viruses – H9N2 + H7N7 + H?N9 = H7N9
 - eg swine viruses + human viruses = H1N1 2009 pandemic
- can create a novel virus - human population susceptible
- may lead to explosive spread – causing pandemics

Major pandemics 20th - 21st centuries



- Novel virus, no prior immunity towards the new virus
- Ability to infect from human to human
- High transmissibility
- High impact to human, high death rate



1918

A(H1N1)

40 to 50 million deaths

1957

A(H2N2)

1 to 4 million deaths

1968

A(H3N2)

1 to 4 million deaths

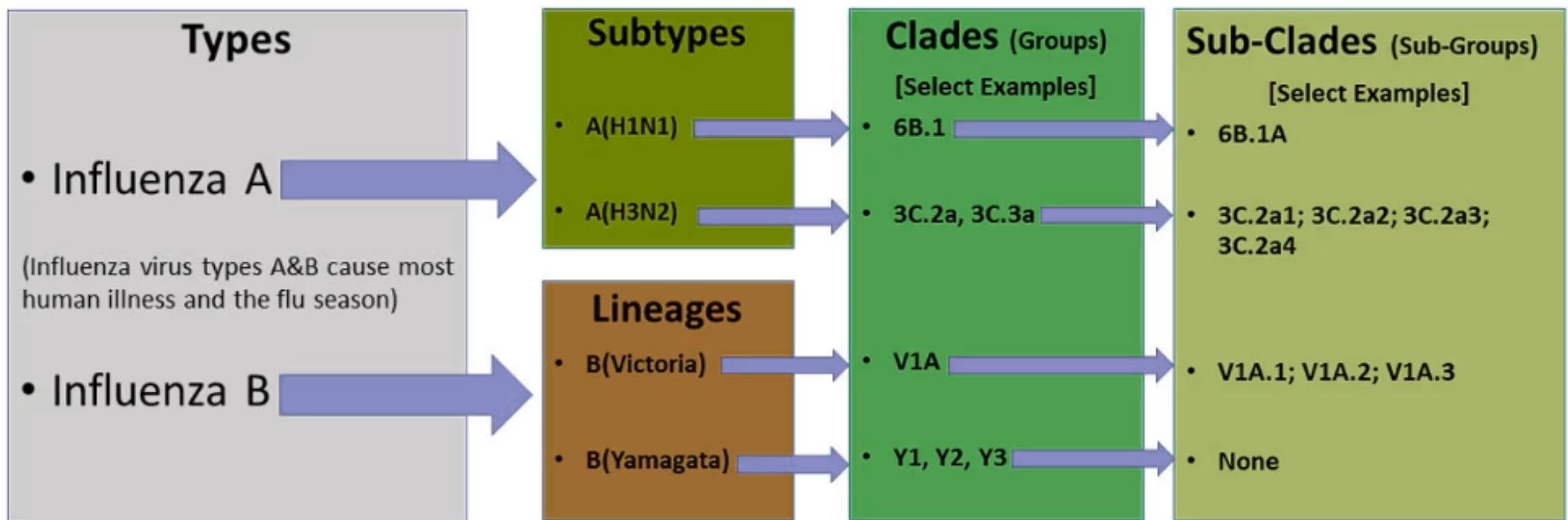
2009

A(H1N1)pdm09

<0.1 million deaths

SARS-CoV-19

Human seasonal influenza viruses



<https://www.cdc.gov/flu/about/viruses/types.htm>

Information from influenza genome sequencing



- Diversity: Evolutionary trends
 - New clades/predominant clades
- Key Mutations: Antigenic sites, Receptor Binding sites
- Resistance markers for antivirals (mixture):
 - Neuraminidase inhibitors (Oseltamivir), M2 channel blocker (adamantanes), polymerase inhibitors (Baloxivir)
- Virus ID and Reassortment:
 - Novel subtypes H5Nx, H3N2v etc

Some of the key influenza gene substitutions



Segments	Example of significant amino acid substitution	Importance to virus characterisation
PB2 (polymerase basic protein 2)	E627K, D701N	Increased transmissibility to mammalian host/mammalian adaptation
PA (polymerase acidic protein)	I38T/M/F	Reduced susceptibility to Baloxavir
HA (hemagglutinin)	various antigenic sites	Antibody escape mutations Receptor specificity
NA (neuraminidase)	H275Y (N1)	Reduced susceptibility to oseltamivir
M1 &2 (matrix proteins 1 & 2)	L26F, V27A, A30V/T/S, S31N , G34E for M2	Reduced susceptibility to Admantene

Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

Influenza viruses amplification for NGS

- Full genome amplification of influenza viruses Multi-RTPCR

JOURNAL OF VIROLOGY, Oct. 2009, p. 10309–10313
0022-538X/09/\$08.00 + 0 doi:10.1128/JVI.01109-09
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

Vol. 83, No. 19

Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses^v

Bin Zhou,^{1,2} Matthew E. Donnelly,¹ Derek T. Scholes,¹ Kirsten St. George,¹ Masato Hatta,³ Yoshihiro Kawaoka,^{3,4} and David E. Wentworth^{1,2,*}

Wadsworth Center, New York State Department of Health,¹ and Department of Biomedical Sciences, School of Public Health, State University of New York,² Albany, New York 12201; Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin—Madison, Madison, Wisconsin 53706;³ and Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan⁴

Received 31 May 2009/Accepted 8 July 2009

Pandemic influenza A viruses that emerge from animal reservoirs are inevitable. Therefore, rapid genomic analysis and creation of vaccines are vital. We developed a multisegment reverse transcription-PCR (M-RTPCR) approach that simultaneously amplifies eight genomic RNA segments, irrespective of virus subtype. M-RTPCR amplicons can be used for high-throughput sequencing and/or cloned into modified reverse-genetics plasmids via regions of sequence identity. We used these procedures to rescue a contemporary H3N2 virus and a swine origin H1N1 virus directly from human swab specimens. Together, M-RTPCR and the modified reverse-genetics plasmids that we designed streamline the creation of vaccine seed stocks (9 to 12 days).



Universal Influenza B Virus Genomic Amplification Facilitates Sequencing, Diagnostics, and Reverse Genetics

Bin Zhou,^a Xudong Lin,^a Wei Wang,^a Rebecca A. Halpin,^b Jayati Bera,^b Timothy B. Stockwell,^a Ian G. Barr,^b David E. Wentworth^a
Virology, J. Craig Venter Institute, Rockville, Maryland, USA;^a WHO Collaborating Centre for Reference and Research on Influenza, Melbourne, Australia^b

Although human influenza B virus (IBV) is a significant human pathogen, its great genetic diversity has limited our ability to universally amplify the entire genome for subsequent sequencing or vaccine production. The generation of sequence data via next-generation approaches and the rapid cloning of viral genes are critical for basic research, diagnostics, antiviral drugs, and vaccines to combat IBV. To overcome the difficulty of amplifying the diverse and ever-changing IBV genome, we developed and optimized techniques that amplify the complete segmented negative-sense RNA genome from any IBV strain in a single tube/well (IBV genomic amplification [IBV-GA]). Amplicons for >1,000 diverse IBV genomes from different sample types (e.g., clinical specimens) were generated and sequenced using this robust technology. These approaches are sensitive, robust, and sequence independent (i.e., universally amplify past, present, and future IBVs), which facilitates next-generation sequencing and advanced genomic diagnostics. Importantly, special terminal sequences engineered into the optimized IBV-GA2 products also enable ligation-free cloning to rapidly generate reverse-genetics plasmids, which can be used for the rescue of recombinant viruses and/or the creation of vaccine seed stock.

Downloaded from <http://jvi.asm.org>

Downloaded from <http://jvi.jci.org>

- Partial genome amplification of influenza A and B viruses multi-RTPCR



VIROLOGY

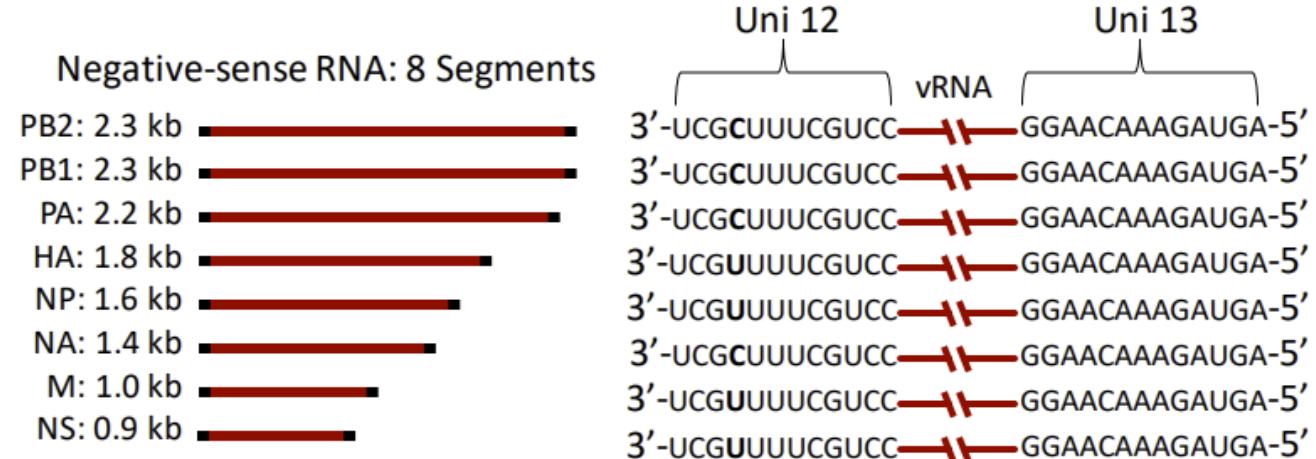
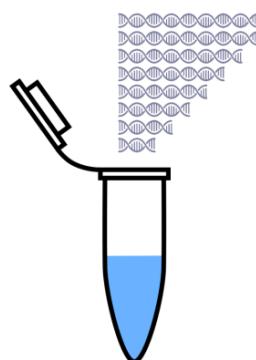
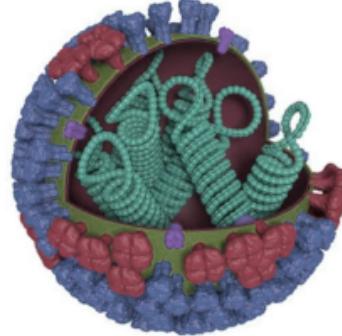


Multiplex Reverse Transcription-PCR for Simultaneous Surveillance of Influenza A and B Viruses

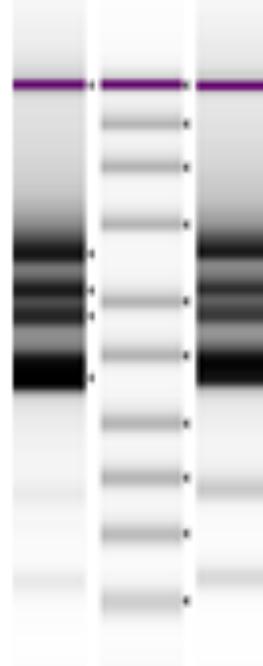
Bin Zhou,^{a,b} Yi-Mo Deng,^c John R. Barnes,^d October M. Sessions,^e Tsui-Wen Chou,^a Malanla Wilson,^d Thomas J. Stark,^d Michelle Volk,^a Natalie Spirason,^c Rebecca A. Halpin,^b Uma Sangamathi Kamara,^e Tao Ding,^a Timothy B. Stockwell,^b Mirella Salvatore,^f Elodie Ghedin,^{a,g} Ian G. Barr,^c David E. Wentworth^{b,d}

Center for Genomics and Systems Biology, Department of Biology, New York University, New York, New York, USA;^a J. Craig Venter Institute, Rockville, Maryland, USA;^b World Health Organization Collaborating Centre for Reference and Research on Influenza, The Peter Doherty Institute for Infection and Immunity, Melbourne, Australia;^c Influenza Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA;^d Program in Emerging Infectious Diseases, Duke-NUS Graduate Medical School Singapore, Singapore;^e Department of Medicine and Department of Healthcare Policy and Research, Weill Cornell Medical College of Cornell University, New York, New York, USA;^f College of Global Public Health, New York University, New York, New York, USA^g

Flu WGS MRT-PCR: multi-segment RT-PCR with Flu universal primers



FluA FluB

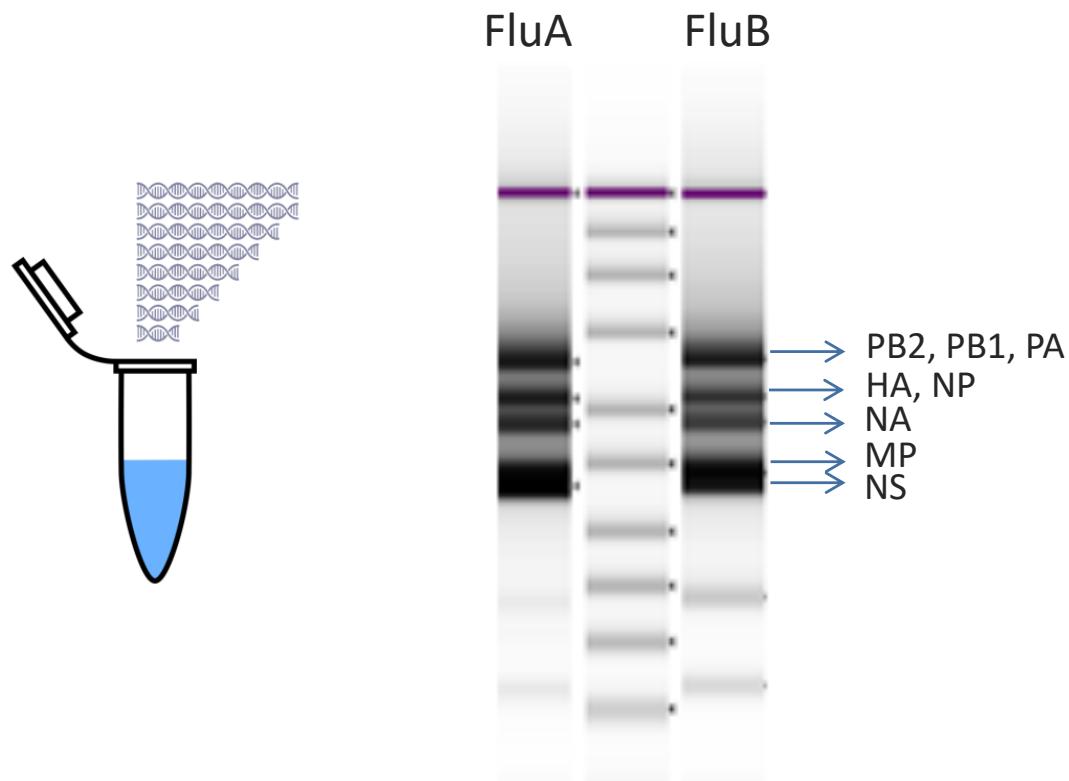


- PB2, PB1, PA
- HA, NP
- NA
- MP
- NS

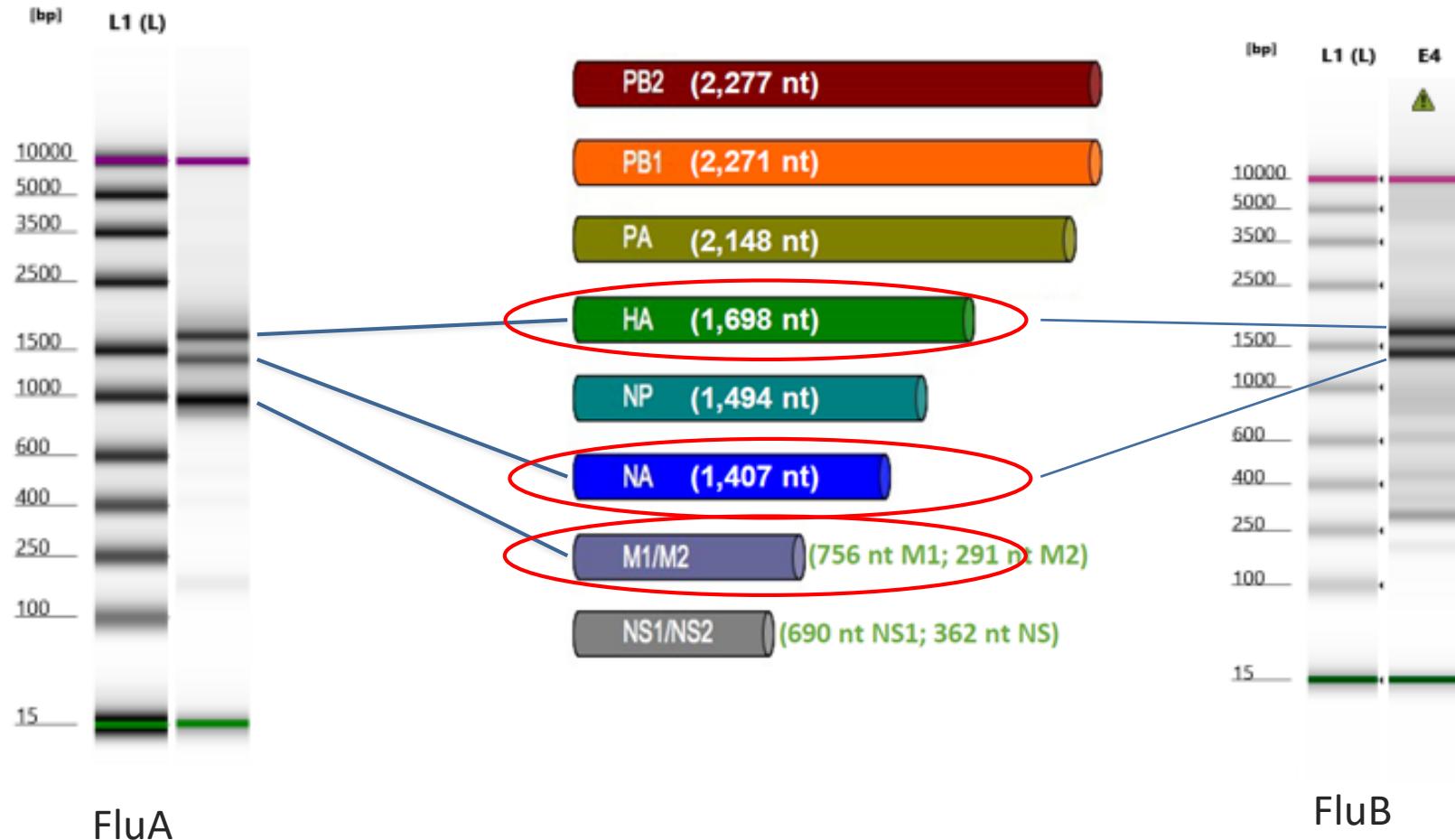
Influenza mRT-PCR for Whole Genome Sequencing (WGS)



- multisegments co-amplification using FluA or FluB universal primers



Influenza A/B mRT-PCR for partial gene NGS



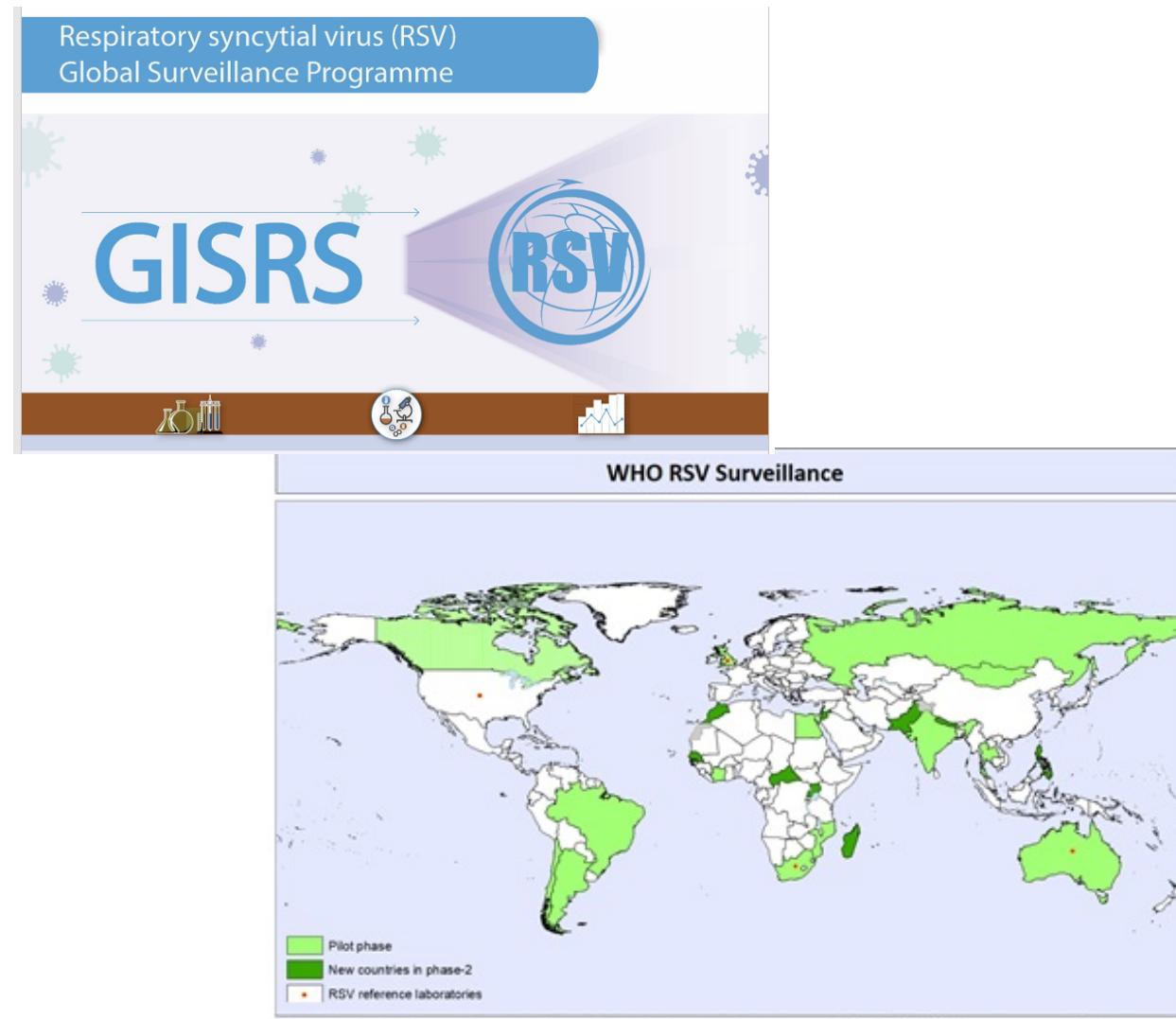
Amplify HA/NA/MP of FluA and HA/NA of FluB in a single tube using primer pools.

Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

WHO RSV Surveillance program



Piloted a RSV surveillance strategy Phase I based on the Global Influenza Surveillance and Response System (GISRS) between 2016 and 2018 in 14 countries.

Launched phase II of the global RSV surveillance between 2018 and 2021 in 11 more countries.

We are one of the 4 WHO Global RSV Surveillance Reference Centre

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement. [1] All references to Kosovo in the document should be understood to be in the context of United Nations Security Council resolution 1244 (1999).

Data Source: GISRS EZCollab Survey
Map Production: WHO Global Influenza Programme
World Health Organization



© WHO 2021. All rights reserved.

Respiratory Syncytial Virus (RSV)

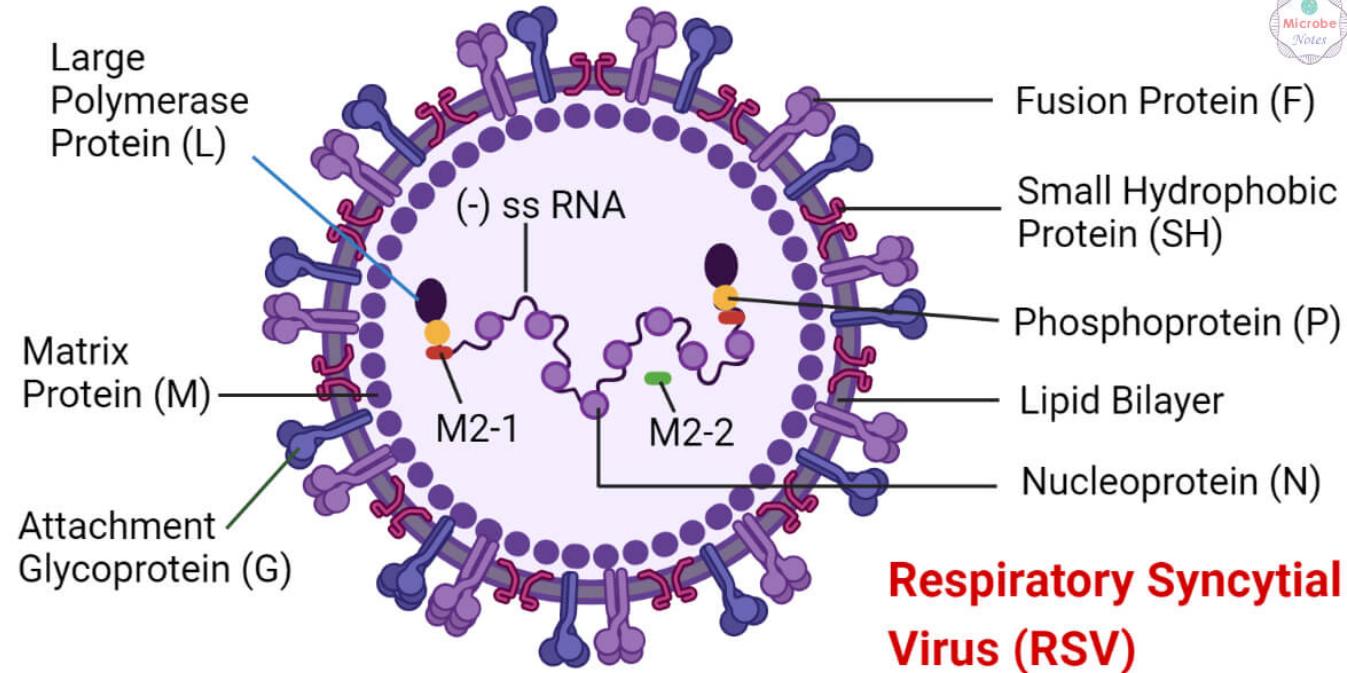


Image from <https://microbenotes.com/respiratory-syncytial-virus-rsv/>

- Two envelope proteins responsible for attachment and fusion of the viral particle: F and G
- F & G are also targets for vaccine and mono-clonal antibody treatment

Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

RSV A2 (15.2 kb)

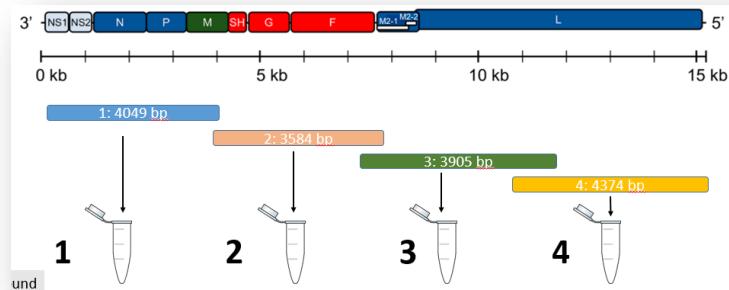


PLOS ONE

RESEARCH ARTICLE

Sequencing and Analysis of Globally Obtained Human Respiratory Syncytial Virus A and B Genomes

Michael E. Bose¹, Jie He¹, Susmita Srivastava², Martha I. Nelson³, Jayati Bera^{2*}, Rebecca A. Halpin², Christopher D. Town², Hernan A. Lorenzi², Daniel E. Noyola⁴, Valeria Falcone⁵, Giuseppe Germa⁶, Hans De Beenhouwer⁷, Cristina Videla⁸, Tuckweng Kok⁹, Marietjie Venter^{10ab}, John V. Williams¹¹, Kelly J. Henrickson^{1*}



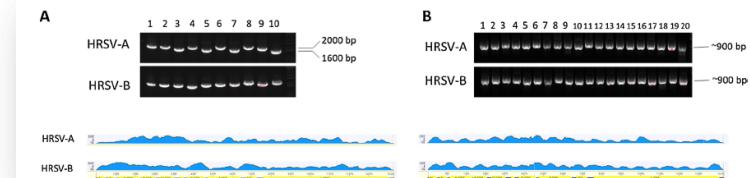
Contents lists available at ScienceDirect

Journal of Virological Methods

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

Next-generation sequencing of human respiratory syncytial virus subgroups A and B genomes

Lijuan Wang ^a, Terry Fei Fan Ng ^a, Christina J. Castro ^b, Rachel L. Marine ^a, Laura C. Magaña ^{b,1}, Mathew Esona ^a, Teresa C.T. Peret ^{a,1}, Natalie J. Thornburg ^a



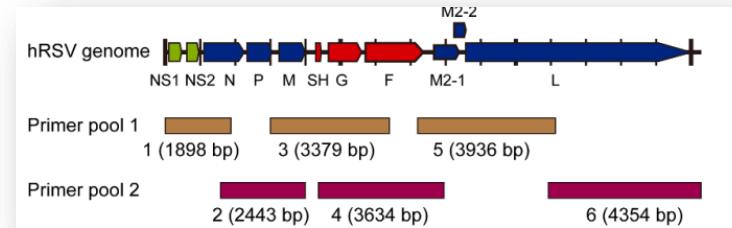
Contents lists available at ScienceDirect

Journal of Clinical Virology

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

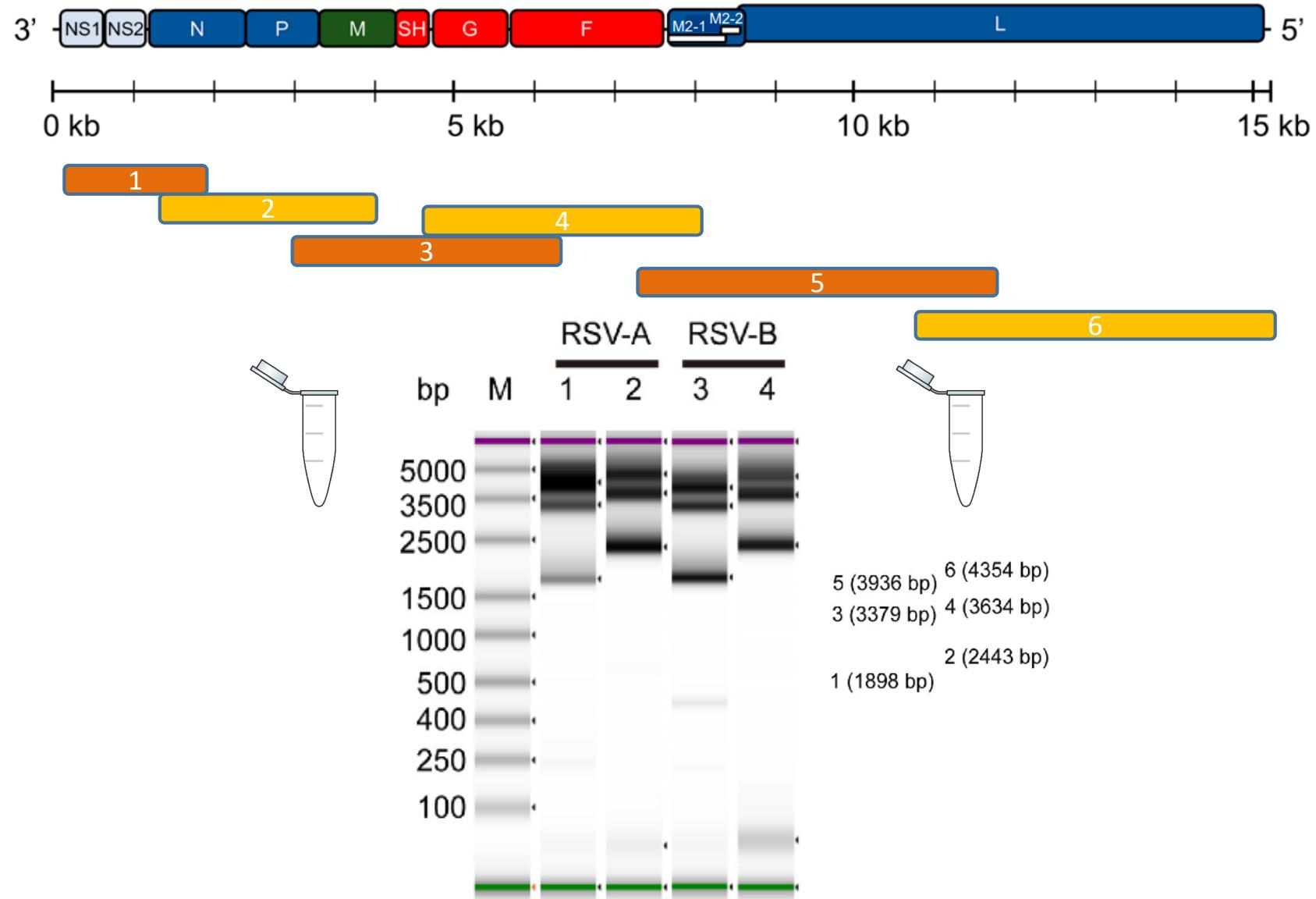
A simplified, amplicon-based method for whole genome sequencing of human respiratory syncytial viruses

Xiaomin Dong ^{a,b}, Yi-Mo Deng ^{a,b}, Ammar Aziz ^{a,b}, Paul Whitney ^{a,b}, Julia Clark ^{c,d}, Patrick Harris ^{e,f}, Catherine Bautista ^f, Anna-Maria Costa ^g, Gregory Waller ^g, Andrew J. Daley ^h, Megan Wieringa ⁱ, Tony Korman ⁱ, Ian G. Barr ^{a,b,*}



Two-tube mRT-PCR for RSV

RSV A2 (15.2 kb)



Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

Challenges

- Which chemistry/platform to choose
- Original specimens
 - Host gene/other pathogens background
 - Enrichment of specific target
- Defective interfering (DI) virus particles
 - Partial sequences for certain genes
- Bottleneck: data analysis (time-consuming, often requires bioinformatics expertise)
- Common problems for data analysis:
 - gaps, uneven coverage, deletion/insertion detection, reference selection
 - De novo analysis for unknown influenza
- A pipeline is useful, but is not the answer for everything

Outline



- WHO GIP eGISRS
- WHO CC Melbourne
- NGS technologies and platforms
- Influenza viruses
- Influenza virus NGS
- RSV
- RSV WGS
- Challenges
- WHO Guidance for NGS

WHO Guidance on NGS



- Objective:
 - To provide practical information and guidance to
 - Evaluate and choose sequencing technologies – Sanger/NGS
- Target Audience:
 - NICs who are considering to implement NGS technology
- Goal:
 - For genetic characterisation and surveillance of influenza viruses

Main considerations

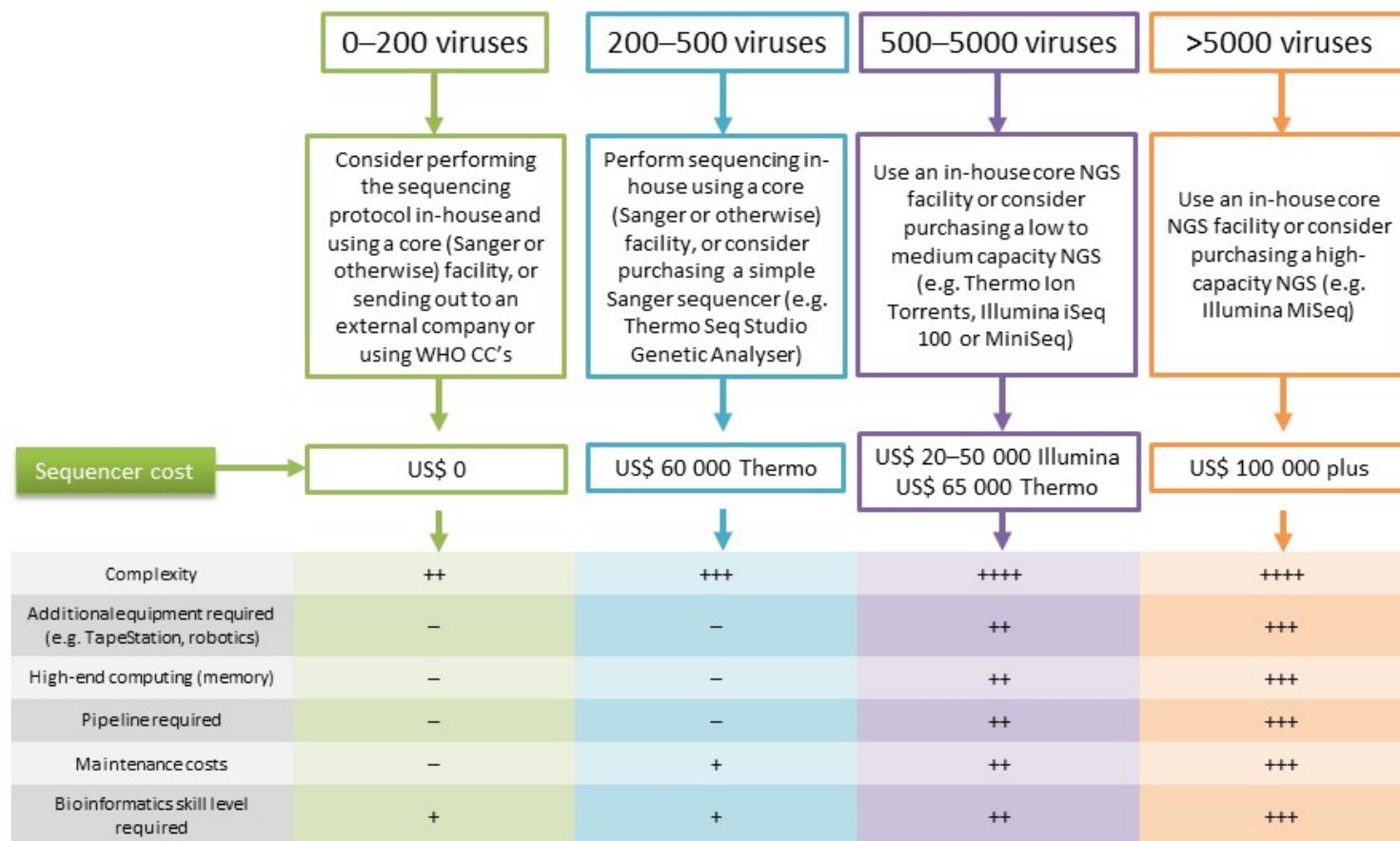
- Sample size and collection pattern
- Budget
 - Initial instrument cost
 - Ongoing annual maintenance cost
 - Reagent/kit cost
 - Other equipment cost (robotic, extra data storage capacity)
- Local technical support, stable reagent supply
- Bioinformatics expertise/support
- Reagent storage
- Resources (computational, stable power supply)

Sample size vs sequencing platform



Selecting the right sequencing approach: a guide for NICs

How many viruses does your laboratory currently sequence or would like to sequence, per year?
(Assume HA and NA genes only/virus)



Summary

- NGS offers a powerful tool for new influenza virus discovery and ongoing surveillance in a timely manner
- High throughput for routine sequencing, cheaper than Sanger sequencing if a large number of viruses sequenced simultaneously
- Bioinformatics analysis is challenging, even using a pipeline
- Evaluate and plan well if consider to get an NGS platform

Acknowledgements



The Melbourne WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australian Government Department of Health and Aged Care.