ANALYSIS OF INFLUENZA A GENOMES AND BACTERIAL MICROBIOTA IN UPPER RESPIRATORY TRACTS OF INFLUENZA PATIENTS

Miss Somruthai Rattanaburi

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biomedical Sciences
Inter-Department of Biomedical Sciences
GRADUATE SCHOOL
Chulalongkorn University
Academic Year 2021

การวิเคราะห์จีโนมของเชื้อไวรัสไข้หวัดใหญ่ชนิด เอ และ สังคมแบคทีเรียในระบบทางเดินหายใจส่วนต้นของผู้ป่วยไข้หวัดใหญ่

น.ส.สมฤทัย รัตนบุรี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเวชศาสตร์ (สหสาขาวิชา) สหสาขาวิชาชีวเวชศาสตร์
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2564

Thesis Title	ANALYSIS OF INFLUENZA A GENOMES AND BACTERIAL MICROBIOTA IN UPPER RESPIRATORY TRACTS OF INFLUENZA PATIENTS
By	Miss Somruthai Rattanaburi
Field of Study	Biomedical Sciences
Thesis Advisor	Associate Professor SUNCHAI PAYUNGPORN, Ph.D.

Accepted by the GRADUATE SCHOOL, Chulalongkorn University in Partial Fulfillment of the Requirement for the Doctor of Philosophy

..... Dean of the GRADUATE SCHOOL
(Associate Professor YOOTTHANA CHUPPUNNARAT, Ph.D.)

DISSERTATION COMMITTEE

..... Chairman
(Assistant Professor Chalisa louicharoen Cheepsunthorn, Ph.D.)

..... Thesis Advisor
(Associate Professor SUNCHAI PAYUNGPORN, Ph.D.)

..... Examiner
(Professor Yong Poovorawan, M.D.)

..... Examiner
(Associate Professor AUNYARATANA THONTIRAVONG, D.V.M.,Ph.D.)

..... External Examiner
(Associate Professor Nattanan Panjaworayan T-Thienprasert, Ph.D.)

สมฤทัย รัตนบุรี : การวิเคราะห์จีโนมของเชื้อไวรัสไข้หวัดใหญ่ชนิด เอ และ สังคมแบคทีเรียในระบบทางเดินหายใจส่วนต้นของผู้ป่วยไข้หวัดใหญ่. (ANALYSIS OF INFLUENZA A GENOMES AND BACTERIAL MICROBIOTA IN UPPER RESPIRATORY TRACTS OF INFLUENZA PATIENTS) อ. ที่ปรึกษาหลัก : รศ. ดร.สัญชัย พยุงภร

ไข้หวัดใหญ่ตามฤดูกาลเป็นโรคติดต่อทางระบบทางเดินหายใจที่เกิดจากไวรัสไข้หวัดใหญ่เข้าติดเชื้อในระบบทางเดินหายใจส่วนต้นของมนุษย์ ซึ่งก่อให้เกิดความเสียหายทางด้านสาธารณะสุขทั่วโลก นอกจากนี้ระบบทางเดินหายใจส่วนบนยังเป็นที่อยู่ของสังคมจุลชีพที่หลากหลาย ที่มีบทบาทสำคัญในการป้องกันสิ่งแปลกปลอม รวมถึงอาจส่งผลทางตรง หรือทางอ้อมต่อการติดเชื้อไวรัสได้ ดังนั้น วัตถุประสงค์ของการศึกษาแรกจึงมุ่งเน้นไปที่การจำแนกจีโนมเชื้อไวรัสไข้หวัดใหญ่ชนิด เอ และ วิเคราะห์การกลายพันธุ์โดยใช้เทคโนโลยี NGS ผลการวิเคราะห์วงศ์วานวิวัฒนาการ (phylogenetic) พบว่า วัคซีนไข้หวัดใหญ่สายพันธุ์ เอ (H1N1) ที่องค์การอนามัยโลกแนะนำ อาจมีประสิทธิภาพในการป้องกันได้น้อยกว่า ในขณะที่สายพันธุ์ เอ (H3N2) อาจมีประสิทธิภาพในการป้องกันเชื้อสายพันธุ์ เอ (H3N2) ที่ระบาดในประเทศไทยช่วงปีพ.ศ. 2560-2561 ได้ดีกว่า นอกจากนี้พบการกลายพันธุ์แบบเปลี่ยนกรดอะมิโน (nonsynonymous mutation) ทั้ง 8 ยีนของเชื้อไวรัสไข้หวัดใหญ่ทั้ง 2 สายพันธุ์ โดยเฉพาะในยีน *HA* และ *NA* ในการวิเคราะห์ความหลากหลายทางพันธุกรรม (nucleotide diversity) พบการคัดเลือกเชิงลบ (negative selection) ที่ยีน *PB1 PA HA* และ *NA* ในไวรัสไข้หวัดใหญ่สายพันธุ์ เอ (H1N1) สำหรับวัตถุประสงค์ของการศึกษาที่สองคือ เปรียบเทียบข้อมูลสังคมแบคทีเรียในระบบทางเดินหายใจส่วนต้นระหว่างผู้ป่วยไข้หวัดใหญ่ (ชนิดเอ และบี) กลุ่มติดโควิด 19 และกลุ่มที่ไม่ได้ติดทั้งไข้หวัดใหญ่รวมทั้งโควิด-19 ด้วยวิธี 16S rDNA sequencing โดยการวิเคราะห์ Shannon diversity พบว่าในกลุ่มติดไข้หวัดใหญ่มีความหลากหลายของสังคมแบคทีเรียอย่างมีนัยสำคัญเมื่อเทียบกับกลุ่มที่ไม่ได้ติดทั้งไข้หวัดใหญ่รวมทั้งโควิด-19 และ Beta diversity พบว่าองค์ประกอบของสังคมแบคทีเรียมีความแตกต่างกันอย่างมีนัยสำคัญทั้ง 4 กลุ่ม นอกจากนี้ผลปริมาณสัมพัทธ์ (relative abundance) แสดงให้เห็นว่าแบคทีเรียวงศ์ Enterobacteriaceae เพิ่มขึ้นในกลุ่มติดไข้หวัดใหญ่ ขณะที่แบคทีเรียสกุล *Streptococcus, Prevotella, Veillonella* และ *Fusobacterium* เพิ่มขึ้นในกลุ่มที่ไม่ได้ติดทั้งไข้หวัดใหญ่รวมทั้งโควิด-19 โดยจากการศึกษาครั้งนี้ทำให้ได้องค์ความรู้พื้นฐานเพื่อนำไปหาความปฏิสัมพันธ์ระหว่าง เจ้าบ้านและจุลินทรีย์ซึ่งอาจเป็นประโยชน์ในการทำนายสภาวะสุขภาพและสามารถนำไปประยุกต์ใช้กับวิศวกรรมไมโครไบโอมเพื่อเพิ่มภูมิคุ้มกันต่อการติดเชื้อที่ก่อโรคในอนาคตได้

สาขาวิชา	ชีวเวชศาสตร์ (สหสาขาวิชา)	ลายมือชื่อนิสิต .....
ปีการศึกษา	2564	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

5987805520 : MAJOR BIOMEDICAL SCIENCES

KEYWORD: Influenza A virus, Genomes, Mutations, Bacterial microbiota, Upper respiratory tract, 16S rDNA, High-throughput sequencing

Somruthai Rattanaburi : ANALYSIS OF INFLUENZA A GENOMES AND BACTERIAL MICROBIOTA IN UPPER RESPIRATORY TRACTS OF INFLUENZA PATIENTS. Advisor: Assoc. Prof. SUNCHAI PAYUNGPORN, Ph.D.

Seasonal influenza is a contagious respiratory illness caused by influenza viruses that infect the upper respiratory tract (URT) of humans and represent a major burden for public health. Furthermore, the URT is colonized by a diverse microbial community which is a key factor protective from pathogens and may directly or indirectly affect viral infection. Therefore, the first aim of this study focuses on influenza A genome characterization and mutation analysis based on NGS technology. The phylogenetic analysis based on the deduced amino acid sequences revealed that the recommended vaccine (A/H1N1) strain might be less effective, whereas the recommended vaccine (A/H3N2) was more effective against the circulating influenza viruses in Thailand during 2017-2018. In addition, several nonsynonymous mutations occurred across eight segmented genes of both viruses, particularly in *HA* and *NA* genes. Indeed, nucleotide diversity analysis was observed negative selection in the *PB1*, *PA*, *HA*, and *NA* genes of A/H1N1 viruses. Then, the second aim of this study is to compare the bacterial microbiota profile in the URT with influenza (Flu A or Flu B groups) and non-influenza (COVID-19 or Non-Flu & COVID-19) patients by 16S rDNA sequencing. The Shannon diversity for the influenza group was significantly lower than Non-Flu & COVID-19 group. The beta diversity revealed that microbial compositions were significantly different among groups. The relative abundance showed that the family of Enterobacteriaceae was increased the in influenza group, whereas *Streptococcus*, *Prevotella*, *Veillonella*, and *Fusobacterium* were predominated in Non-Flu & COVID-19. In summary, our data provide fundamental knowledge to investigate the association host-microbe interaction that might be useful for predicting health status and applied for microbiome engineering to enhance immunity system to future infections.

Field of Study: Biomedical Sciences Student's Signature .....

Academic Year: 2021 Advisor's Signature .....

ACKNOWLEDGEMENTS

While a completed dissertation only has a single name of students, the long process of completing a dissertation is always achieved with the dedicated work of numerous people during this work. So, I would like to thank all the people who contributed in some way to complete my success with my deepest heart.

I would like to express my sincere thanks to my advisor, Associate Professor Dr. Sunchai Payungporn, for his invaluable help and opportunity to gain experience in the research field of molecular virology. I thank my past and present labmates of the Research Unit of Systems Microbiology for all supporting, sharing their knowledge, and helping each other with problem-solving. Especially, I would like to thank Mr. Vorthon Sawaswong, who is an important person in part of my data analysis.

I would like to express gratitude to the Center of Excellence in Clinical Virology (Professor Yong Poovorawan, Dr. Nungruthai Suntronwong, Miss Preeyaporn Vichaiwattan, and Mr. Sumeth Korkong), the staff from Bangpakok 9 International Hospital, Chum Phae Hospital, and Institute IUDC, Thailand for providing the sample collection and performing influenza diagnosis. Moreover, I would like to thank Assistant Professor Aunyaratana Thontiravong and the department of Virology, Faculty of Veterinary Medicine Chulalongkorn University, for support in influenza virus isolation.

A special debt of gratitude is expressed to the Royal Golden Jubilee (RGJ) Ph.D. Programme scholarship (PHD/0150/2558), National Research Council of Thailand (NRCT) [2564NRCT321520], and Thailand Research Fund (TRF)

Lastly, I would like to thank my family, my friends (Miss Pattaraporn Nimsamer and Miss Orapan Mayuramart), Miss Suwalak Chitcharoen, and Mr. Pavaret Sivapornnukul for supporting me spiritually throughout my study advice in every step I have taken in my life.

Somruthai Rattanaburi

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

%	Percent
℃	Degree Celsius
µg	Microgram
µL	Microliter
µm	Micrometre
µM	Micromolar
16S rDNA	16S ribosomal deoxyribonucleic acid
16S rRNA	16S ribosomal ribonucleic acid
AMPs	Antiviral-antimicrobial peptides
ASVs	Amplicon Sequence Variants
bp	Base pairs
CCD	Coupled-charge device
CDC	Centres for Disease Control and Prevention
cDNA	Complementary deoxyribonucleic acid
CK	Chicken kidney
CO_2	Carbon dioxide
COVID-19	Coronavirus disease of 2019
CPE	Cytopathic effect
Ct	Cycle threshold values
DMEM	Dulbecco's Modified Eagle Medium
dNTPs	Deoxynucleotide triphosphates
dsDNA	Double-stranded deoxyribonucleic acid
FBS	Fetal bovine serum
Flu	Influenza disease
GI	Gastrointestinal tract
HA	Hemagglutinin
HMP	Human Microbiome Project
IgA	Immunoglobulin A
ILI	Influenza-like illness

IUDC	Institute for Urban Disease Control and Prevention
LEfSe	Linear discriminant analysis Effect Size
LRT	Lower respiratory tract
M	Matrix
MDCK	Madin-Darby canine kidney cell
$MgCl_2$	Magnesium Chloride
min	Minute
mL	Milliliter
mM	Millimolar
NA	Neuraminidase
NEP	Nuclear export protein
NGS	Next-generation sequencing
NK	Natural killer cell
nM	Nanomolar
No.	Number
NP	Nasopharyngeal swab
NP	Nucleoprotein
NS	Non-structural protein
M	Matrix
OTUs	Operational Taxonomic Units
P	Passages
PA	Polymerase acidic
PB1	Polymerase basic 1
PB2	Polymerase basic 2
PBS	Phosphate buffer saline
PCoA	Principal-coordinate analysis
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
RdRp	RNA dependent RNA polymerase
RNP	Ribonucleoprotein
RT	Reverse transcriptase

sec	Second
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SBS	Sequencing by synthesis
SRA	Sequence Read Archive
TPCK-trypsin	Tosyl phenylalanine chloromethyl ketone- trypsin
U	Unit
URT	Upper respiratory tract
UV	Ultraviolet
V	Hypervariable regions
v/v	Volume by volume
vRNA	Viral ribonucleic acid
VTM	Viral transport media
WGS	Whole-genome sequencing
WHO	World Health Organization
πN	Nonsynonymous nucleotide variation
πS	Synonymous nucleotide variation

Chapter 1

INTRODUCTION

Annually, the widespread circulation of seasonal influenza viruses has been observed in humans and caused upper respiratory tract (URT) disease. They are comprised of two subtypes of influenza A viruses (A/H1N1 and A/H3N2) and two lineages of influenza B virus (B/Victoria and B/Yamagata lineages). These viruses cause significant morbidity and mortality each year and contribute to worldwide public health problems. Especially, the children and elderly have a high risk of influenza infection. The study in Thailand during 2005-2008 reported that an annual average of 36,400 influenza-associated hospitalizations and 300 deaths occurred (1). Therefore, the influenza vaccine has an important role in preventing influenza infection and reducing the severity.

Influenza A virus is a negative single-strand RNA virus in the Family of *Orthomyxoviridae* with a 13.5-kb genome consisting of eight segments (ranging from 890-2341 nucleotides in length) encoding 18 proteins. Indeed, the influenza A virus has diverse subtypes and infects many hosts such as humans, aquatic birds and swine. A remarkable feature of the influenza A virus is that it can cause more severe illnesses because they evolve with a high mutation rate and contribute to antigenic variation through antigenic drift and antigenic shift mechanisms (2). Antigenic drift is caused by the continuous accumulation of point mutations in the antigenicity of the viral surface glycoproteins and selection pressure from host immunity. Contrarily, an antigenic shift is a major change by genetic reassortment within eight segments from two strains infected in the same call, resulting in novel strains (3). The previous study report that each replicated genome of influenza A carries an average of 2–3 mutations per genome (4). Such constantly evolving allows the viral variation to escape host immune response and vaccination. These have been to develop the composition of influenza vaccine strain and reformulated in every year, preventing influenza infection. Nowadays, the influenza vaccine is available for the seasonal trivalent influenza vaccine (TIV) and quadrivalent influenza vaccine (QIV), which are

derived from two influenzas A train and one (in TIV) or two (in QIV) influenza B strain. Due to rapid evolution of influenza viruses, the global influenza surveillance was performed to select the best match vaccine strain against circulating strains and provided the influenza vaccine production for temporal regions, including the northern and southern hemispheres. For example, the WHO recommends influenza vaccine compositions for the southern hemisphere are A/Michigan/45/2015 (H1N1), A/Hong Kong/4801/2017 (H3N2), B/Brisbane/60/2008 (Victoria), and B/Phuket/3073/2013 (Yamagata) in 2017. However, the significant position of amino acid change contributes to enhancing viral replication, transmission, virulence and drug resistance. Due to this virus continually evolving, it is essential to surveillance circulating influenza infection in genetic and mutation. Nowadays, Next-generation sequencing (NGS) has been potential for characterization of the viral whole-genome and mutations that provide comprehensive information on the influenza genome.

Human bacterial microbiota in the upper respiratory tract is a commensal bacterial community that plays a crucial role in protecting the mucosal surface from pathogens. The diversity of the bacterial community is associated with exposure to external triggers, including environmental factors (pollution, smoke and allergy), the host factors (age, diet, lifestyle, and antibiotic usage) and the microbe-host immune response. Alterations of the microbiota from respiratory viral infections contribute to a higher susceptibility to secondary infections and disease severities. A healthy respiratory status is associated with an abundance of *Prevotella*, *Veillonella*, *Streptococcus*, *Staphylococcus*, and *Moraxella*. In contrast, bacterial dysbiosis of respiratory virus infections results in a lower bacterial diversity and a higher abundance of specific bacteria such as *Streptococcus*, *Pseudomonas*, and *Neisseria* (5). Specifically, the relative abundances of *Staphylococcus* and *Bacteroides* genera were changed when the influenza A viruses infection (6). Comparing the bacterial profiles in Thai patients with influenza and non-influenza could reduce the confounding factors and incompatible techniques used in each previous study. 16S rDNA sequencing by NGS technologies allows for the taxonomic assignment of microbial species in complex microbial communities.

Research questions

1. What are the genomes and mutations of influenza A viruses that outbreak in Thailand during 2017-2018?

2. What is the bacterial community in upper respiratory tracts of influenza patients?

3. How different are the bacterial microbiota in upper respiratory tracts between patients infected with influenza and non-influenza?

Objectives

1. To characterize the genomes and mutations of influenza A viruses that outbreak in Thailand during 2017-2018 based on NGS technology.

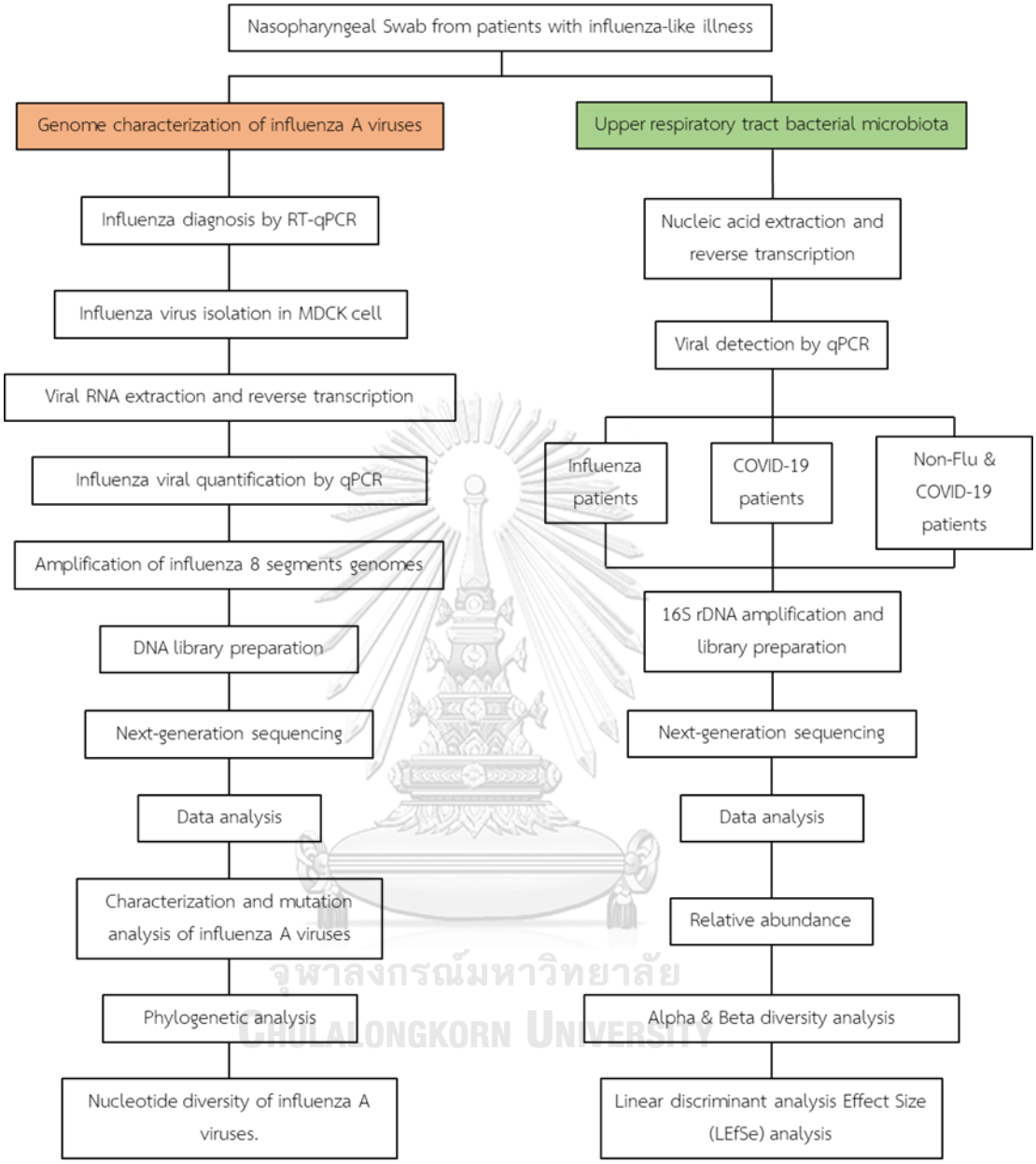
2. To investigate the bacterial community in upper respiratory tracts of influenza patients.

3. To compare the bacterial microbiota between patients infected with influenza and non-influenza.

Keywords

Influenza A virus, Genome characterization, Thailand, Mutation, Bacterial microbiota, Upper respiratory tract, 16S rDNA, High-throughput sequencing

Experimental design



Expected benefits of the study

The results of this study will provide us with a better understanding of the characteristics and the mutation of the influenza A virus that is currently outbreak in Thailand, which would be crucial for preparation against pandemic and epidemic outbreaks in the future. Moreover, the findings of bacterial profiles in the URT of Thai patients with influenza will provide the basic knowledge for further investigation of the relationship between upper respiratory microbiota and respiratory diseases.

Chapter 2

REVIEW OF RELATED LITERATURE

Epidemiology of Influenza viruses

Influenza viruses are identified as acute respiratory pathogens demonstrating clinical signs from asymptomatic to death, especially in persons with chronic condition diseases, pregnant women, infants, and the elderly (7). After the influenza infection, the incubation period of this virus estimates at 2-8 days. The patients develop the symptoms of fever, muscle pain, severe headache, coughing, sore throat, weakness, and fatigue. For severe cases, influenza infection can cause heart or lung disease, immunological disorders, renal failure and diabetes, which may become potentially lethal (8). The seasonal influenza virus causes an infected rate of 5-10% in adults and 20-30% in children (9).

Annually recurrent epidemic outbreaks of influenza A and B viruses are seasonal influenza (10). Influenza A viruses can infect many hosts, including avian and mammalian species (11). Due to the wild bird migration, influenza A viruses can spread worldwide. Influenza B viruses can infect humans, horses, seals, and swine (12, 13). However, the genome of the influenza A virus can rapidly accumulate mutations due to the RNA polymerase virus and the transfer of genomic segments between different influenza virus strains (14). The low fidelity of the influenza virus RNA polymerase contributes to the high rate of replication errors, which occur at approximately 1 in 104 bases per replication cycle (15). Thus, each round of replication leads to a population with more variants (16). These changes of HA and NA can modify the virus that contributes to antiviral drug resistance, evading the host immunity and crossing the barriers of the host species (17). The new influenza virus strains enter the human population from other species due to the lack of pre-existing immunity. Thus, influenza A viruses cause pandemics while influenza B viruses cause only seasonal epidemics (3).

Other subtypes of influenza A viruses from animals have occasionally infected humans. However, these strains are not caused widespread outbreaks due to the limitation of human-human transmissions such as swine A/H3N2, avian A/H5N1, and

avian A/H7N9 (18, 19). In 1918, the first cases of deadly influenza infection reported in Spain caused the deaths of 20-50 million people worldwide by a strain of influenza A (A/H1N1), and the origin of this virus came from an avian-descended A/H1N1 virus (20, 21). After that, the 1957 Asian flu was first identified in China; the A/H2N2 virus is thought out to have emerged after human A (A/H2N2) combined with a mutant strain in ducks. This virus killed 2 million people (22). In 1968, the first pandemic of influenza A (A/H3N2) was also known as Hong Kong flu; this strain caused around 1 million deaths, particularly the elderly aged more than 65 years old (23, 24). The recent influenza pandemic was "swine flu" first isolated in Mexico in 2009. A novel strain generated from the triple reassortment of mixed segmented genes among birds, swine, and humans. So far, seasonal influenza A (A/H1N1), which caused the seasonal epidemic, has been replaced by pandemic influenza A (A/H1N1) 2009 (25-27). During the last century, the A/H1N1, A/H3N2 and A/H2N2 subtypes of influenza A viruses have circulated in humans. The incidence of influenza B virus infections varies between influenza seasons. The Victoria lineage was most predominated in the 1980s, while Yamagata lineage took over in the 1990s. Nevertheless, there have been circulated annually since 2001, but they have only one lineage, predominantly circulating strain each year (28-30). To date, the seasonal influenza epidemic has been caused by influenza A/H1N1, A/H3N2, B/Victoria and B/Yamagata.

Molecular virology of influenza viruses

Classification

Influenza A and B viruses are members of the *Orthomyxoviridae* family. They have segmented negative single strands RNA genome. Depending on genetically and antigenically, the influenza viruses can be classified into four types: A, B, C, and D viruses (Figure 1). Influenza A and B viruses have eight segmented genomes, while influenza C and D viruses have seven-segmented genomes. Different subtypes of influenza A viruses are classified by the antigenic variation of surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). To date, there are 18 HA and 11 NA subtypes (19, 31, 32). Newly, H17-18 and N10-11 have been discovered from the bat (19). However, a few subtypes of seasonal human influenza viruses comprise H1, H2,

H3, N1 and N2 that cause the epidemic outbreaks. On the other hand, influenza B viruses can be divided into two lineages depending on the genetic and antigenically difference of HA surface glycoprotein.

Structure and genome

Influenza virus particles have a size of approximately 80-120 nm in diameter. These viruses consist of the host cell-derived lipid membrane, two types of surface glycoprotein: HA and NA, and ion channel protein (M2). The M2 is anchored and responsible for the viral uncoating process in the infected cell. The HA is a homotrimer whose receptor binding and membrane fusion function. In contrast, the NA is a homotetramer, responding to destroying receptors by hydrolyzing sialic acid groups and releasing the viral progeny (33). The M1 protein is responsible for viral capsid where the ribonucleoprotein (RNP) complex adheres. Inside the virion, the influenza A and B viruses genome is approximately 13.5 kb, and the size of each RNA segment varies from 890 to 2,341 nucleotides (33).

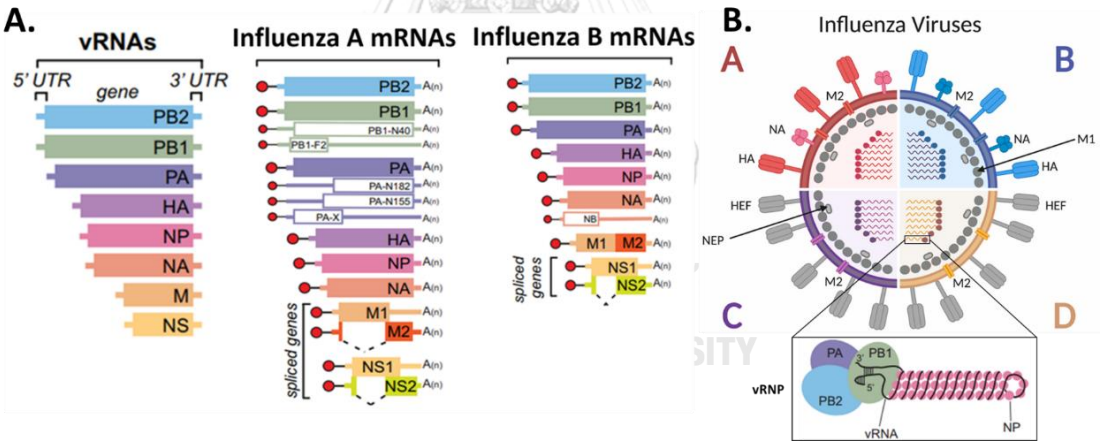


Figure 1. Schematic representation of the influenza virus genome and structure .

(A.) The viral RNA segmented genome of the influenza A and B viruses

(B.) Structure of influenza virus particle (34, 35).

The viral genome consists of polymerase basic 2 (*PB2*), polymerase basic 1 (*PB1*), polymerase acidic (*PA*), hemagglutinin (*HA*), nucleoprotein (*NP*), neuraminidase (*NA*), matrix (*M*), and a non-structural protein (*NS*) (Figure 1.) (36). All eight segments of the genomic RNA (vRNA) are bound to the NP and three viral RNA polymerase

subunits (*PB2*, *PB1*, and *PA*), resulting in the formation of the RNP complex. The low amount of NS2 protein functions as a nuclear export protein for vRNA in infected cells (37). The viral genome of influenza A can encode 17 proteins, while influenza B can translate up to 11 proteins (38). The function of proteins encoded by each viral RNA segment is described in Table 1.

Table 1. Characterization of influenza viruses and their functions (39).

RNA segment	Influenza A Protein	Influenza B Protein	Functions
1	PB2	PB2	Component of RNA polymerase; cap recognition
2	PB1	PB1	Catalytical subunit of RNA polymerase: elongation
	PB1-F2	-	Pro-apoptotic activity; IFN antagonist, interaction with PB1 to regulate polymerase activity
	PB1-N40	-	Viral life cycle
3	PA	PA	RNA polymerase; Cap-snatching endonuclease subunit
	PA-X	-	Modulates host gene expression, negative virulence regulator
	PA-N155	-	N-terminally truncated version PA; function unknown
	PA-N182	-	N-terminally truncated version PA; function unknown
4	HA	HA	Receptor binding site, fusion activity; antigenic determinant, assembly and binding
5	NP	NP	RNA binding RNA synthesis, RNP nuclear import
6	NA	NA	Neuraminidase activity, release novel particles after budding; antigenic determinant
		NB	Ion channel activity; function unknown
7	M1	M1	Interaction with RNP and glycoproteins, RNP nuclear export, viral assembly and budding
	M2	BM2	Membrane protein ion channel activity; essential for uncoating and role viral budding
	M42	-	M2 isoform; ion channel activity
8	NS1	NS1	Multifunctional protein involved in virus-host interactions, regulation viral RNA polymerase complex
	NS2/NEP	NS2/NEP	RNP nuclear export regulation of RNA synthesis (transcription and replication)
	NS3	-	NS1 isoform with an internal deletion

The life cycle of influenza viruses

The influenza infection begins with the attachment between the viral receptor (HA protein) and sialyl oligosaccharide receptors that are present on the host cell surface. In the receptor binding process, this virus can infect by recognizing two types of sialic acid such as α 2,6 sialic acid, which is found in the epithelial cell of the human upper respiratory tract, and α 2,3 sialic acid, which is found in the epithelial cell of the avian intestine and the epithelial cell of the human lower respiratory tract (40). Then, the virus is taken up into the host cell by receptor-mediated endocytosis (41). In the fusion and uncoating, late endosomal vesicles with low pH can induce virion to change the conformational structure of HA. HA protein is hydrolyzed by protease enzyme, resulting in the dimerization of HA1 and HA2 linked by the disulfide bond. HA2 plays a role in mediating fusion with the host endosomal membrane (42). Subsequently, the viral M2 ion channel is opened, resulting in the influx of proton ions into the viral particle and inducing eight segmented viral RNP complexes releasing into the cytoplasm. The RNP complexes are imported to the host nucleus, where the genome of the influenza virus is transcribed and replicated by viral RdRp (43). After these processes, viral mRNAs are transcribed and then translated using host cell machinery in the cytoplasm. In the mean times, the replication is also triggered to create positive-sense RNA, which is used as the intermediate template to generate the viral RNA. Viral proteins are processed to the post-translation process, such as glycosylation and the viral packaging (44). After viral assembly, the virus is cleaved by the NA protein to release the virus out of the infected cell (Figure 2.).

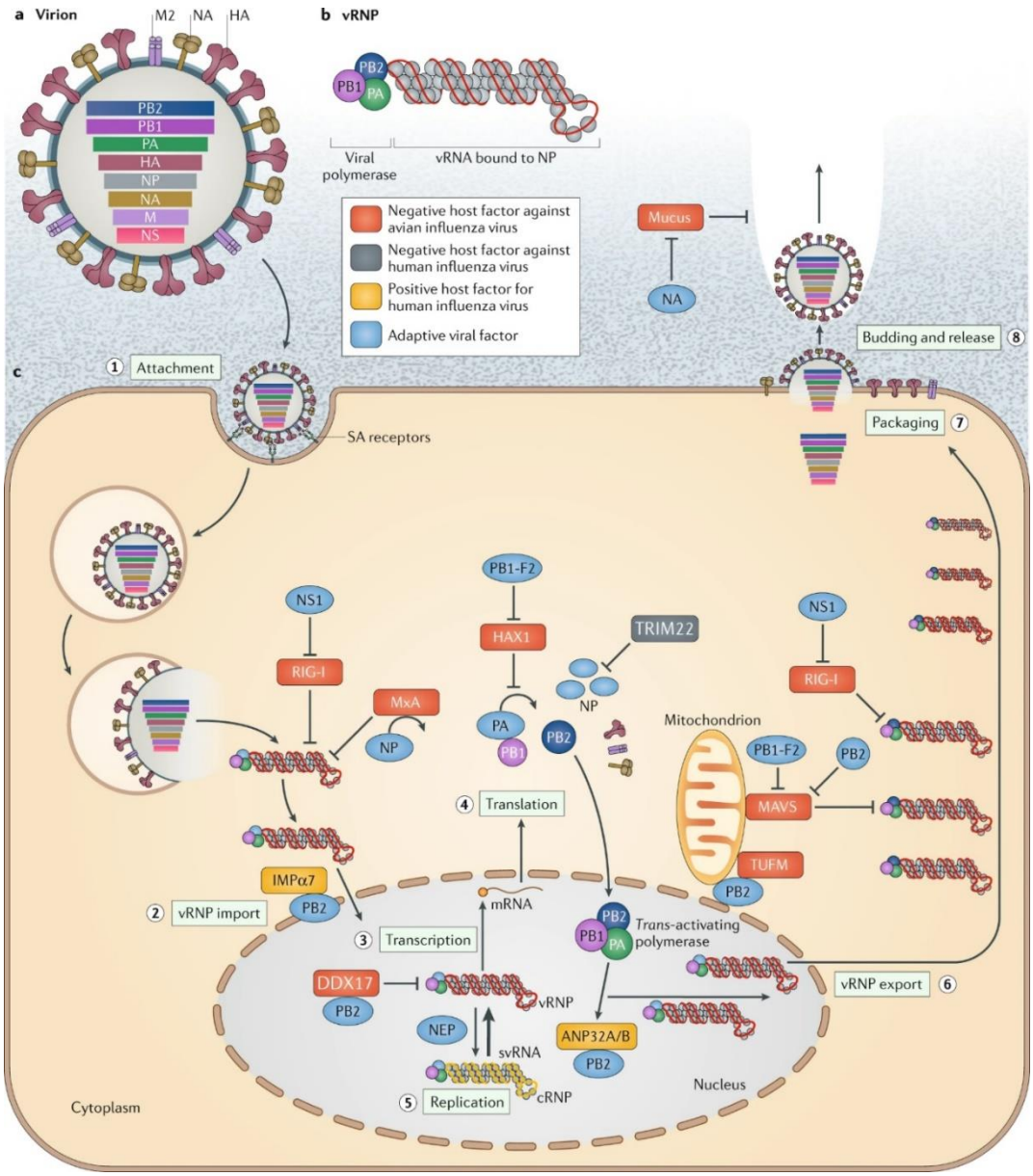


Figure 2. The life cycle of influenza viruses. The life cycle of the influenza virus consists of several processes; viral attachment, viral entry by endocytosis, fusion and uncoating, genome replication and transcription, translation, packaging and budding from the infected cell (45).

Microbiome and influenza infection

Microbiome

A diverse microbial community colonizes on the human body's surface in a defined environment known as the human microbiota. Humans are holobionts; hosts with associated communities of microorganisms. Therefore, microorganisms (bacteria, fungi, viruses, bacteriophages, archaea, and eukaryotes) colonize in a particular environment, and their collective genome defines the term microbiome (46). The microbiota resides in the human body surface, such as the skin, vagina, oral cavity, gut, upper respiratory tract (URT), and lung (47-50). Indeed, the body site looks to be a factor for diversity in these microbiota communities.

In 2007, the National Institutes of Health started the Human Microbiome Project (HMP) to study microbiota communities and their relationship with human health and disease using the metagenomic sequencing methods (51). Human health is associated with the microbiota. Microbiota are necessitated for optimal human development and against various pathogens. For example, the human gastrointestinal (GI) tract microbiota is essential for human physiology, including digestion, vitamin synthesis, development and maintenance of the immune system (52, 53).

Several factors impact the potentially disrupted individual microbiomes divided into environmental factors (pollution, hygiene, demographic change, migration, and geography), host factors (genetics, age, gender, and physiological parameters; such as pH and hormonal fluctuations), exposure factors (diet, lifestyle, medications and infection), and the microbe-host immune response (54, 55). The relationship between bacterial communities, viruses and host physiology is complex. Understanding the pathophysiological consequences; first, either bacterial or viral infection may change the richness and evenness of the microbiota compositions. Second, some microbiome status may increase the host's susceptibility to infection and disease. Third, among the microbiome, environment and pathogens regulate the host immune response, impacting the outcome of diseases. Therefore, balancing beneficial bacteria or adding probiotics offers the microbiota potential to enhance antiviral immunity (Figure 3.).

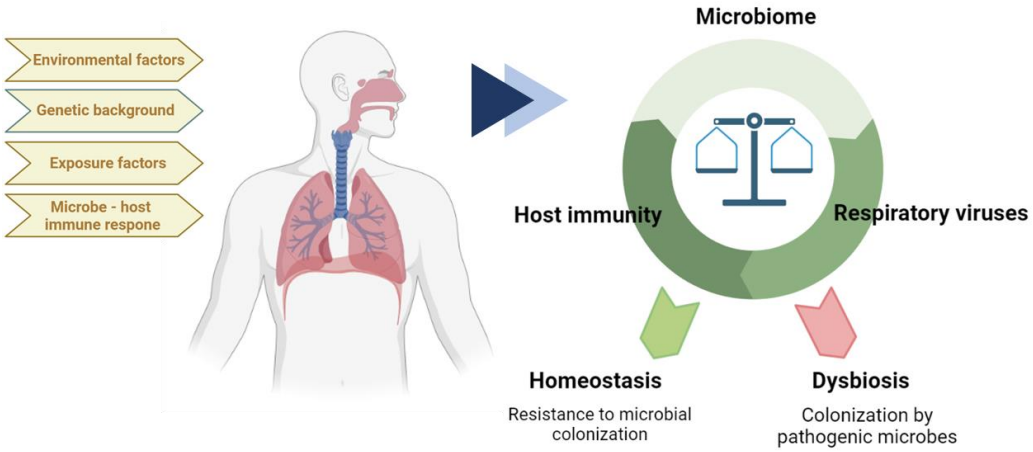


Figure 3. The relationship between bacterial communities, viruses, and hosts (5).

The microbiome and antiviral mechanism

Virus infections still cause morbidity and mortality worldwide. However, previous research revealed that commensal microbiota influenced host defense against viral pathogens (56). Mucosal surfaces (respiratory, GI, or vaginal) are a barrier to defense against viral infection divided into three broad defense lines: the mucus layer, innate immune response, and adaptive immune response. The antiviral microbiome mechanisms are directly and indirectly responsible (Figure 4.), including 1.) Enhanced mucosal barrier function: the mucous layer of wet epithelial surfaces (lung, GI, and vagina) are produced mucins, which allow physical barriers to protect invasive pathogens into epithelial cells (57). The composition of GI microbiota have affected the mucus layer function and mucin production (58). In particular, porcine gastric mucins have been found to protect epithelial cells from several virus infections, such as influenza A virus and human papillomavirus type 16 (57). 2.) Secretion of antiviral-antimicrobial peptides (AMPs), also known as bacteriocins, are antiviral activity (virucidal) and produced by certain bacterial species. For example, duramycin bacteriocins produced by *Streptomycetes* have been observed to prevent the Zika virus before viral entry into a human cell (59). 3.) Inhibition of viral attachment to host epithelial cells: the previous study in the culture model showed that two strains of bacteria, *Lactobacillus* (*sensu lato*) and *Bifidobacterium*, interfere with vesicular stomatitis virus attachment and entry into cells (60). 4.) Modulation of

the immune system: In innate immunity, the bacterial strain *Lactococcus lactis* JCM 5805 activates plasmacytoid dendritic cells to produce interferon-alpha (IFN-α). Then, IFN-α induces the cytotoxic activity of natural killer (NK) cells (61). Furthermore, *L. rhamnosus* GG (ATCC53103) has been shown to enhance the soluble factor Msp2 (or p40) protein production, thus stimulating B-cells into IgA antibodies production (62).

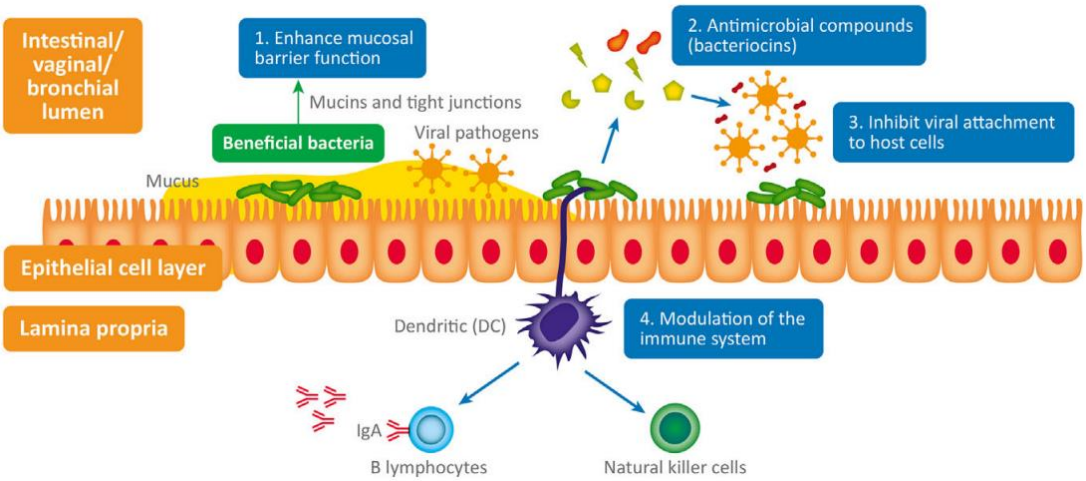


Figure 4. The antiviral microbiome mechanisms: 1. Enhanced mucosal barrier function, 2. Secretion of antiviral-antimicrobial compounds, 3. Inhibition of viral attachment to host cells, and 4. Modulation of the immunity (63).

Bacterial microbiota in the upper respiratory tract

The upper respiratory tract starts from the nostrils to the larynx. The nose and nasopharynx are important niches of the URT, which are commensal bacterial habitats and may cause airway infections by pathogenic species. Bacterial microbiota associated with humans from the first hour of life. Within the first week of healthy URT status, the bacterial profiles comprise *Staphylococcus aureus*, *Corynebacterium spp.*, *Dolosigranulum pigrum*, *Moraxella spp.*, *Streptococcus pneumoniae*, and *Haemophilus influenzae* (64). The nasopharyngeal microbiota of 18 months of age was predominant with bacterial in phyla of *Proteobacteria*, *Firmicutes, Bacteroidetes*, *Actinobacteria*, *Fusobacteria*, and bacterial in genera of *Moraxella*, *Streptococcus*,

Haemophilus, *Flavobacterium*, *Dolosigranulum*, and *Neisseria* (65). The lung microbiome in healthy adults enriched with *Prevotella*, *Veillonella*, *Streptococcus*, *Neisseria*, *Haemophilus*, *Fusobacterium* and *Pseudomonas* (66). Overview of the bacterial community composition in the upper respiratory tract (URT; nasopharynx, and oropharynx) and the lower respiratory tract (LRT; lung) in health and disease as illustrated in Figure 5.

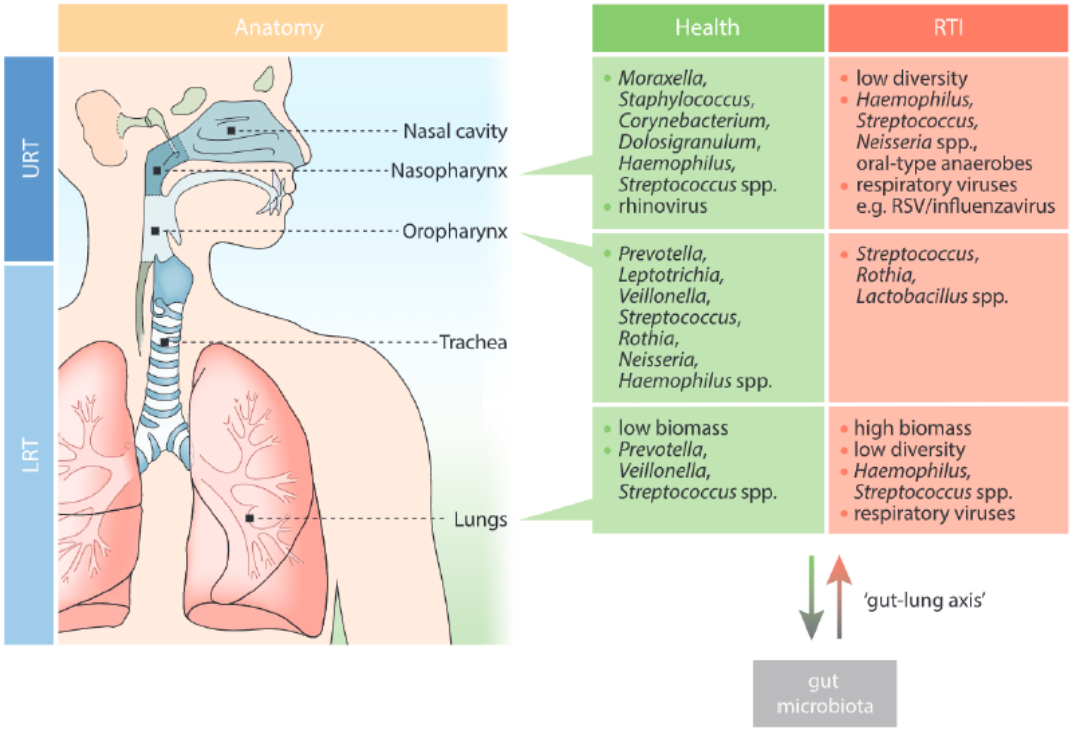


Figure 5. The bacterial composition in the respiratory tract of health and disease(67).

The relationship between gut-lung microbiota

The microbiota in the respiratory system directly influences host immune responses. Alteration of immunity or inflammation can also impact host immune responses. Moreover, microbiota in the GI tract and the respiratory system communicate and affect each other, defined as the gut-lung axis (Figure 6.). A healthy gut, the gut microbiome contributes to maintaining immune response and homeostasis through signaling by derived immunomodulating compounds (lipopolysaccharides, peptidoglycan, and short-chain fatty acids) from these

microbiome. However, infection or medication result in the gut microbiome alteration and dysbiosis, which changes the lung's signaling and antiviral immune responses. Therefore, intake of probiotics may help to improve effective immune response directly in the gut and indirectly in the lung.

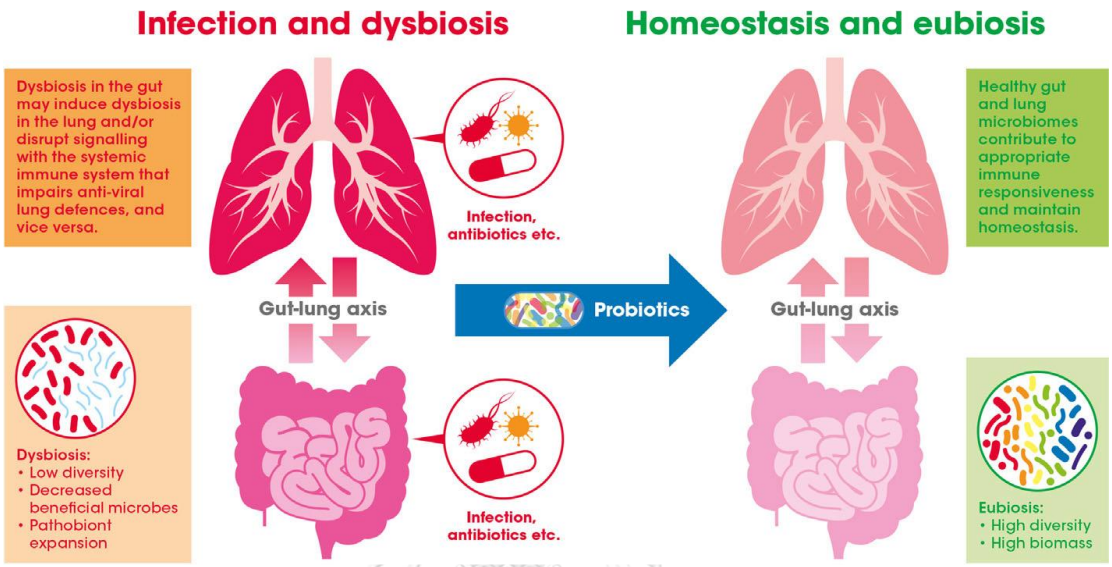


Figure 6. The gut-lung axis crosstalk in the viral respiratory infections circumstance (63).

The microbiome and influenza infection

Alteration in the GI microbiome affects lung immunity, responding to viral infection. For example, the evidence demonstrates that ingestion of *L. paracasei* CNCM I-1518 in mice decreased susceptibility to influenza infection and increased viral clearance (68). Bacterial strain *S. aureus* in the mucosal upper respiratory tract can reduce lung injury by influenza infection via inducing M2 alveolar macrophages and anti-inflammatory cytokines (69). Alteration of pathogenic *S. aureus* to commensal bacteria by co-culture with *Corynebacterium striatum* due to this bacteria has been reported to reduce the transcription of virulence genes of *S. aureus* (70). Interestingly, bacteria-bacteria interaction may define respiratory bacterial communities and also host immune response to viral infection.

Next-generation sequencing (NGS)

Since the complete human genome sequence project was reported in 2003, genome sequencing technology has rapidly developed and played an important role in medical laboratories to understand the complexity of genomes in health and disease (71). Nowadays, second and third-generation NGS technologies are widely used on the market due to providing a short time, low cost, unbiased, generating millions of short sequences at high, and reducing fragment-cloning methods, compared to the traditional Sanger sequencing methods (Figure 7).

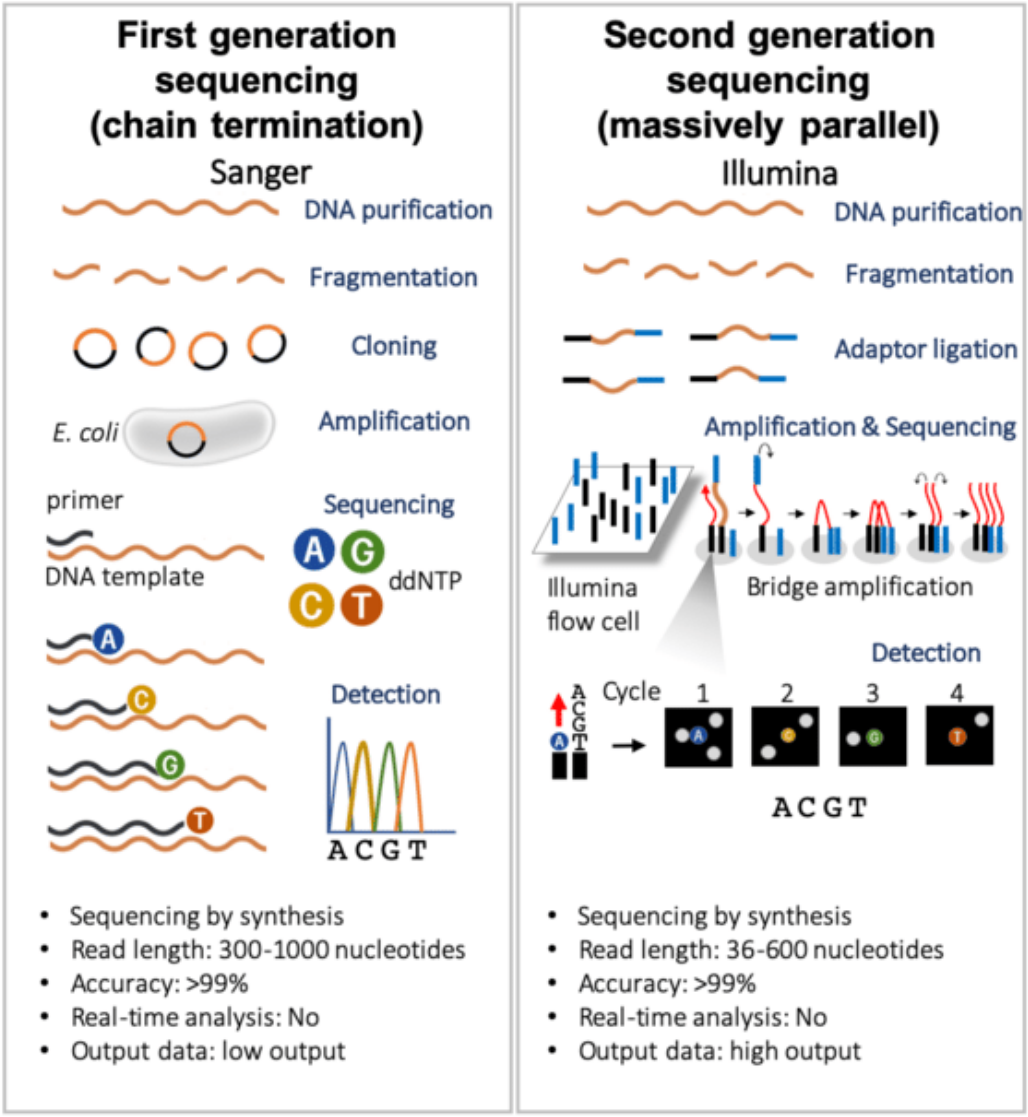


Figure 7. Schematic of first and second-generation sequencing (72).

Illumina sequencing

Illumina sequencing process includes cluster generation, sequencing, image acquisition and base calling and starts with the surface of flow cell occurs bridge clonal amplification of adaptor-ligated DNA fragments. This process generated copies of numerous clonal template DNA that conducted miniaturized colonies. A cyclic reversible termination strategy is a process to read nucleotide bases on the template strand at a time through cumulative rounds of base incorporation, washing, imaging, and cleavage. The fluorescently labeled 3′-O-azidomethyl- dNTPs nucleotides are used to stop the polymerization reaction, and then unincorporated bases are washed. The fluorescent imaging and coupled-charge device (CCD) camera are identified the incorporated nucleotide base (Figure 8.) (73). The sequence data output per run of the Illumina MiSeq platform ranges from 0.3 to 15 Gb (74). Accordingly, illumine NGS technology applies clonal amplification and sequencing by synthesis (SBS) chemistry to capacitate rapid and accurate sequencing. This highly scalable technology can sequence the DNA from any organism in targeted regions or the entire genome through various methods and provide comprehensive information to understand health and disease better. DNA sequencing methods such as untargeted and targeted NGS techniques. Whole-genome sequencing (WGS) is an untargeted, unbiased solution and comprehensive method for investigating the genome which rapidly dropping sequencing costs and able to obtain valuable information about the entire genetic code whereas targeted sequencing can used to sequence and analyze subset of genes or regions of the genome on specific areas of interest such as 16S amplicon sequencing, which is technique based on the amplification of small fragments of hypervariable regions of the 16S ribosomal RNA (rRNA) gene of prokaryotes.

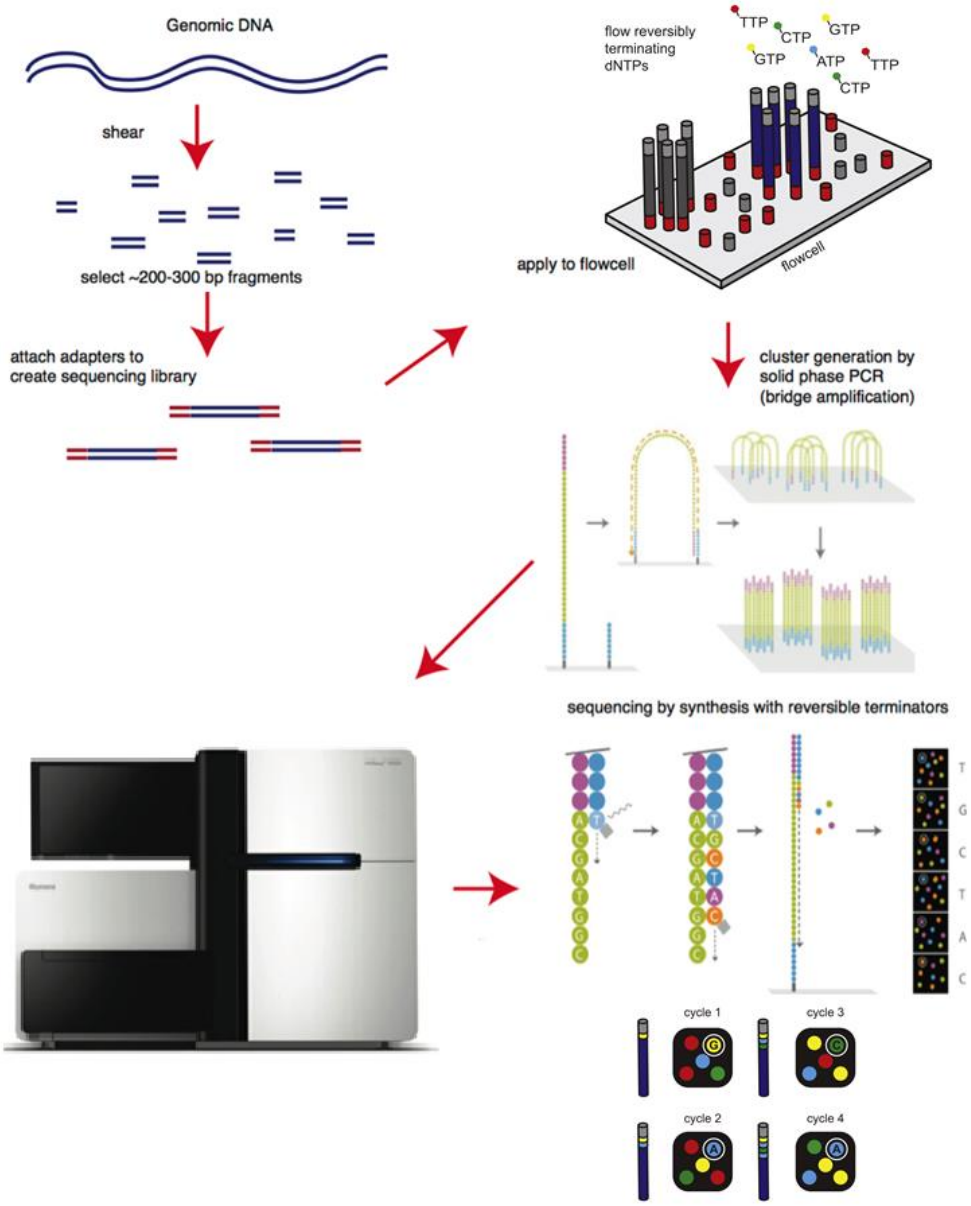


Figure 8. Schematic diagram of NGS process based on Illumina platform (73).

Whole-genome sequencing

NGS technologies have revolutionized the field of genomics and molecular biology. Whole-genome sequencing can provide extensive detailed data on the identification of pathogens and higher resolution data than conventional methods. For example, traditional Sanger sequencing is frequently targeted to define genome fractions that are missing precious information. Moreover, WGS can provide additional genotypic information, tracking the origin of outbreaks and forecasting the spread of diseases such as influenza A and B viruses. Mutations in influenza nucleotide sequences are used to monitor and detect genetic evolution, select appropriate vaccine strains, the emergence of antiviral resistance, and immune evasion, which is important to public health surveillance (75). WGS generates an enormous amount of data in the form of sequence reads. The process for data analysis after obtaining sequence reads includes 1. raw read quality control; 2. data preprocessing; 3. alignment; 4. variant calling; 5. genome assembly; 6. genome annotation; and 7. other advanced analyses such as phylogenetic analysis. The FASTQ files (raw read) were eliminated from poor-quality reads/sequences and adapter sequences by FastQC and read trimming, and then the sequence reads were mapped to a reference genome. After alignment, variants can be identified by comparing the sample genome to the reference genome and then reads are assembled into contiguous sequences, which align overlapping reads to form longer contigs and order the contigs into scaffolds. The identification of genome segments by genome annotation. In-dept of complete virus genome sequencing data could be provided more precise phylogenetic characterization of viral quasispecies within infected hosts, which potential to study in virus transmission (76). Whole-genome sequencing by NGS technologies can detect major and minor nucleotide variants presented in the sequenced virus population, but traditional Sanger sequencing can detect minor nucleotide variants only 10% within a sample (77). Therefore, these NGS technologies efficiently detect low-frequency single nucleotide variants in sequence fragments or whole genomes.

16S amplicon sequencing

16S rDNA Sequencing is a culture-independent method for microbiome analysis. 16S rDNA gene can be found in most bacteria that comprise nine hypervariable regions. These flanking regions are conserved sequences that allow the designing of PCR primers to amplify and sequence hypervariable regions (Figure 9). These can be used for taxonomic classification based on the nucleotide sequence similarity, which is clustered into Operational Taxonomic Units (OTUs). To identify the microorganism within the microbiome, the OTUs are then compared with databases (78). The limitation of traditional Sanger sequencing is missing less abundant members of the microbiomes, which this technique can be sequenced lower proportion of amplicons. Afterward, 16S rDNA sequencing using NGS platforms expanded the capacity to complete the identification of bacterial community members at a much lower cost and provide higher coverage per sample by short amplicon sequencing (79). Due to the 16S rDNA sequencing has become a popular technique to identify the microbiome profile, thus many bioinformatic tools were developed for sequencing data analysis such as USEARCH (ultra-fast sequence analysis) (80), QIIME (Quantitative Insights Into Microbial Ecology) (81), and Mothur (82). However, Illumina sequencing has been limited to identified bacterial classification in the genus level because of a short read sequencing. Challenging the selection of hypervariable regions (V1-V9) for 16S rDNA sequencing depends on published or in-house designed protocols. The most widely used primers set to identify bacterial composition are V3/V4 (primers 341F–805R) and V4 (primers 515F–806R), which provide the region sizes 465 bp and 291 bp, respectively (83, 84). Due to the low cost per sample and the requirement of low input template DNA concentrations for 16S rDNA sequencing are the most popular high throughput sequencing methods.

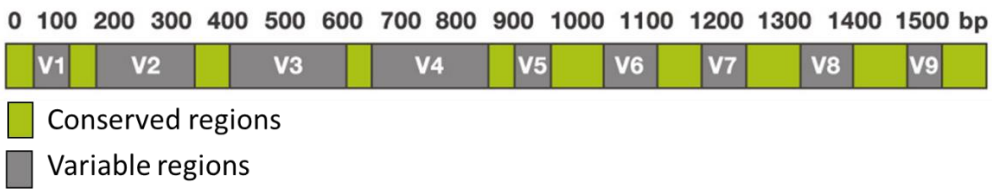


Figure 9. Organization of hypervariable and conserved regions of 16S rRNA gene (85).

Chapter 3

Genome Characterization and Mutation Analysis of Human Influenza A Virus in Thailand

(Accepted to Genomics & Informatics)

Somruthai Rattanaburi[1,2], Vorthon Sawaswong[2], Pattaraporn Nimsamer[2], Oraphan Mayuramart[2], Pavaret Sivapornnukul[2,3], Ariya Khamwut[2], Prangwalai Chanchaem[2], Kritsada Kongnomnan[2], Nungruthai Suntronwong[4], Yong Poovorawan[4], Sunchai Payungporn[2,3,*]

[1] Interdisciplinary Program of Biomedical Sciences, Graduate School, Chulalongkorn University, Bangkok 10330, Thailand

[2] Research Unit of Systems Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

[3] Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

[4] Center of Excellence in Clinical Virology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

***Corresponding author**:

Associate Professor Sunchai Payungporn, Ph.D.

Department of Biochemistry, Faculty of Medicine,

Chulalongkorn University, Bangkok 10330 Thailand

Email: sp.medbiochemcu@gmail.com

Abstract

The influenza A viruses have high mutation rates and cause a serious health problem worldwide. Therefore, this study focused on genome characterization of the viruses isolated from Thai patients based on the next-generation sequencing (NGS) technology. The nasal swabs were collected from patients with influenza-like illness in Thailand during 2017-2018. Then, the influenza A viruses were detected by RT-qPCR and isolated by MDCK cells. The viral genomes were amplified and sequenced by Illumina MiSeq platform. Whole genome sequences were used for characterization, phylogenetic construction, mutation analysis and nucleotide diversity of the viruses. The result revealed that 90 samples were positive for the viruses including 44 of A/H1N1 and 46 of A/H3N2. Among these, 43 samples were successfully isolated and then the viral genomes of 25 samples were completely amplified. Finally, 17 whole genomes of the viruses (A/H1N1, n=12 and A/H3N2, n=5) were successfully sequenced with an average of 232,578 mapped reads and 1,720 genome coverage per sample. Phylogenetic analysis demonstrated that the A/H1N1 viruses were distinguishable from the recommended vaccine strains. However, the A/H3N2 viruses from this study were closely related to the recommended vaccine strains. The nonsynonymous mutations were found in all genes of both viruses, especially in *HA* and *NA* genes. The nucleotide diversity analysis revealed negative selection in the *PB1*, *PA*, *HA* and *NA* genes of the A/H1N1 viruses. High-throughput data in this study allow for genetic characterization of circulating influenza viruses which would be crucial for preparation against pandemic and epidemic outbreaks in the future.

Key words: Influenza A virus, Genome characterization, Thailand, Mutation, NGS

Introduction

Currently, influenza viruses are still a major cause of respiratory disease and can affect all age groups, resulting in a serious public health problem. The estimated infection rate of influenza viruses is approximately 5 to 15% of the population (86). Furthermore, there are more than 500,000 deaths reported worldwide (87). Seasonal influenza is caused by influenza A (A/H1N1 and A/H3N2 subtypes) and influenza B (B/Victoria and B/Yamagata lineages) viruses. However, the influenza A viruses cause more severity, and lead to more epidemics and pandemics due to the high mutation rates which result from antigenic drifts and antigenic shifts (2). First, antigenic drift is caused by the accumulation of point mutations that change the properties of the viral hemagglutinin (HA) and neuraminidase (NA) surface proteins to avoid the host immune system. On the other hand, an antigenic shift is a genetic reassortment process when at least two strains of influenza A viruses have infected within the same cell (3). During viral replication, the high rate of mutation is promoted by error-prone polymerase enzyme (88). The mutation rates of the influenza A virus have been reported within a range of 2.0×10^{-6} to 2.0×10^{-4} mutations per site per round of genome replication (4, 89, 90). Therefore, this evidence suggests that each replicated genome of influenza A carries an average of 2–3 mutations per genome (91). The virus has gradually adapted to its antigenic sites to avoid the host immune response and vaccination (92). Due to the high mutation rates, the influenza vaccine was less effective (only 29% to 61%) against seasonal outbreaks during 2019-2020 (93).

In Thailand, influenza transmission occurs year-round with two annual peaks: a major peak in the rainy season and a minor peak in winter (94). Previous studies have reported that influenza was a major cause of morbidity and mortality in Thailand and resulted in crucial economic costs annually. A study, conducted during 2005–2008, estimated an annual average of 36,400 influenza-associated hospitalizations and 300 deaths occurred, with significantly higher mortality rates in children and the elderly (1). Furthermore, several studies examined the genetic variabilities within *HA* and *NA* genes of influenza A viruses based on Sanger sequencing (95-98). Interestingly, whole genome sequencing (WGS) can be applied to

characterize viral strains and provide the comprehensive information of the influenza genome for better understanding of the viral evolution and novel viral strains (99).

Nowadays, next-generation sequencing (NGS) has the advantages of massively parallel sequencing thus making it the ideal tool for characterization of the viral whole genome, viral reassortment and viral mutations (99, 100). Consequently, whole genome sequencing of influenza viruses based on NGS technology can provide the information necessary to understand the characteristics of influenza viruses. This study aimed to investigate the viral genome and mutations of influenza A viruses circulating in Thailand from 2017 to 2018, and this approach can be further applied for preparation against pandemic and epidemic outbreaks in the future.

Materials and Methods

Sample collection and influenza diagnosis

The study protocol was approved by the Institutional Review Board (IRB No. 337/57) and Institutional Biosafety Committee (MDCU-IBC No. 001/2018) from the Faculty of Medicine, Chulalongkorn University. Briefly, nasal swab samples from patients with influenza-like illness (ILI) were obtained from Bangpakok 9 International Hospital and Chum Phae Hospital during August 2017 to November 2018. The clinical samples were collected and processed for influenza diagnosis and RT-qPCR under the Center of Excellence in Clinical Virology, Faculty of Medicine, Chulalongkorn University. The samples were preserved in viral transport media (VTM) consisting of Hank's Balanced Salt Solution supplemented with 1% bovine serum albumin, amphotericin B (15 µg/mL), penicillin G (100 U/mL) and streptomycin (50 µg/mL). The nasal swab samples were screened for influenza virus infection using a one-step multiplex reverse transcription-quantitative polymerase chain reaction (RT-qPCR) as described previously (101, 102). Briefly, the assay was performed in a 10 µL final volume, containing 1 µL of RNA sample, 5 µL of 2X reaction mix, 0.2 µL of SuperScript™ III RT/Platinum™ *Taq* Mix (Invitrogen, USA), an additional 0.1 mM of $MgCl_2$, 0.25 µM of each primer, and 0.125 µM of each probe. The One-step multiplex RT-qPCR was performed on the StepOnePlus™ Real-Time PCR System (Applied

Biosystems™, USA) using the following thermal cycling conditions: at 55°C for 30 min for reverse transcription, followed by 95°C for 10 min, continuing with 40 cycles of 95°C for 15 sec and 60°C for 30 sec.

Cell cultures

Madin-Darby Canine Kidney (MDCK) cells were obtained from the American Type Culture Collection (ATCC®, USA) and cultured in Dulbecco's Modified Eagle Medium (DMEM) with high glucose (HyClone™, USA) supplemented with 10% fetal bovine serum (FBS) (Gibco, USA) and 1% (v/v) penicillin/streptomycin (Gibco, USA) maintained under humidified 5% CO_2 at 37°C (103).

Influenza virus isolation

MDCK cells were used for influenza virus isolation and propagation as described in the previous study (104). Briefly, MDCK cells were seeded in 60 mm tissue culture dishes (SPL Life Science, Korea) at $5x10^5$cells per dish in DMEM medium without antibiotics. When the cells reached around 80% confluence, the media were removed and then washed by phosphate buffer saline (PBS) (Amresco, USA). Positive influenza samples were used for virus isolation. Briefly, 500 µL of a nasal swab from influenza-positive samples was mixed with 500 µL of DMEM with high glucose (HyClone™, USA) and filtered through 0.22 µm filter (Millipore, USA). The filtrate was immediately processed to influenza viral propagation. Three hundred microliters of each filtered influenza-positive sample were mixed with 200 µL infection medium (DMEM-high glucose supplemented with 2 mM L-glutamine and 0.5 µg/mL TPCK-trypsin). The mixture was added in each dish and then incubated in 5% CO_2 at 37°C for 1 hour. After incubation, the virus suspension was removed and then washed with PBS. Finally, the cells were overlaid with fresh infection medium and incubated under 5% CO_2 at 37°C for 48 hours. After that, the cytopathic effect (CPE) of infected cells was observed and the viral supernatant was collected. Each sample was isolated in three serial passages (P0-P2). The viral titers were quantified by RT-qPCR (105).

Viral RNA extraction and reverse transcription

One hundred and fifty microliters of the supernatant in each isolation passage were extracted in a GenUp™ Viral RNA kit (Biotechrabbit, Germany) according to the manufacturer's instructions, and eluted in 60 µL with warm RNase-free water. The concentration of total viral RNA was quantified by using Nanodrop UV spectrophotometer (Implen, Germany). Three hundred nanograms per microliter of viral RNA were reverse transcribed into cDNA using the RevertAid™ First Strand cDNA Synthesis Kit (Thermo Scientific, USA) according to the manufacturer's instructions with 10 µM RT primer (5′-ACGCGTGATCAGCAAAAGCAGG-3′) that is conserved and complemented with 12 nucleotides at the 3′ ends of each influenza A viral genes (106). The mixtures were incubated at 42°C for 1.5 hours. The cDNAs were stored in -20°C for further analysis.

Quantitative Real-time PCR

To determine the amount of influenza virus in each sample passage, the viral matrix (M) gene was amplified based on StepOnePlus™ Real-time PCR Systems (Applied Biosystem, USA) using SYBR Green Luna® Universal qPCR Master Mix (New England Biolabs, USA) as described above. The results were analyzed by StepOnePlus™ Software v2.3. The samples amplified with Ct values lower than 28 were interpreted as positive influenza viral propagation (107).

Amplifications of influenza genomes

The viral cDNAs from the previous step were used as templates for genome amplifications. The primer sets; forward primer (5′-ACGCGTGATCAGCAAAAGCAGG-3′) and reverse primer (5′-ACGCGTGATCAGTAGAAACAAGG-3′) were used for amplification of influenza A viral genes (8 segments) following the previous study (106). Briefly, PCR master mix composed of 1.25 µM of each primer, 0.35 mM of dNTPs, 0.02 U/µL of Phusion High-Fidelity DNA polymerase (Thermo Scientific, USA), 7.5 µL of cDNA and nuclease-free water to a final volume of 50 µL. Subsequently, 15 µL of the amplicons were analyzed by 1% agarose gel electrophoresis. The amplicons were purified by the QIAquick PCR Purification kit (Qiagen, Germany) following the

manufacturer's protocol. The concentrations of purified PCR products were measured by the Qubit dsDNA High-Sensitivity assay kit (Invitrogen, USA).

DNA library preparation and next-generation sequencing

The purified amplicons (1 µg in 130 µL) from the genome amplification step were sheared to approximately 200 bp fragments by the Covaris™ M220 Focused-ultrasonicator™ (Covaris, UK) with 20% duty factor, 50 unit of peak incident power (W), 200 cycles per burst for 150 seconds. The fragmented DNAs were used for DNA library preparation by using NEBNext Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, USA) following the manufacturer's instructions. Briefly, 50 µL of DNA fragments were ends repaired and subsequently adapters ligated by using the NEB ligase master mix. Then, DNA libraries (approximately 320 bp) were cleaned up and size selected by AMPure XP beads (Beckman Coulter, USA). For library enrichment, PCR amplification was carried out by adding the Illumina MiSeq-compatible indexes to the DNA libraries. Afterwards, the enriched DNA libraries were purified by 2% agarose gel electrophoresis with 100 volts for 20 minutes and size selected (approximately 320 bp). Finally, the total DNA libraries were quantified by real-time PCR using the KAPA Library Quantification Kits (Kapa Biosystems, USA). After that, the concentration of each sample was determined and pooled equally at 2 nM of each library. Subsequently, the pooled library was then diluted to 6 pM and paired-end sequenced (2x150 bp) on an Illumina MiSeq instrument using MiSeq Reagent Kits v2 (300 cycles) according to the manufacturer protocol (Illumina, USA).

Influenza genome analysis

The MiSeq Reporter Software version 2.4 was used for the primary analysis of FASTQ sequencing data. Low-quality reads (Q-score < 30) and adaptors were trimmed. The passing filtered reads (Q-score ≥ 30) were aligned with the vaccine strains of influenza A reference genomes (A/California/07/2009 (H1N1) or A/South Australia/55/2014 (H3N2)) for genome characterization and mutation analysis by using CLC Genomics Workbench software (QIAGEN, Germany). Mutation patterns and frequencies were generated by using GraphPad Prism version 6.01 software. The

FASTQ files and FASTA files of influenza genome sequences were submitted to the Sequence Read Archive (SRA) [BioProject ID: PRJNA576776] and GenBank as shown in Supplementary Table 1.

Phylogenetic analysis

In this study, the HA and NA deduced amino acid sequences of A/H1N1 and A/H3N2 were aligned with reference strains retrieved from the Global Initiative on Sharing All Influenza Data (GISAID) EpiFlu database by using the Clustal W program, implemented in the BioEdit sequence alignment editor software v.7.2.5 (108). Phylogenetic analysis was performed by mean of the maximum likelihood method (1,000 bootstrapping replicates) and LG with Freqs.(+F) model (discreate gamma distribution with 5-rate categories and complete deletion data subset) using the MEGA X sorfware (109).

Sliding windows analysis of nonsynonymous nucleotide variation

The nucleotide diversity (π) within each gene of influenza A/H1N1 and A/H3N2 viruses was evaluated by PoPoolation v.1.2.2 to investigate the genetic variations of viruses within the sample (110). The sliding window analysis of nonsynonymous nucleotide variation (πN) was performed based on Syn-nonsyn-sliding.pl script with the window size of 9 codons and a step size of 1 codon. The average corresponding πN values were calculated and plotted to a middle position of the windows to demonstrate the degree of nonsynonymous substitutions within 8 viral gene segments. In addition, the nonsynonymous nucleotide variations (πN) per synonymous nucleotide variation (πS) were analyzed by the Syn-nonsyn-at-position.pl script to investigate the neutrality of selection in each segment. The πN/πS ratios per gene in each influenza subtype were calculated as the average value from individual samples. Lastly, a paired Wilcoxon signed-rank test (p-value < 0.05) was used to compare pooled average πN and πS values within each subtype of influenza viruses.

Results

Detection and isolation of influenza A viruses

In this study, 500 nasal swab samples were collected from patients with influenza-like illness and detected for influenza A virus by RT-qPCR. Ninety samples were influenza A virus-positive samples including 48.9% (44 samples) of A/H1N1 and 51.1% (46 samples) of A/H3N2 as shown in Table 2. Among these ninety samples, 43 samples (29 of A/H1N1 and 14 of A/H3N2) were successfully isolated in the second passage (P2) of MDCK cells with Ct value ranging from 13 to 28 (Table 3).

Table 2. The amount of positive influenza A samples obtained from RT-qPCR, virus isolation, genome amplification and NGS.

	Positive samples	Positive virus isolation	Genome amplification	NGS
Influenza A/H1N1	44	29	17	12
Influenza A/H3N2	46	14	8	5
Total	**90**	**43**	**25**	**17**

Whole genome sequencing and characterization of influenza A viruses

From 43 isolated samples, 25 samples were completely amplified as a full genome including 17 samples of A/H1N1 and 8 samples of A/H3N2. Finally, 17 samples (12 samples of A/H1N1 and 5 samples of A/H3N2) passed the quality control of libraries preparation for NGS as shown in Table 2. The result revealed that 17 whole genomes of influenza A viruses were successfully sequenced with an average of 424,151 total reads per sample, 232,578 mapped reads per sample and 1,720 genome coverage per sample (Table 3). Therefore, these results were highly confident for genome annotation and mutation analysis. The FASTQ data were deposited as BioProject accession no. PRJNA576776 and influenza genome sequences were submitted into GenBank database as summarized in Supplementary Table 1.

Table 3. The sample characteristic, virus isolation and NGS data of influenza A virus in this study.

No.	Sample name	Age	Sex	Ct from each passage			Total reads	Mapped reads	Average length of mapped read (bp)	Average genome coverage
				P0	P1	P2				
1	A/Thailand/CU-B23883/2017 (H1N1)	2	M	38	30	27	318,034	54,644	67.2	237.7
2	A/Thailand/CU-B24063/2017 (H1N1)	19	F	46	34	17	171,064	163,257	82.4	905.6
3	A/Thailand/CU-B24069/2017 (H1N1)	39	M	36	23	18	325,424	309,926	84.4	1,784.8
4	A/Thailand/CU-B24076/2017 (H1N1)	12	M	36	33	28	143,698	134,494	79.6	726.3
5	A/Thailand/CU-B24660/2017 (H1N1)	51	F	33	18	16	518,852	442,915	96.0	2,907.9
6	A/Thailand/CU-B25124/2017 (H1N1)	3	M	31	32	20	134,102	129,280	87.8	771.9
7	A/Thailand/CU-B25506/2017 (H1N1)	38	F	32	27	18	165,684	159,654	88.7	952.3
8	A/Thailand/CU-B27534/2017 (H1N1)	31	F	28	13	13	124,720	112,724	61.6	477.5
9	A/Thailand/CU-B29642/2018 (H1N1)	30	F	32	29	16	1,355,512	520,427	113.6	3,927.2
10	A/Thailand/CU-B30312/2018 (H1N1)	59	F	31	27	27	199,458	180,341	120.6	1,460.6
11	A/Thailand/CU-B30648/2018 (H1N1)	29	F	33	22	15	702,634	367,078	113.6	2,771.4
12	A/Thailand/CU-E1180/2018 (H1N1)	2	M	20	30	13	555,488	357,666	100.3	2,351.7
13	A/Thailand/CU-B24411/2017 (H3N2)	61	F	34	26	23	217,002	68,933	60.0	284.1
14	A/Thailand/CU-B24666/2017 (H3N2)	2	F	27	31	15	139,802	90,676	112.1	696.8
15	A/Thailand/CU-B28277/2017 (H3N2)	24	M	23	19	24	583,222	467,163	129.5	4,229.7
16	A/Thailand/CU-B29296/2017 (H3N2)	52	F	30	23	22	557,710	22,569	132.3	2,013.4
17	A/Thailand/CU-B30632/2018 (H3N2)	53	M	31	24	24	998,160	372,082	107.3	2,747.7
Average							**424,151**	**232,578**	**96.3**	**1,720.4**

Phylogenetic analysis of influenza A viruses in Thailand

The HA and NA deduced amino acid sequences were used for phylogenetic analysis. The sequences were aligned with the HA and NA deduced amino acid sequences of the influenza vaccine strains (southern hemisphere influenza seasons during 2012-2019) recommended by the World Health Organization (WHO). The influenza A/H1N1 viruses isolated from this study during 2017-2018 belonged to genetic subclade 6B.1. Interestingly, the results demonstrated that A/H1N1 viruses were closely related to influenza (A/California/7/2009) strain and distinguishable from the recommended influenza vaccine strains for use in 2017-2019 (A/Michigan/45/ 2015 (H1N1)) as shown in Figure 10A and Figure 10B. On the other hand, the circulating A/H3N2 strains, classified into subclade 3C.2a1 and 3C.2a2, were comparatively more closely related to the 2017-2018 WHO influenza vaccine strains (A/Hong Kong/4801/2014 (H3N2) and A/Singapore/INFIMH-16-0019/2016 (H3N2)) as shown in Figure 11A and Figure 11B.

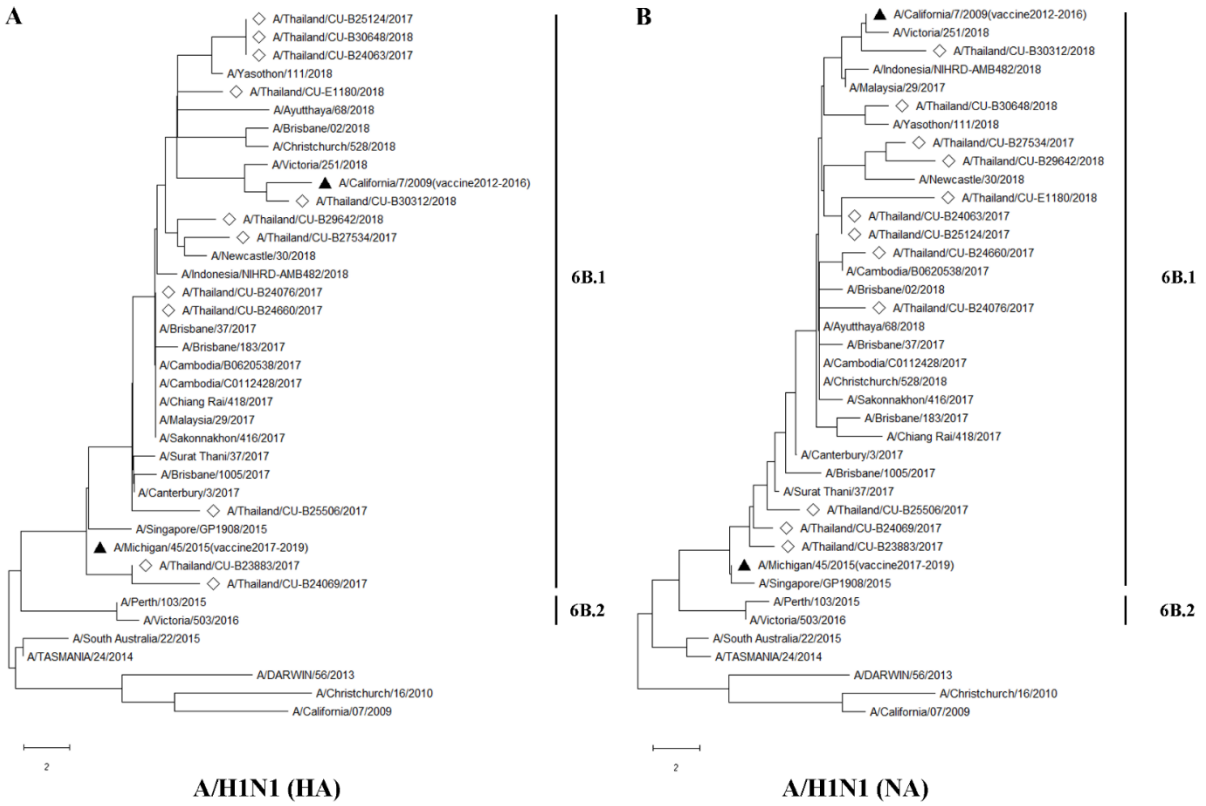


Figure 10. The phylogenetic analysis of influenza A viruses (H1N1) circulating in Thailand during 2017-2018 (diamond) compared with several WHO recommended influenza vaccine strains (black triangle). The HA (A) and NA (B) deduced amino acid sequences were analyzed based on mean of maximum likelihood with 1,000 bootstrapping and LG with Freqs.(+F) model (discreate gamma distribution with 5-rate categories and complete deletion data subset).

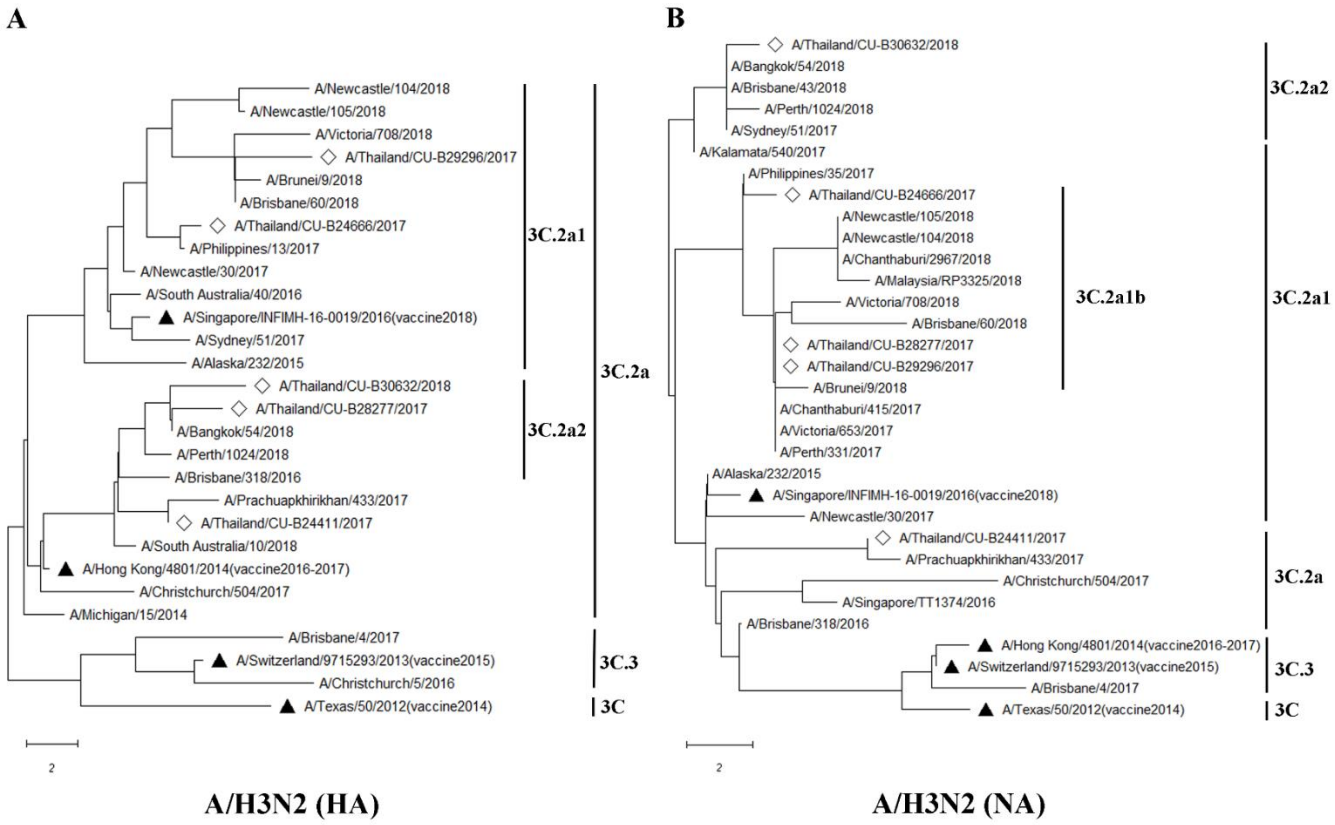


Figure 11. The phylogenetic analysis of influenza A viruses (H3N2)circulating in Thailand during 2017-2018 (diamond) compared with several WHO recommended influenza vaccine strains (black triangle). The HA (A) and NA (B) deduced amino acid sequences were analyzed based on maximum likelihood with 1,000 bootstrapping and LG with Freqs.(+F) model (discreate gamma distribution with 5-rate categories and complete deletion data subset).

Nucleotide diversity of influenza A viruses.

The variations of nonsynonymous within influenza A viruses among the samples in this study are summarized in Figure 12. As shown in Figure 12A, strong signals appeared in the polymerase (*PB2*, *PB1*, and *PA*) genes as well as in the NP gene in A/H1N1 virus. However, the *NA*, *M* and *NS* genes showed the low nonsynonymous variations. Interestingly, the *HA* gene contained the pattern of the variation signals around the middle position of this A/H1N1 gene. As for the results of A/H3N2 (Figure 12B), the polymerase genes were presented as sharp and multiple peaks of the nonsynonymous nucleotide diversity. In addition, the *HA*, *NP*, and *NA* genes of A/H3N2 displayed sharp signals at the beginning and the end of these genes. Furthermore, the *M* and *NS* genes of the A/H3N2 only had peaks around the middle of the genes.

Exploring deeper detail about the direction of diversity, the ratios of nonsynonymous to synonymous nucleotide diversity (πN/πS) were introduced to examine the changes in nucleotide variations. In brief, the πN/πS ratios >1 indicate that selective pressure promotes the new variations (positive selection). In contrast, the πN/πS ratios < 1 refer to the new variation being unfavored (negative selection). In addition, the πN/πS ratios ≈ 1 suggests that neutral evolution occurs in these new variations. According to the results shown in Figure 13, there was no significant positive selection occurring in this study. However, the statistically significant negative selections (p-value < 0.05) were found in *PB1*, *PA*, *HA*, and *NA* genes of A/H1N1. Meanwhile, the A/H3N2 exhibited random selections within these 10 genes due to there being no significant difference observed in the πN/πS ratios.

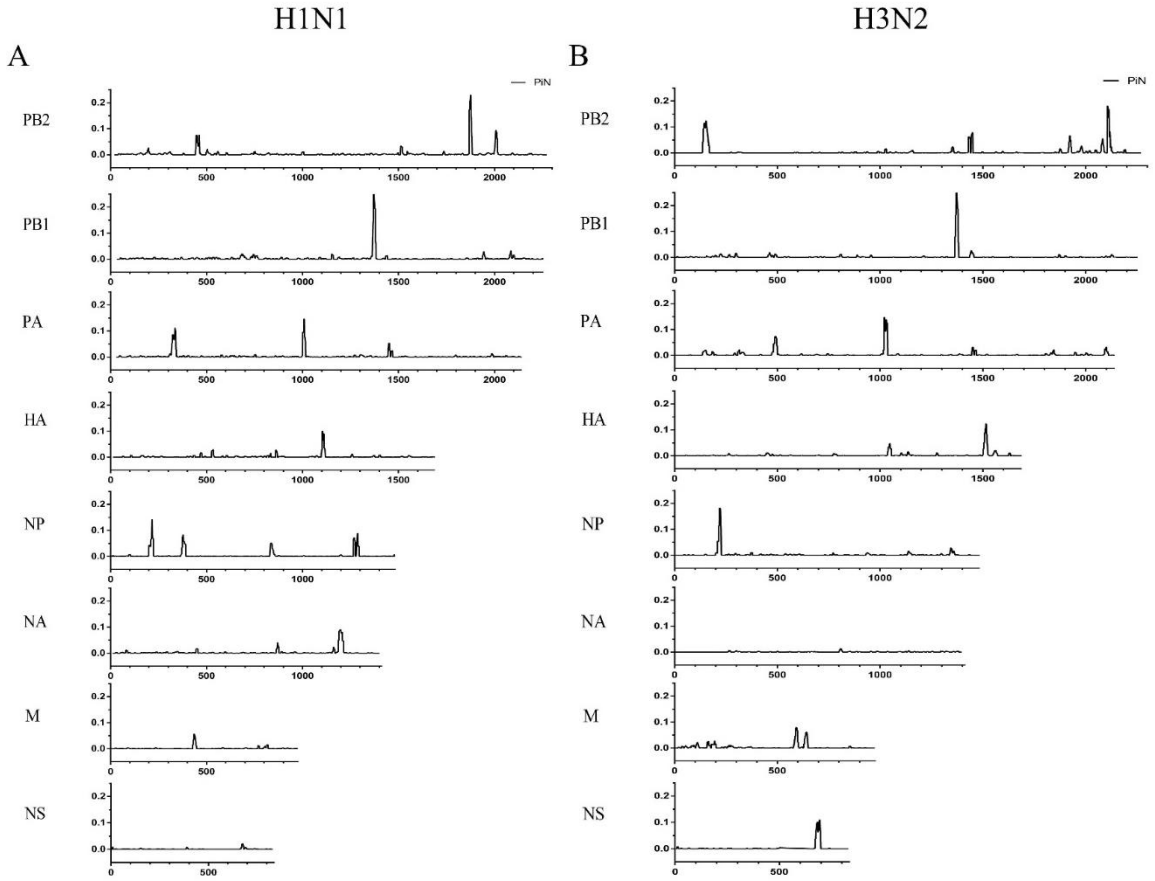
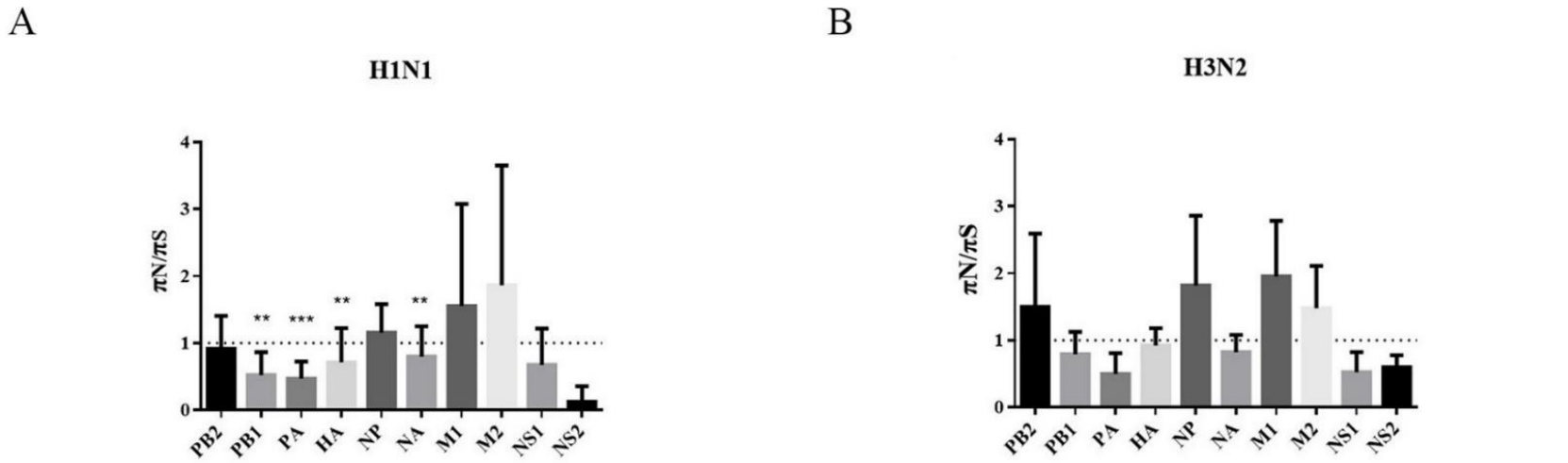


Figure 12. Sliding windows analysis of nonsynonymous nucleotide variation (**π**N) in eight genes of influenza A virus subtypes H1N1 (A) and H3N2 (B). The **π**N values were determined by sliding windows with the window size of 9 codons and a step size of 1 codon. The mean corresponding **π**N values were calculated and plotted to a middle site of the windows.



Gene	PB2	PB1	PA	HA	NP	NA	M1	M2	NS1	NS2
mean πN/πS	0.917	0.574	0.516	0.711	1.157	0.797	1.555	2.035	0.811	0.501
S.D.	0.490	0.312	0.214	0.514	0.426	0.454	1.525	1.771	0.490	0.126
P-value	0.151	0.002	0.001	0.003	0.151	0.009	0.850	0.052	0.266	0.519

Gene	PB2	PB1	PA	HA	NP	NA	M1	M2	NS1	NS2
mean πN/πS	1.498	0.793	0.621	0.924	1.819	0.820	1.955	1.479	0.657	0.600
S.D.	1.092	0.333	0.164	0.259	1.043	0.263	0.831	0.634	0.073	0.177
P-value	0.438	0.313	0.125	0.188	0.125	0.313	0.063	0.125	0.125	0.063

Figure 13. The ratio of nonsynonymous nucleotide variation (πN) to synonymous nucleotide variation (πS) analysis in eight genes of influenza A virus subtypes H1N1 (A) and H3N2 (B). The significant at *p*-value * < 0.05, ** < 0.01 and *** < 0.001. The ratio πN/πS > 1: positive selection, πN/πS < 1: negative selection and πN/πS ≈ 1: neutral evolution. The mean πN/πS, S.D, and *p*-value (Student's t-test) in each segment were summarized at the bottom of the figure.

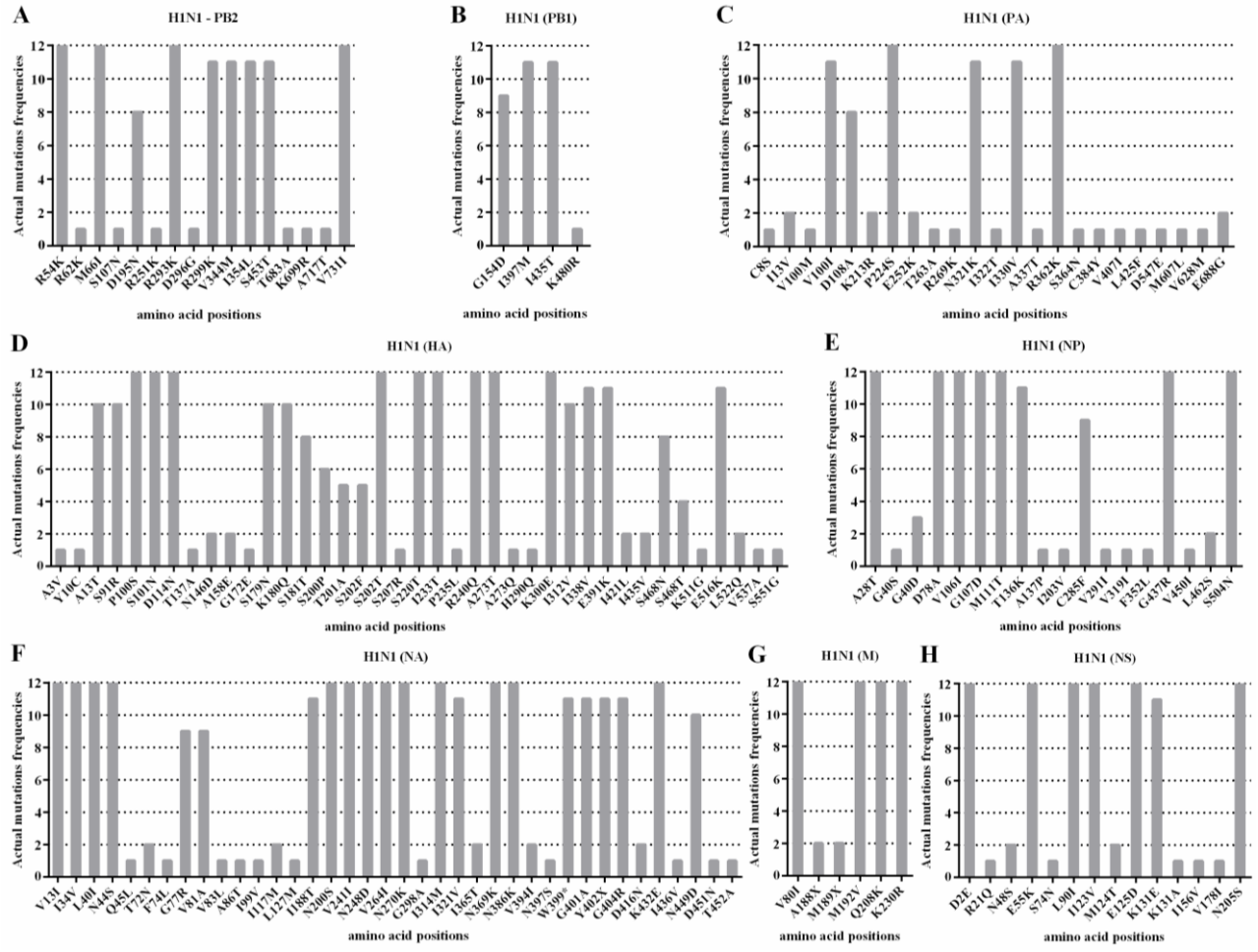


Figure 14. Mutation patterns with actual mutations frequencies observed in each viral gene segments of influenza A virus (H1N1). Amino acid changes were compared to the reference sequence (A/California/07/2009 (H1N1)).

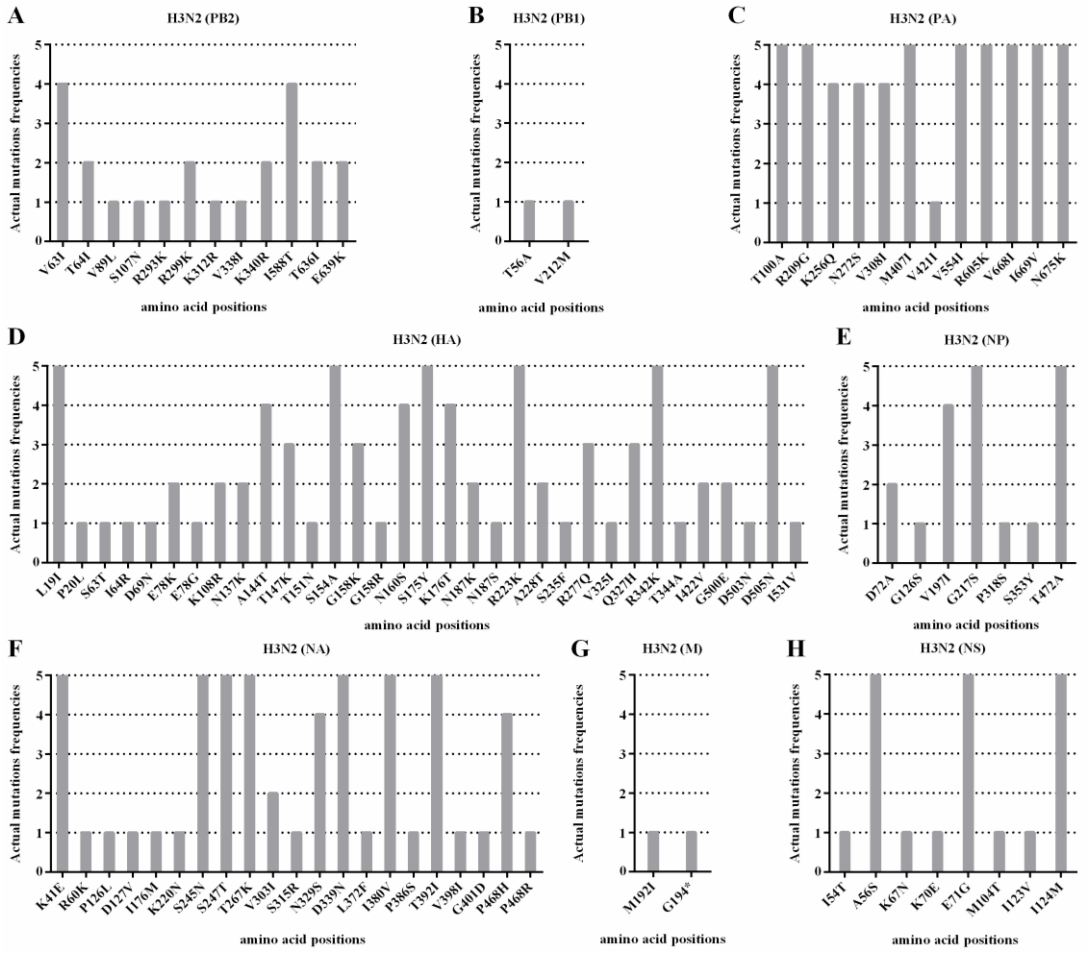


Figure 15. Mutation patterns with actual mutations frequencies observed in each viral gene segments of influenza A virus (H3N2). Amino acid changes were compared to the reference sequence (A/South Australia/55/2014 (H3N2)).

Discussion

In this study, 90 out of 500 (18%) nasal swabs obtained from Thai patients with influenza-like-illness during 2017-2018 were positive for influenza A virus detection based on RT-qPCR detection. The percentage of influenza A virus positive in this research was slightly higher than those reported in the previous study (13.2%) during 2016-2017 in Thailand (111). Previously, several studies have demonstrated that the appropriate quality and quantity of DNA are important for the successful NGS platform sequencing (112-114). In particular, this study has successfully isolated 47.78% (43 of 90 samples) which are positively identified as the influenza A virus, which is higher than the previous study (3.04% of isolation rate) (115). Also, the positive virus isolations (58.14%, 25 of 43 isolates) can be amplified with universal primers, following the study of Zhou et al., which is appropriate for whole genome characterization and mutation analysis of influenza A virus (116). For NGS analysis, the result revealed that 17 whole genomes of influenza viruses were successfully sequenced with 232,578 mapped reads (424,151 total reads), average read length of 96.3 bp and average 1,720.4 genome coverage. Furthermore, the complete sequences of the viral genomes provide reliable and highly informative data despite the average genome coverages, depth coverage, which ranged from 237.7 to 4,229.7 (Table 3). The advantages of the NGS-based technique are that it provides the full genome segment and whole genome of influenza virus, as well as effectively reducing both the turnaround time and cost per nucleotide sequence for the whole genome when compared to the Sanger sequencing method (117-119). However, the sequencing with the Sanger method does not provide the data for quasispecies and nucleotide diversity analysis. Interestingly, the NGS provides more information for minor mutations and selection pressures within the viral genome. Indeed, the nucleotide variations obtained from NGS can be applied for calculation of viral nucleotide diversity within each sample (120, 121).

The number of mutations in the *HA* and *NA* genes of A/H1N1 might affect the efficiency of a vaccine, and related to deduced amino acid sequences of phylogenetic tree (Figure 10A and Figure 10B). The vaccine effectiveness of the 2017-

2018 flu vaccine against both influenza A viruses is approximately 25 to 52% in Europe and 27 to 44% in the United States (122, 123). Interestingly, the result of the influenza A/H1N1 phylogenetic tree with deduced amino acid sequences, which belongs to clade 6B.1, showed a distinction between vaccine strains for 2017-2018 (A/Michigan/42/2015) and our A/H1N1 samples. This result implied that the vaccine might be less effective against A/H1N1 in Thailand. Moreover, the report from the US Centres for Disease Control and Prevention (CDC) also showed the vaccine effectiveness against A/H1N1 was 65% (124). However, the phylogenetic analysis of both HA and NA deduced amino acid sequences revealed the closer relationship between A/H3N2 isolates (clade 3C.2a1 and 3C.2a2) and A/Hong Kong/4801/2014 strain which was the recommended vaccine virus for A/H3N2 (125). Therefore, these results implied that the recommended vaccine was more effective against the influenza A/H3N2 in Thailand during 2017-2018. Indeed, the phylogenetic trees of influenza A/H1N1 and A/H3N2 obtained in this study were correlated with recent genetic and antigenic characterizations of influenza viruses in Thailand (126).

In this study, the genome of circulating influenza A viruses in Thailand during 2017-2018 was characterized. The result from NGS analysis not only provided the full genome of the virus but also acquired the amino acid substitutions across 8 segmented genes. Moreover, there were several known functional mutations of influenza A/H1N1 that had been already characterized. Firstly, the mutations at I354, V344M, and S453T in the PB2 could regulate in the cap-snatching from host RNAs during the viral RNA transcription process (127). Furthermore, N321K in the PA was reported to increase the polymerase complex activity and the viral replication in the cell culture (127, 128). The amino acid substitution at V100I in the PA-X could trigger down-regulated innate immune response genes. Indeed, the amino substitutions at S91R, S181T, I312V, and E391K in the HA might be related to adaptive genetic variations that alter the salt bridge pattern and the membrane fusion stability for major antigenic sites and glycan specificity (127, 128). Three mutations (K180Q, S202T and S220T) were located in the HA antigenic sites, which might be involved in the pathogenicity and contributed to the epidemic (129). Moreover, the mutation (S220T) was observed to affect the infectivity and transmissibility of the virus in humans

(130). The mutation (R240Q) was found in the receptor-binding domain of the HA, which has been reported to increase virus growth (131). The amino acid variations (D114N, K180Q, S202T, S220T and K300E) were responsible for loss of antibody neutralization and decreased overall vaccine effectiveness (132-134). The amino acid substitutions (N44S, V241I, and N369K) in the NA have been reported to facilitate the stability of the virus (135). The I188T and N449D mutations in the NA found in this study are similar to those reported in the previous study (136); however, the function of the mutations has not been well characterized. Additionally, the nonsynonymous mutation at E55K, L90I, I123V, E125D, K131E, and N205S in the NS1 involve the inhibition of host gene expressions related to the interferon response (137, 138). Indeed, the E125D mutation in NS1 interacts with cellular cleavage and polyadenylation specificity factor 30 (CPSF30), which is considered potential in host adaptation to influenza A/H1N1 virus (139, 140).

In the influenza A/H3N2, the previous reports found R277Q and D69N at the antigenic epitope C, N137K and N187K at the antigenic epitope D and E78K/G at the antigenic epitope E of the HA (141, 142). Among these, four amino acid substitutions (N137K, N187K, I422V and G500E) belong to clade 3C.2a.1 and are represented by A/Singapore/INFIMH-16-0019/2016(H3N2) virus (143). The T151N substitution in HA protein was related to the potential N-glycosylation site, affecting antigenic and other viral properties. Moreover, the Q327H substitution in the HA was suggested to bind host proteins (143). Since 2016, the accumulation of mutation at S245N of the NA has contributed to an N-glycosylation site. These mutations (S245N, S247T and P468H) were introduced to the NA antigenic drift of the circulating A/H3N2 virus (144). However, N329S mutation could result in a loss of N-glycosylation in the NA (145). The V303I substitution has been observed in the NA protein with a low resistance to neuraminidase inhibitors (146). Indeed, most mutations of influenza A viruses observed in this study were identified as novel mutations which have not been reported yet (Figure 14 and Figure 15). However, the function of the novel mutations needs to be further investigated.

Nonsynonymous (πN) and synonymous (πS) mutations of the viruses can be accessed by NGS leading to nucleotide diversity (πN/πS) analysis. According to the previous study, the deep sequencing of A/Wisconsin/67/2005 (H3N2) revealed that the positive selection was observed in the viruses isolated from the chicken kidney (CK), Vero cell culture and embryonated chicken eggs, whereas the negative selection was found in virus from direct intranasal inoculation in the human challenge (121). There was no significant nucleotide diversity observed in A/H3N2 viruses in our study, and this might be due to the strain of the virus, host cell, or limited numbers of the sample. For πN/πS analysis of influenza A/H1N1 viruses, the mutations existing in the viral genes with statistical significance were *PB1*, *PA*, *HA*, and *NA* genes in which these mutations were suggested as negative selection. Therefore, to further investigate the πN and πS variations, the sliding window analysis of those genes was performed to ensure that the negative selections were not the outcome of the averaging value across the entire gene. The results of sliding window analysis were consistent with the negative selections from the πN/πS analysis in those genes at which the πN signals were high and sharp at some regions of the genes, while the rest of the genes were relatively low in the πNsignals

Conclusion

In summary, the NGS was successfully applied for whole genome characterizations of influenza A/H1N1 and A/H3N2 viruses that provide the high-throughput data for phylogenetic construction, mutation analysis and nucleotide diversity. The results revealed that the recommended vaccine A/H1N1 strain might be less effective against the A/H1N1 virus. Moreover, several mutations were demonstrated in both A/H1N1 and A/H3N2, especially in *HA* and *NA* genes. Finally, the negative selections were found in the *PB1*, *PA*, *HA* and *NA* genes of the A/H1N1. Unfortunately, limited number of samples were successfully propagated, amplified and sequenced in this study. Nevertheless, the whole genome data obtained from this study might be useful for mutation analysis and can be compared with data obtained from other studies in the future.

Supplementary Table 1. The summary of accession numbers of influenza A virus in this study.

Strain name	SRA accession no. FASTQ data	FASTA accession no.							
		PB2	*PB1*	*PA*	*HA*	*NP*	*NA*	*M*	*NS*
A/Thailand/CU-B23883/2017(H1N1)	SRR10256725	MN560613	MN560614	MN560615	MN560616	MN560617	MN560618	MN560619	MN560620
A/Thailand/CU-B24063/2017(H1N1)	SRR10256724	MN560778	MN560779	MN560780	MN560781	MN560782	MN560783	MN560784	MN560785
A/Thailand/CU-B24069/2017(H1N1)	SRR10256713	MN560975	MN560976	MN560977	MN560978	MN560979	MN560980	MN560981	MN560982
A/Thailand/CU-B24076/2017(H1N1)	SRR10256710	MN561022	MN561023	MN561024	MN561025	MN561026	MN561027	MN561028	MN561029
A/Thailand/CU-B24660/2017(H1N1)	SRR10256709	MN561043	MN561044	MN561045	MN561046	MN561047	MN561048	MN561049	MN561050
A/Thailand/CU-B25124/2017(H1N1)	SRR10256708	MN561055	MN561056	MN561057	MN561058	MN561059	MN561060	MN561061	MN561062
A/Thailand/CU-B25506/2017(H1N1)	SRR10256707	MN561178	MN561179	MN561180	MN561181	MN561182	MN561183	MN561184	MN561185
A/Thailand/CU-B27534/2017(H1N1)	SRR10256706	MN561198	MN561199	MN561200	MN561201	MN561202	MN561203	MN561204	MN561205
A/Thailand/CU-B29642/2018(H1N1)	SRR10256705	MN561252	MN561253	MN561254	MN561255	MN561256	MN561257	MN561258	MN561259
A/Thailand/CU-B30312/2018(H1N1)	SRR10256704	MN561279	MN561280	MN561281	MN561282	MN561283	MN561284	MN561285	MN561286
A/Thailand/CU-B30648/2018(H1N1)	SRR10256723	MN561288	MN561289	MN561290	MN561291	MN561292	MN561293	MN561294	MN561295
A/Thailand/CU-E1180/2018(H1N1)	SRR10256722	MN561304	MN561305	MN561306	MN561307	MN561308	MN561309	MN561310	MN561311
A/Thailand/CU-B24411/2017(H3N2)	SRR10256721	MN561330	MN561331	MN561332	MN561333	MN561334	MN561335	MN561336	MN561337
A/Thailand/CU-B24666/2017(H3N2)	SRR10256720	MN561322	MN561323	MN561324	MN561325	MN561326	MN561327	MN561328	MN561329
A/Thailand/CU-B28277/2017(H3N2)	SRR10256719	MN561350	MN561351	MN561352	MN561353	MN561354	MN561355	MN561356	MN561357
A/Thailand/CU-B29296/2017(H3N2)	SRR10256718	MN561279	MN561280	MN561281	MN561282	MN561283	MN561284	MN561285	MN561286
A/Thailand/CU-B30632/2018(H3N2)	SRR10256717	MN560639	MN560640	MN560641	MN560642	MN560643	MN560644	MN560645	MN560646

Bacterial microbiota in upper respiratory tract of COVID-19 and influenza patients

(Published in Experimental Biology and Medicine 2021 Nov 13:15353702211057473
DOI: https://doi.org/10.1177/15353702211057473)

Somruthai Rattanaburi[1,2], Vorthon Sawaswong[2], Suwalak Chitcharoen[2], Pavaret Sivapornnukul[2,3], Pattaraporn Nimsamer[2], Nungruthai Suntronwong[4], Jiratchaya Puenpa[4], Yong Poovorawan[4], Sunchai Payungporn[2,3,*]

[1]Interdisciplinary Program of Biomedical Sciences, Graduate School, Chulalongkorn University, Bangkok 10330, Thailand
[2]Research Unit of Systems Microbiology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand
[3]Department of Biochemistry, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand
[4]Center of Excellence in Clinical Virology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

***Corresponding author**:
Associate Professor Sunchai Payungporn, Ph.D.
Department of Biochemistry, Faculty of Medicine,
Chulalongkorn University, Bangkok 10330 Thailand
Email: sp.medbiochemcu@gmail.com

Impact statement

In this study, the bacterial microbiota in the upper respiratory tract (URT) of Thai patients with influenza or COVID-19 were evaluated. The results revealed the different bacterial community in the URT between patients with influenza and COVID-19. The common opportunistic pathogens were found in both influenza and COVID-19 patients. Thus, different respiratory viral infections may trigger dysbiosis of the commensal microbiota, which might be associated with more susceptibility to secondary infections and disease progression. However, the change in URT microbiome among the influenza viruses and COVID-19 infected patients has not been clearly investigated. Thus, these findings would be useful for better understanding of the interaction between microbiota and viral infectious diseases.

Abstract

The upper respiratory tract is inhabited by diverse range of commensal microbiota which plays a role in protecting the mucosal surface from pathogens. Alterations of the bacterial community from respiratory viral infections could increase the susceptibility to secondary infections and disease severities. We compared the upper respiratory bacterial profiles among Thai patients with influenza or COVID-19 by using 16S rDNA high-throughput sequencing based on MiSeq platform. The Chao1 richness was not significantly different among groups, whereas the Shannon diversity of Flu A and Flu B groups were significantly lower than Non-Flu & COVID-19 group. The beta diversity revealed that the microbial communities of influenza (Flu A and Flu B), COVID-19, and Non-Flu & COVID-19 were significantly different; however, the comparison of the community structure was similar between Flu A and Flu B groups. The bacterial classification revealed that Enterobacteriaceae was predominant in influenza patients, while Staphylococcus and Pseudomonas were significantly enriched in the COVID-19 patients. These implied that respiratory viral infections might be related to alteration of upper respiratory bacterial community and susceptibility to secondary bacterial infections. Moreover, the bacteria that observed in Non-Flu & COVID-19 patients had high abundance of Streptococcus, Prevotella,

Veillonella, and Fusobacterium. This study provides the basic knowledge for further investigation of the relationship between upper respiratory microbiota and respiratory disease which might be useful for better understanding the mechanism of viral infectious diseases.

Keywords: Bacterial microbiota, upper respiratory tract, COVID-19, influenza, 16S rDNA, high-throughput sequencing

Introduction

The upper respiratory tract (URT) is inhabited by a diverse range of commensal microbes, contributing beneficial roles in protecting the mucosal surface from pathogens. However, the microbiome could be altered, or undergo dysbiosis, by pathogenic respiratory viruses, resulting in a higher susceptibility to infection and respiratory diseases (147).

The re-emergence of respiratory viruses such as seasonal influenza viruses has a considerable impact on global public health. Seasonal influenza outbreaks are mainly caused by influenza A (H1N1 and H3N2 subtypes) and influenza B (Victoria and Yamagata lineages). Currently, the emergence and rapid transmission of a novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causing Coronavirus Disease 2019 (COVID-19) has become a serious public health problem and also causes significant economic loss. The common symptoms of influenza and COVID-19 are also known as influenza-like illness (ILI) (148). Several studies have reported that healthy individuals and patients with respiratory virus infection have different upper respiratory tract microbiota (6, 149). However, the comparison of microbiome alteration caused by different virus infections has not been clearly investigated, especially between influenza and COVID-19 which cause similar ILI symptoms.

The data obtained from previous studies has demonstrated that the differences in respiratory bacterial communities associated with the respiratory infection were dependent on several factors, including environmental factors (geography, pollution, and hygiene), the host factors (race, age, gender, diet, lifestyle, and antibiotic usage) and the microbe-host immune response (55). Comparison of the bacterial profiles obtained from different studies might be difficult due to the several confounding factors and incompatible techniques used in each study. To reduce these confounding factors, this study aimed to compare the bacterial profiles in the upper respiratory tract of Thai patients with influenza or COVID-19 by using 16S high-throughput sequencing based on the MiSeq platform.

Materials and Methods

Study population and setting

A retrospective observational study was carried out to identify the bacterial profile in the upper respiratory tract of patients with influenza viruses or SARS-CoV-2 virus. The study protocols were reviewed and approved by the Institutional Review Board (IRB No. 337/57 for the influenza patient's cohort study and IRB No. 302/63 for the COVID-19 patient's cohort study) and Institutional Biosafety Committee (MDCU-IBC No. 001/2018) of the faculty of Medicine, Chulalongkorn University.

Sample collection

All clinical samples used in this study were routinely diagnosed for influenza and SARS-CoV-2 from patients with influenza-like illness (ILI) symptoms. The ILI was defined as fever (>38°C) and with respiratory symptoms (cough, runny nose, nasal congestion, and sore throat). Nasopharyngeal swab (NP) samples of the two cohorts were collected in a viral transport medium (VTM) (MP Biomedicals, USA) at 4°C and then transported to laboratory within 2 hours. First, the influenza and Non-Flu & COVID-19 cohort samples were obtained during 2017-2018 (before COVID-19 pandemic) from Bangpakok 9 International Hospital and Chum Phae Hospital. The samples were stored at -80°C freezer until tested. The Non-Flu & COVID-19 samples were collected from ILI patients and were negative with influenza viruses and SARS-CoV-2 based on RT-qPCR. Second, the samples with COVID-19 suspected cases were collected from the Institute for Urban Disease Control and Prevention (IUDC), Thailand during the pandemic of the SARS-CoV-2 virus in 2020. Samples from both cohorts were processed at the same time of the experiment.

Nucleic acid extraction, reverse transcription, and virus detection

Nucleic acids were extracted from 200 µL of NP swab samples using the magLEAD 12gC instrument with a magLEAD Consumable Kit (Precision System Science, Japan) according to the manufacturer's instructions. The extracted RNA was converted into cDNA with the random hexamer primer by using the RevertAid™ First Strand cDNA Synthesis Kit (Thermo Scientific, USA) following the manufacturer's

instructions. Detection and subtyping of influenza viruses were performed by using a multiplex real-time reverse-transcription polymerase chain reaction as described previously (101, 150) . The SARS-CoV-2 viral detection was carried out by real-time PCR with Allplex™ 2019-nCoV Assay (Seegene, Korea) following the manufacturer's instructions (151).

16S amplification and high-throughput sequencing

The DNA library was constructed by amplification of the 16S rDNA within V4 region using *Taq* DNA Polymerase (New England Biolabs, USA) according to the manufacturer's instructions. Primers, master mix components, and thermo cycles followed the procedure of previous studies (152, 153). The DNA libraries (~ 430 bp) were subsequently purified by the QIAquick PCR Purification kit (Qiagen, Germany) and quantified by KAPA Library Quantification Kits for Illumina platforms (Kapa Biosystems, USA). The DNA libraries from all samples were pooled with equal concentration. The pooled libraries were paired-end (2×250 cycles) sequenced using the Illumina MiSeq platform (Illumina, USA).

Data analysis

The raw sequence datasets were de-multiplexed by MiSeq reporter software (version 2.6.2.3), and the paired-end FASTQ files were processed through the QIIME2 pipeline (version 2019.7) (154). The paired-end reads were quality filtered, merged, and denoised by DADA2 (155) plugin implemented in QIIME2. The construction of the Amplicon sequence variant (ASV) and chimeric sequence filtering were subsequently performed based on DADA2 algorithm. Then, the ASVs were classified based on the 16S Greengene database (156) by using the VSEARCH algorithm. The alpha (Shannon and Chao1) and beta (weighted UniFrac distance) diversities were calculated using a plugin implemented in QIIME2 (157). Linear discriminant analysis Effect Size (LEfSe) was used to identify different taxa among groups (158). Statistical analyses were performed with the Wilcoxon matched pairs test using GraphPad Prism version 6.01 software.

Results

In this study, the influenza A, influenza B, and SARS-CoV-2 viruses were detected by real-time PCR and then the samples were divided into four groups, including the patients infected with influenza A virus (Flu A, N=24), influenza B virus (Flu B, N=24), SARS-CoV-2 virus (COVID-19, N= 24), and the patients without the influenza viruses and SARS-CoV-2 virus (Non-Flu & COVID-19, N= 24). The age of patients among groups was not statistically different (*p*-value= 0.065). The high-throughput sequencing data are shown in Table 4. Briefly, the total reads and pass-filtered reads obtained from each group were sufficient for ASVs classification as demonstrated in the rarefaction curve (Supplementary Figure 1). Moreover, the pass-filtered reads of all groups were efficiently classified as bacterial taxa (more than 98%), indicating that the sequencing data obtained in this study were suitable for 16S rRNA amplicon-based analysis.

Table 4. The subject characteristics and sequencing summary of each group.

Parameters	Flu A (N=24)	Flu B (N=24)	COVID-19 (N=24)	Non-flu & COVID-19 (N=24)
Age	25.96±20.28	26.50±19.30	36.09±12.17	34.46±9.34
Total reads	79,425±13,712	48,227±11,358	79,504±13,480	53,092±14,065
Pass-filtered reads	70,290±13,712	41,811±9,514	69,663±11,450	44,037±12,197
% Classified	98.58±3.06	99.74±0.73	99.96±0.14	99.86±0.21
Chao1	22.46±9.38	24.21±8.06	24.21±9.86	23.13±8.17
Shannon	1.26±0.83	1.28±0.55	2.08±.13	2.91±0.93

The alpha diversity based on Chao1 analysis revealed that bacterial richness was not different among groups as shown in Figure 16(a). On the other hand, the Shannon diversity of Flu A and Flu B groups were significantly lower than the Non-Flu & COVID-19 group (*p*-value < 0.05). Moreover, Flu A group was significantly decreased compared to COVID-19 group (*p*-value < 0.05) as demonstrated in Figure 16(b). The beta diversity was analyzed by weighted UniFrac distance to compare the microbial community distance among groups. The result revealed that microbial communities of influenza (Flu A and Flu B) groups were highly similar, but they were significantly distinct from those of the COVID-19 group and Non-Flu & COVID-19 group (PERMANOVA analysis, *p*-value <0.05) as illustrated in Figure 16(c).

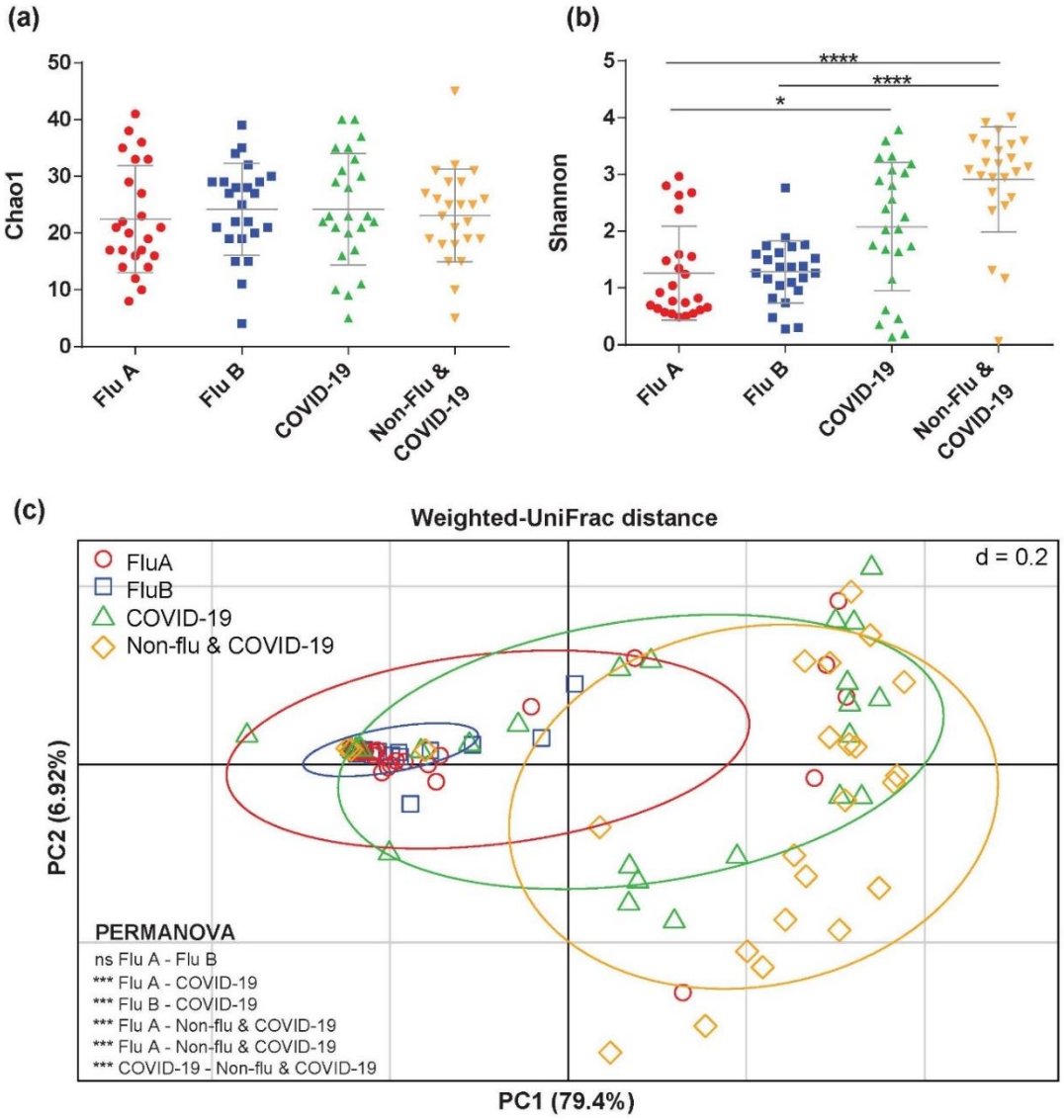


Figure 16. The comparison of the bacterial diversities in Flu A, Flu B, COVID-19, and Non-Flu & COVID-19 patients.The bars indicate average diversity of **(a)** Chao1 richness and **(b)** Shannon diversity index with standard deviation and significant differences among 4 groups. **(c)** Principal-coordinate analysis (PCoA) of the specific microbial communities weighted UniFrac distance colored by sample groups shows the significant difference between groups (tested by PERMANOVA analysis with *p*-value < 0.05). * *p*-value< 0.05, *** *p*-value < 0.001, **** *p*-value < 0.0001 and ns (not significant).

The taxonomic classification demonstrated that the relative abundances of upper respiratory bacteria were different among groups (Figure 17). The influenza (Flu A and Flu B) groups were highly predominated Proteobacteria, while bacterial profile in COVID-19 group was mostly dominated with Firmicutes and Bacteroidetes. Moreover, the Non-Flu & COVID-19 group was enriched by Firmicutes, Bacteroidetes, and Proteobacteria (Figure 17(a)).

Comparing the relative abundance of bacterial composition at the genus level showed that Enterobacteriaceae had the highest abundance in influenza groups (Flu A and Flu B). Nevertheless, *Streptococcus* was the most common bacterial composition in COVID-19 and Non-Flu & COVID-19 groups. The differences in bacterial compositions between COVID-19 and Non-Flu & COVID-19 groups were also observed. Enterobacteriaceae, *Staphylococcus*, *Lautropia*, *Pseudomonas*, and *Corynebacterium* were increased in the COVID-19 group, whereas *Prevotella*, *Veillonella*, *Capnocytophaga*, and *Fusobacterium* were slightly increased in the Non-Flu & COVID-19 group (Figure 17(b)).

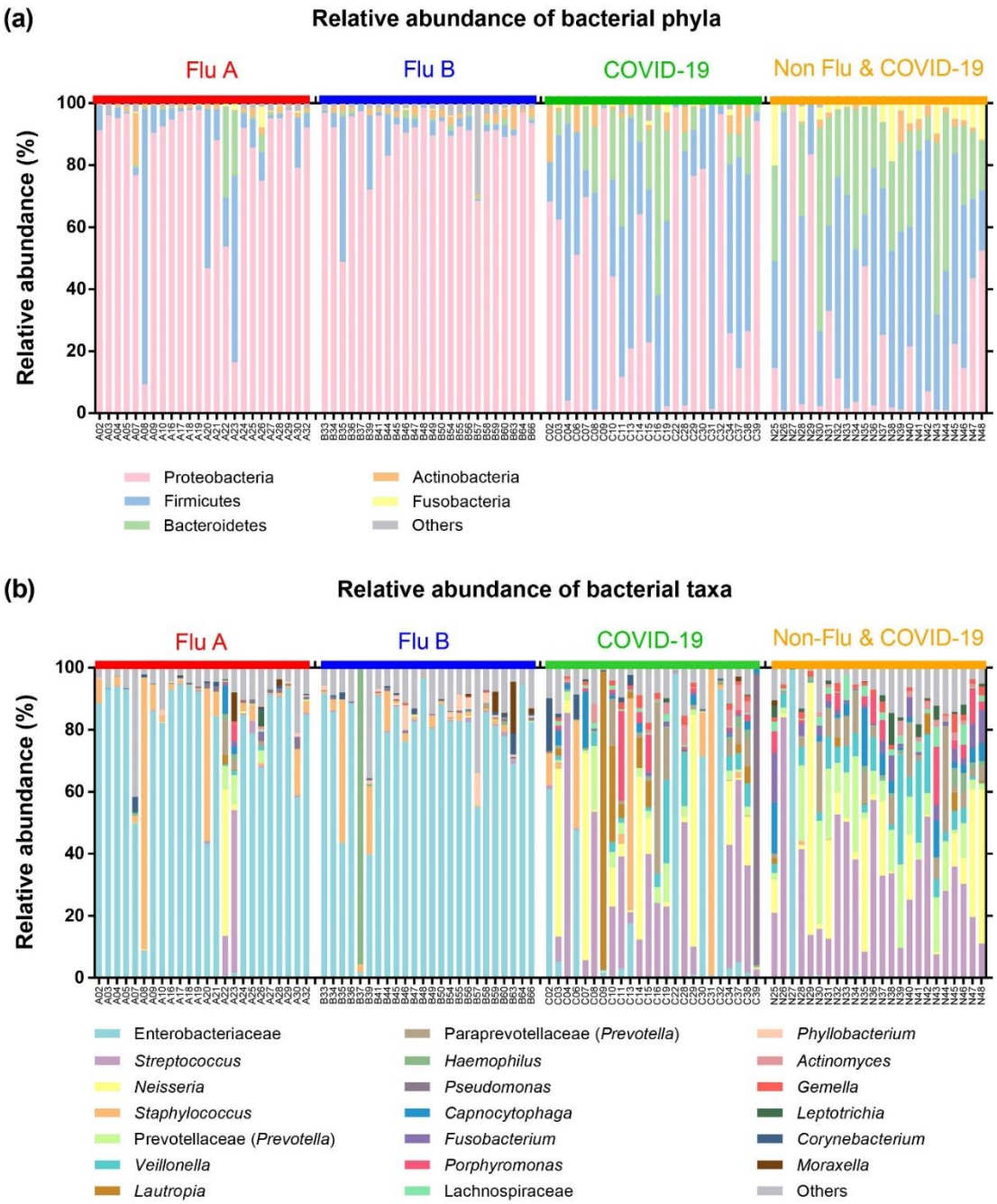


Figure 17. The bar plots illustrate the relative abundance of bacterial phyla (a) and genera (b) in each group (Flu A, Flu B, COVID-19, and Non-Flu & COVID-19) as analyzed by QIIME2 pipeline.

The differentially enriched bacteria among influenza, COVID-19 and Non-Flu & COVID-19 groups were analyzed using the Linear discriminant analysis Effect Size (LEfSe) method with LDA score > 4 and *p*-value <0.05, as shown in Supplementary Figure 2. Subsequently, the characteristics of significant bacterial composition were selected from LEfSe and relative abundance, as demonstrated in Table 5. The results revealed that Enterobacteriaceae was significantly enriched in the influenza group, which contributed approximately 74.51%. *Staphylococcus* and *Pseudomonas* genera were significantly dominant in the COVID-19 group, which accounted for 9.83% and 4.01%, respectively. Lastly, the Non-Flu & COVID-19 group showed significant enrichment of *Streptococcus*, *Prevotella*, *Veillonella* and *Fusobacterium,* which accounted for 29.85%, 11.28%, 7.21% and 3.42%, respectively.

Table 5. The characteristics of significantly enriched bacteria in each group

Group	Bacterial composition	Gram strain	Commonly found	Previous reports
Influenza	Enterobacteriaceae (74.51%)	negative	Soil, water and GI tract	(159)
COVID-19	*Staphylococcus* (9.83%)	positive	Skin, nasal vestibule and hospital	(160)
	Pseudomonas (4.01%)	negative	Soil, water and vegetation	(161)
Non-Flu& COVID-19	*Streptococcus* (29.85%)	positive	Throat, skin, ear and sinus	(162, 163)
	Prevotella (11.28%)	negative	Human intestine, the rumen of cattle and sheep	(163, 164)
	Veillonella (7.21%)	negative	Mammal and human intestines	(164)
	Fusobacterium (3.42%)	negative	Upper respiratory and GI tracts	(165)

Note: the parenthesis behind the family or genus indicated percentage of the relative abundance belonged to the family or genus

Discussion

The respiratory microbiota has been investigated in several previous studies. The bacterial profiles associated with the healthy respiratory status included *Prevotella*, *Veillonella*, *Streptococcus*, *Staphylococcus*, and *Moraxella*. On the other hand, dysbiotic microbiota related with lower microbial diversity and increased amount of specific bacteria such as *Streptococcus*, *Pseudomonas*, and *Neisseria* in patients with respiratory virus infections (5).

This study explored the bacterial compositions in the URT of Thai patients infected with influenza viruses, SARS-CoV-2 virus, and the patients without either influenza viruses or SARS-CoV-2 virus. The data obtained from the present study showed that different viral infection (influenza or SARS-CoV-2) might be related to the alteration in bacterial diversity and community composition of the URT. From our observation, the most abundant bacterial phylum in the influenza group was Proteobacteria, while the COVID-19 and Non-Flu & COVID-19 groups were dominated by Firmicutes. The results correlated with previous studies, in which Proteobacteria and Firmicutes were classified as the normal flora in URT (166).

For other lower respiratory viruses such as respiratory syncytial virus (RSV) and rhinoviruses, previous study revealed that the Proteobacteria was highly dominated in the infected patients. Indeed, the predominated genera were belonged to the *Moraxella* and *Streptococcus* (167). In our study also found the *Streptococcus* in which was highly contributed in the COVID-19 and Non-flu & COVID-19 groups. Moreover, the abundance of *Acinetobacter* and *Klebsiella* genera in Proteobacteria phylum were higher in the fatal cases than those in nonfatal cases of the MERS-CoV infected patients in Saudi Arabia (168).

Our finding shown that the Enterobacteriaceae family was predominant in influenza infected patients, which is similar to the previous study of Chinese patients with H1N1 influenza A virus infection resulting in hypothermia and pneumonia (159). The *Staphylococcus* genus is significantly enriched in the Thai patients with COVID-19 in the present study, and this corresponds to the recent finding that *Staphylococcus* was significantly increased in the Chinese patients with severe COVID-19 illness (160).

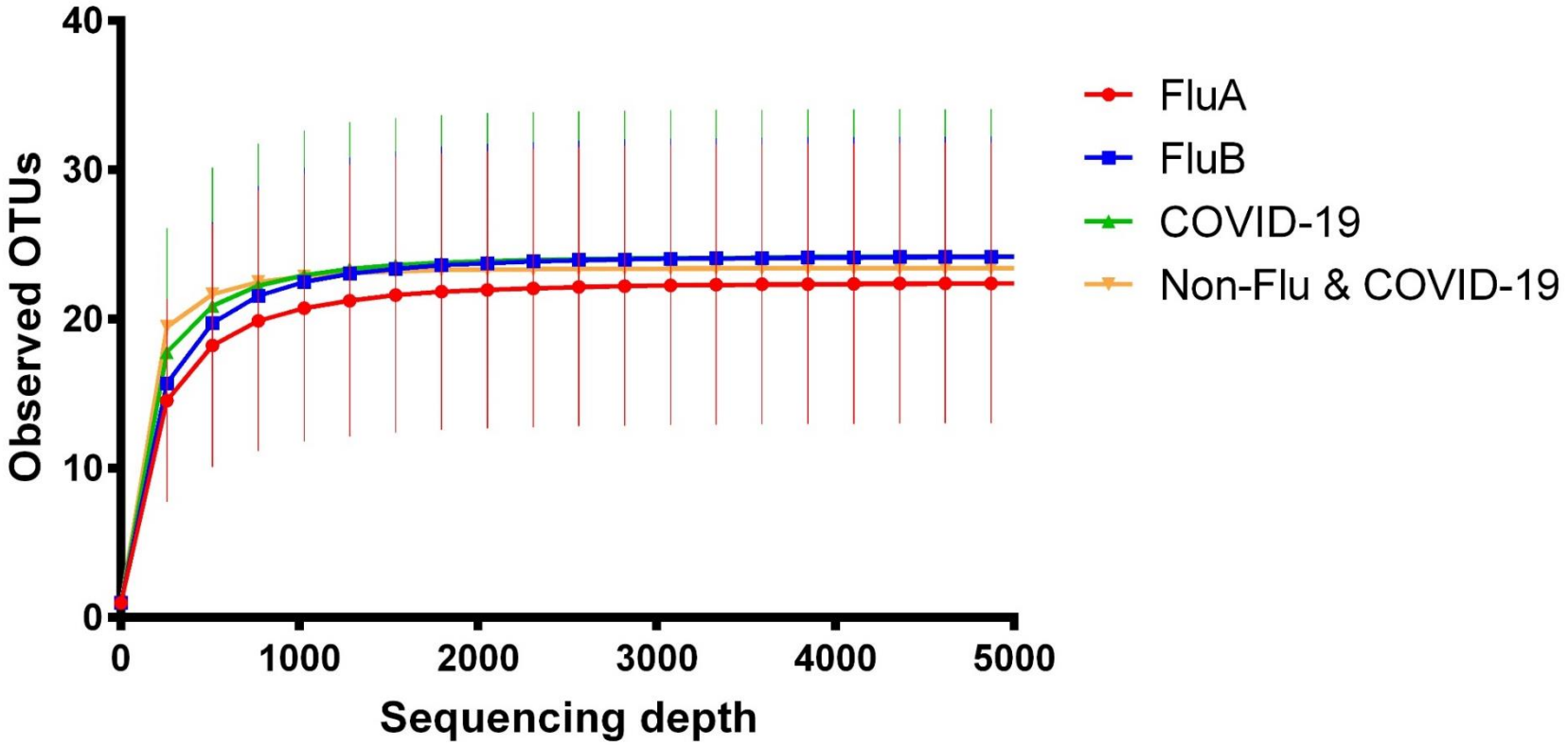
Staphylococcus is considered to be one of the causative agents in hospital-acquired infections (169) Moreover, the secondary infection with *Staphylococcus aureus* bacteremia was associated with high mortality rates in hospitalized COVID-19 (170). The *Pseudomonas* genus was also found in COVID-19 patients of our study. This is similar to the previous report which found that *Pseudomonas* was significantly increased in both the upper respiratory tract and the gut of children with COVID-19 infection (161). Interestingly, viral infections such as influenza viruses or SARS-CoV-2 might modulate the bacterial communities which potentially favor the expansion of opportunistic pathogens. These pathogens can cause hospital-acquired infections involving multiple antibiotic resistances, associated with pneumonia and high mortality rates. These data suggested that the bacterial composition of the URT can impact host susceptibility and disease outcomes after respiratory infection.

The previous studies showed that microbiota in healthy individuals were enriched with commensal bacteria in the nasopharynx, such as *Streptococcus* and *Fusobacterium,* while in the oropharyngeal cavity, *Prevotella* and *Veillonella* predominated (162-165). Our data revealed that *Streptococcus*, *Prevotella*, *Veillonella*, and *Fusobacterium* were significantly dominant in the patients without the influenza viruses and SARS-CoV-2 infections, indicating that bacterial communities in these patients might be closely related to healthy individuals.

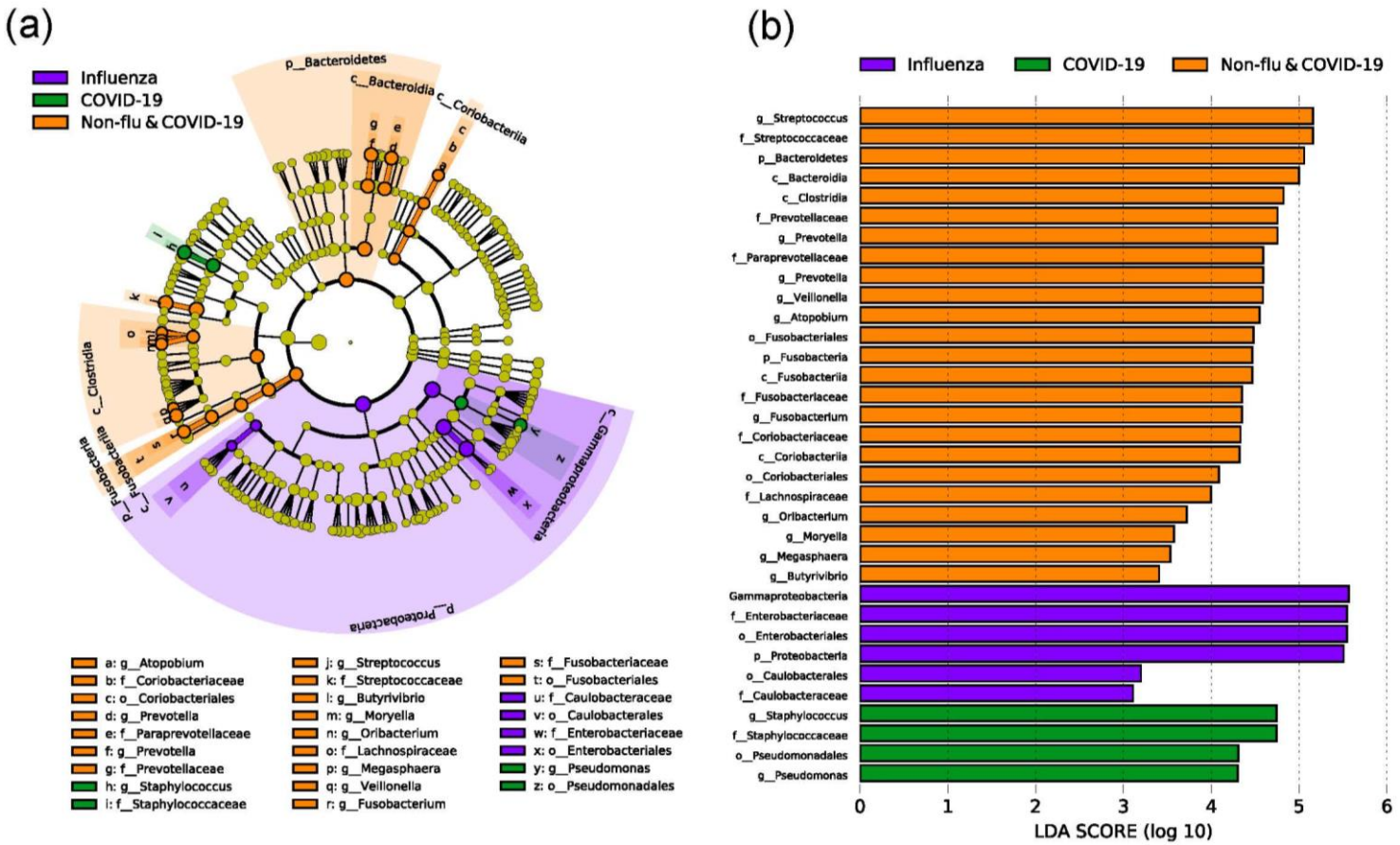
Conclusion

In summary, this study provides a comparison of the upper respiratory bacterial communities of Thai patients with either influenza or COVID-19. The common opportunistic pathogens such as family Enterobacteriaceae was found in patients with influenza illness, whereas *Staphylococcus* and *Pseudomonas* were observed in patients with COVID-19, indicating that respiratory viral infections might be associated with an altered upper respiratory bacterial community and a differential induction of secondary bacterial infections. However, there are several limitations to this study, including small sample sizes, unmatched ages/genders, unknown severity/symptoms and different periods of samples collection among groups. Therefore, further studies with larger cohorts and matched demographic data

should be performed to confirm the bacterial communities which are associated with the respiratory viral infections. Finally, the findings of this study provide the basic knowledge for further investigation of the relationship between upper respiratory microbiota and respiratory disease.



Supplementary Figure 1. The rarefaction curve of upper respiratory bacterial profile in 4 groups.



Supplementary Figure 2. Linear discriminant analysis Effect Size (LEfSe) analysis of bacterial profiles in 3 groups. **(a)** A cladogram of the differentially enriched taxa (LDA scores > 4 and p-value<0.05). **(b)** The bar graph shows the LDA scores of significantly enriched taxa (influenza = purple, COVID-19 = green, and Non-Flu & COVID-19 = orange).

Chapter 4
CONCLUSION

Conclusion, limitations, and suggestions

In this study, NGS technology can be applied to analyze whole-genome sequencing of influenza A viruses and URT bacterial microbiota profiling of influenza patients in Thailand. In our first study, the high-throughput data demonstrated that 17 nasal swab samples from ILI patients were successfully sequenced and characterized the whole genome of influenza A (A/H1N1 and A/H3N2) viruses. The phylogenetic analysis based on the deduced amino acid sequences revealed that the recommended vaccine (A/H1N1) strain might be less effective, whereas the recommended vaccine (A/H3N2) was more effective against the circulating influenza viruses in Thailand during 2017-2018. Furthermore, several amino acid substitutions across eight segmented genes, especially involved in the antigenic site of *HA* and *NA* genes of both A/H1N1 and A/H3N2 viruses. Lastly, the nucleotide diversity showed negative selection in the *PB1*, *PA*, *HA*, and *NA* genes of A/H1N1 viruses. Ideally, direct sequencing from clinical samples might be represented the influenza virus mutation, however low viral titers and also more inhibitors within the sample; thus numbers of successfully processed samples were limited. In this study, influenza isolation was performed in MDCK cells, following the WHO standard method for propagating influenza viruses. The percentage of successful isolation was approximately 47.78%, which higher than the previous study (114). In the genome amplification step, full-length of all influenza eight-segments were amplified using universal influenza primers. However, the longest segments of influenza genes (PB2, PB1 and PA genes) are quite difficult to perform the amplification. Therefore, specific primers might improve the success of the longest segments (polymerase genes) amplification.

Moreover, several mutations found in this study were novel mutations that need further functional prediction by in silico analysis and confirmation by reverse genetics. Therefore, high-throughput data obtained from this study might be useful

for mutation analysis and can be compared with data obtained from other studies in the future.

The second study compares bacterial microbiota profiling in the upper respiratory among influenza (Flu A or Flu B), COVID-19, and Non-Flu & COVID-19 (ILI patients with negative influenza viruses and SARS-CoV-2 virus) groups. High-throughput 16S rDNA sequencing data were analyzed for the alpha diversity, beta diversity, relative abundances, and LEfSe. In alpha diversity, the Shannon diversity was significantly lower in influenza (Flu A and Flu B) groups compared to Non-Flu & COVID-19 group. The beta diversity showed that bacterial compositions were significantly different among influenza, COVID-19, and Non-Flu & COVID-19 groups, whereas similar between Flu A and Flu B groups. The common opportunistic pathogens were observed in both influenza (Enterobacteriaceae family) and COVID-19 groups (*Staphylococcus* and *Pseudomonas*), which implied that alteration of bacterial communities in the upper respiratory tract might be associated with viral infections. Furthermore, commensal bacteria in URT, such as *Streptococcus*, *Prevotella*, *Veillonella*, and *Fusobacterium* were highly abundant in the Non-Flu & COVID-19 group. Nevertheless, this study was limited in the number of samples, healthy control group, unmatched ages/genders, unknown severity, and different periods of samples collection among groups. Another confounding factor is unknown antibiotic treatment record, which could affect the URT bacterial microbiota. Moreover, a short read from Illumina sequencing has been limited to identified bacterial classification at the genus level, which can be improved by long-read sequencing. Although, alteration in URT bacterial microbiota between influenza and non-influenza patients has not been clearly examined as related to other viral infections. Our data provide fundamental knowledge to investigate the association host-microbe interaction that might be useful for predicting health status and applied for microbiome engineering to enhance immunity against pathogen infection in the future.

REFERENCES

1. Simmerman JM, Chittaganpitch M, Levy J, Chantra S, Maloney S, Uyeki T, et al. Incidence, seasonality and mortality associated with influenza pneumonia in Thailand: 2005-2008. PLoS One. 2009;4(11):e7776.

2. Shao W, Li X, Goraya MU, Wang S, Chen JL. Evolution of Influenza A Virus by Mutation and Re-Assortment. Int J Mol Sci. 2017;18(8):1650.

3. Cox NJ, Subbarao K. Global epidemiology of influenza: past and present. Annu Rev Med. 2000;51:407-21.

4. Pauly MD, Procario MC, Lauring AS. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. Elife. 2017;6:e26437.

5. Porto BN, Moraes TJ. The triad: respiratory microbiome - virus - immune response in the pathophysiology of pulmonary viral infections. Expert Rev Respir Med. 2021;15(5):635-48.

6. Kaul D, Rathnasinghe R, Ferres M, Tan GS, Barrera A, Pickett BE, et al. Microbiome disturbance and resilience dynamics of the upper respiratory tract during influenza A virus infection. Nature communications. 2020;11(1):1-12.

7. Rasmussen SA, Jamieson DJ, Macfarlane K, Cragan JD, Williams J, Henderson Z. Pandemic influenza and pregnant women: summary of a meeting of experts. Am J Public Health. 2009;99 Suppl 2(Suppl 2):S248-54.

8. Moghadami M. A Narrative Review of Influenza: A Seasonal and Pandemic Disease. Iran J Med Sci. 2017;42(1):2-13.

9. Neidich SD, Green WD, Rebeles J, Karlsson EA, Schultz-Cherry S, Noah TL, et al. Increased risk of influenza among vaccinated adults who are obese. Int J Obes (Lond). 2017;41(9):1324-30.

10. Bouvier NM, Palese P. The biology of influenza viruses. Vaccine. 2008;26 Suppl 4(Suppl 4):D49-53.

11. Taubenberger JK, Kash JC. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe. 2010;7(6):440-51.

12. Ohishi K, Ninomiya A, Kida H, Park CH, Maruyama T, Arai T, et al. Serological

evidence of transmission of human influenza A and B viruses to Caspian seals (Phoca caspica). Microbiol Immunol. 2002;46(9):639-44.

13. Ran Z, Shen H, Lang Y, Kolb EA, Turan N, Zhu L, et al. Domestic pigs are susceptible to infection with influenza B viruses. J Virol. 2015;89(9):4818-26.

14. Ito T, Goto H, Yamamoto E, Tanaka H, Takeuchi M, Kuwayama M, et al. Generation of a highly pathogenic avian influenza A virus from an avirulent field isolate by passaging in chickens. J Virol. 2001;75(9):4439-43.

15. Steinhauer DA, Holland JJ. Rapid evolution of RNA viruses. Annu Rev Microbiol. 1987;41:409-33.

16. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. Microbiol Rev. 1992;56(1):152-79.

17. Simonsen L, Viboud C, Grenfell BT, Dushoff J, Jennings L, Smit M, et al. The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. Mol Biol Evol. 2007;24(8):1811-20.

18. Kawaoka Y, Neumann G. Influenza viruses: an introduction. Methods Mol Biol. 2012;865:1-9.

19. Schrauwen EJ, Fouchier RA. Host adaptation and transmission of influenza A viruses in mammals. Emerg Microbes Infect. 2014;3(2):e9.

20. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DA. Incubation periods of acute respiratory viral infections: a systematic review. Lancet Infect Dis. 2009;9(5):291-300.

21. Harper SA, Bradley JS, Englund JA, File TM, Gravenstein S, Hayden FG, et al. Seasonal influenza in adults and children--diagnosis, treatment, chemoprophylaxis, and institutional outbreak management: clinical practice guidelines of the Infectious Diseases Society of America. Clin Infect Dis. 2009;48(8):1003-32.

22. Jayasundara K, Soobiah C, Thommes E, Tricco AC, Chit A. Natural attack rate of influenza in unvaccinated children and adults: a meta-regression analysis. BMC Infect Dis. 2014;14:670.

23. Scholtissek C, Rohde W, Von Hoyningen V, Rott R. On the origin of the human influenza virus subtypes H2N2 and H3N2. Virology. 1978;87(1):13-20.

24. Kawaoka Y, Krauss S, Webster RG. Avian-to-human transmission of the PB1 gene

of influenza A viruses in the 1957 and 1968 pandemics. J Virol. 1989;63(11):4603-8.

25. Taubenberger JK, Morens DM. Influenza: the once and future pandemic. Public Health Rep. 2010;125 Suppl 3(Suppl 3):16-26.

26. Nicoll A. A new decade, a new seasonal influenza: the Council of the European Union Recommendation on seasonal influenza vaccination. Euro Surveill. 2010;15(1).

27. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459(7250):1122-5.

28. Ambrose CS, Levin MJ. The rationale for quadrivalent influenza vaccines. Hum Vaccin Immunother. 2012;8(1):81-8.

29. Vijaykrishna D, Holmes EC, Joseph U, Fourment M, Su YC, Halpin R, et al. The contrasting phylodynamics of human influenza B viruses. Elife. 2015;4:e05055.

30. Paul Glezen W, Schmier JK, Kuehn CM, Ryan KJ, Oxford J. The burden of influenza B: a structured literature review. Am J Public Health. 2013;103(3):e43-51.

31. Tan Y, Guan W, Lam TT, Pan S, Wu S, Zhan Y, et al. Differing epidemiological dynamics of influenza B virus lineages in Guangzhou, southern China, 2009-2010. J Virol. 2013;87(22):12447-56.

32. Tong S, Li Y, Rivailler P, Conrardy C, Castillo DA, Chen LM, et al. A distinct lineage of influenza A virus from bats. Proc Natl Acad Sci U S A. 2012;109(11):4269-74.

33. Cheung TK, Poon LL. Biology of influenza a virus. Ann N Y Acad Sci. 2007;1102:1-25.

34. Dou D, Revol R, Östbye H, Wang H, Daniels R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. Front Immunol. 2018;9:1581.

35. Rioux M, McNeil M, Francis ME, Dawe N, Foley M, Langley JM, et al. The Power of First Impressions: Can Influenza Imprinting during Infancy Inform Vaccine Design? Vaccines (Basel). 2020;8(3).

36. Krug RM, Yuan W, Noah DL, Latham AG. Intracellular warfare between human influenza viruses and human cells: the roles of the viral NS1 protein. Virology. 2003;309(2):181-9.

37. O'Neill RE, Talon J, Palese P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. Embo j. 1998;17(1):288-96.

38. Vasin AV, Temkina OA, Egorov VV, Klotchenko SA, Plotnikova MA, Kiselev OI. Molecular mechanisms enhancing the proteome of influenza A viruses: an overview of recently discovered proteins. Virus Res. 2014;185:53-63.

39. van de Sandt CE, Bodewes R, Rimmelzwaan GF, de Vries RD. Influenza B viruses: not to be discounted. Future Microbiol. 2015;10(9):1447-65.

40. de Graaf M, Fouchier RA. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. Embo j. 2014;33(8):823-41.

41. Rust MJ, Lakadamyali M, Zhang F, Zhuang X. Assembly of endocytic machinery around individual influenza viruses during viral entry. Nat Struct Mol Biol. 2004;11(6):567-73.

42. Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem. 2000;69:531-69.

43. Tripathi S, Batra J, Lal SK. Interplay between influenza A virus and host factors: targets for antiviral intervention. Arch Virol. 2015;160(8):1877-91.

44. Hutchinson EC, Fodor E. Transport of the influenza virus genome from nucleus to nucleus. Viruses. 2013;5(10):2424-46.

45. Long JS, Mistry B, Haslam SM, Barclay WS. Host and viral determinants of influenza A virus species specificity. Nat Rev Microbiol. 2019;17(2):67-81.

46. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. Microbiome. 2015;3:31.

47. Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol. 2011;9(4):244-53.

48. Ding T, Schloss PD. Dynamics and associations of microbial community types across the human body. Nature. 2014;509(7500):357-60.

49. Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, Sinha R, et al. Disordered microbial communities in the upper respiratory tract of cigarette smokers. PLoS One. 2010;5(12):e15216.

50. Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. Nat Rev Immunol. 2014;14(12):827-35.

51. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. Nature. 2007;449(7164):804-10.

52. Olszak T, An D, Zeissig S, Vera MP, Richter J, Franke A, et al. Microbial exposure

during early life has persistent effects on natural killer T cell function. Science. 2012;336(6080):489-93.

53. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.

54. Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. Dysbiosis and the immune system. Nat Rev Immunol. 2017;17(4):219-32.

55. Gupta VK, Paul S, Dutta C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. Front Microbiol. 2017;8:1162.

56. Li N, Ma WT, Pang M, Fan QL, Hua JL. The Commensal Microbiota and Viral Infection: A Comprehensive Review. Front Immunol. 2019;10:1551.

57. Lieleg O, Lieleg C, Bloom J, Buck CB, Ribbeck K. Mucin biopolymers as broad-spectrum antiviral agents. Biomacromolecules. 2012;13(6):1724-32.

58. Schroeder BO. Fight them or feed them: how the intestinal mucus layer manages the gut microbiota. Gastroenterol Rep (Oxf). 2019;7(1):3-12.

59. Tabata T, Petitt M, Puerta-Guardo H, Michlmayr D, Wang C, Fang-Hoover J, et al. Zika Virus Targets Different Primary Human Placental Cells, Suggesting Two Routes for Vertical Transmission. Cell Host Microbe. 2016;20(2):155-66.

60. Botić T, Klingberg TD, Weingartl H, Cencic A. A novel eukaryotic cell culture model to study antiviral activity of potential probiotic bacteria. Int J Food Microbiol. 2007;115(2):227-34.

61. Jounai K, Ikado K, Sugimura T, Ano Y, Braun J, Fujiwara D. Spherical lactic acid bacteria activate plasmacytoid dendritic cells immunomodulatory function via TLR9-dependent crosstalk with myeloid dendritic cells. PLoS One. 2012;7(4):e32588.

62. Wang Y, Liu L, Moore DJ, Shen X, Peek RM, Acra SA, et al. An LGG-derived protein promotes IgA production through upregulation of APRIL expression in intestinal epithelial cells. Mucosal Immunol. 2017;10(2):373-84.

63. Harper A, Vijayakumar V, Ouwehand AC, Ter Haar J, Obis D, Espadaler J, et al. Viral Infections, the Microbiome, and Probiotics. Front Cell Infect Microbiol. 2020;10:596166.

64. Bosch A, Levin E, van Houten MA, Hasrat R, Kalkman G, Biesbroek G, et al. Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. EBioMedicine. 2016;9:336-45.

65. Bogaert D, Keijser B, Huse S, Rossen J, Veenhoven R, van Gils E, et al. Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. PLoS One. 2011;6(2):e17035.

66. Charlson ES, Bittinger K, Chen J, Diamond JM, Li H, Collman RG, et al. Assessing bacterial populations in the lung by replicate analysis of samples from the upper and lower respiratory tracts. PLoS One. 2012;7(9):e42786.

67. de Steenhuijsen Piters WAA, Binkowska J, Bogaert D. Early Life Microbiota and Respiratory Tract Infections. Cell Host Microbe. 2020;28(2):223-32.

68. Belkacem N, Serafini N, Wheeler R, Derrien M, Boucinha L, Couesnon A, et al. Lactobacillus paracasei feeding improves immune control of influenza infection in mice. PLoS One. 2017;12(9):e0184976.

69. Wang J, Li F, Sun R, Gao X, Wei H, Li LJ, et al. Bacterial colonization dampens influenza-mediated acute lung injury via induction of M2 alveolar macrophages. Nat Commun. 2013;4:2106.

70. Ramsey MM, Freire MO, Gabrilska RA, Rumbaugh KP, Lemon KP. Staphylococcus aureus Shifts toward Commensalism in Response to Corynebacterium Species. Front Microbiol. 2016;7:1230.

71. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17(6):333-51.

72. Patterson JD, Wongsurawat T, Rodriguez A. A Glioblastoma Genomics Primer for Clinicians. Med Res Arch. 2020;8(2).

73. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Mol Cell. 2015;58(4):586-97.

74. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012;2012:251364.

75. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data - from vision to reality. Euro Surveill. 2017;22(13).

76. McNaughton AL, D'Arienzo V, Ansari MA, Lumley SF, Littlejohn M, Revill P, et al.

Insights From Deep Sequencing of the HBV Genome-Unique, Tiny, and Misunderstood. Gastroenterology. 2019;156(2):384-99.

77. Tsiatis AC, Norris-Kirby A, Rich RG, Hafez MJ, Gocke CD, Eshleman JR, et al. Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications. J Mol Diagn. 2010;12(4):425-32.

78. Neefs JM, Van de Peer Y, De Rijk P, Chapelle S, De Wachter R. Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res. 1993;21(13):3025-49.

79. Claesson MJ, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, et al. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS One. 2009;4(8):e6669.

80. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460-1.

81. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6.

82. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537-41.

83. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature. 2019;569(7758):655-62.

84. Wu WK, Chen CC, Liu PY, Panyod S, Liao BY, Chen PC, et al. Identification of TMAO-producer phenotype and host-diet-gut dysbiosis by carnitine challenge test in human and germ-free mice. Gut. 2019;68(8):1439-49.

85. Renvoisé A, Brossier F, Sougakoff W, Jarlier V, Aubry A. Broad-range PCR: past, present, or future of bacteriology? Med Mal Infect. 2013;43(8):322-30.

86. Stöhr K. Influenza--WHO cares. Lancet Infect Dis. 2002;2(9):517.

87. World Health Organization. Up to 650000 people die of respiratory diseases

linked to seasonal flu each year 2017 [Available from: http://www.who.int/mediacentre/news/releases/2017/seasonal-flu/en/.

88. Boivin S, Cusack S, Ruigrok RW, Hart DJ. Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. J Biol Chem. 2010;285(37):28411-7.

89. Bloom JD. An experimentally determined evolutionary model dramatically improves phylogenetic fit. Mol Biol Evol. 2014;31(8):1956-78.

90. Nobusawa E, Sato K. Comparison of the mutation rates of human influenza A and B viruses. J Virol. 2006;80(7):3675-8.

91. Suarez DL. Evolution of avian influenza viruses. Vet Microbiol. 2000;74(1-2):15-27.

92. Doyle JD, Chung JR, Kim SS, Gaglani M, Raiyani C, Zimmerman RK, et al. Interim Estimates of 2018-19 Seasonal Influenza Vaccine Effectiveness - United States, February 2019. MMWR Morb Mortal Wkly Rep. 2019;68(6):135-9.

93. Rose A, Kissling E, Emborg HD, Larrauri A, McMenamin J, Pozo F, et al. Interim 2019/20 influenza vaccine effectiveness: six European studies, September 2019 to January 2020. Euro Surveill. 2020;25(10).

94. Chittaganpitch M, Supawat K, Olsen SJ, Waicharoen S, Patthamadilok S, Yingyong T, et al. Influenza viruses in Thailand: 7 years of sentinel surveillance data, 2004-2010. Influenza Other Respir Viruses. 2012;6(4):276-83.

95. Tewawong N, Vichiwattana P, Korkong S, Klinfueng S, Suntronwong N, Thongmee T, et al. Evolution of the neuraminidase gene of seasonal influenza A and B viruses in Thailand between 2010 and 2015. PLoS One. 2017;12(4):e0175655.

96. Tewawong N, Suntronwong N, Korkong S, Theamboonlers A, Vongpunsawad S, Poovorawan Y. Evidence for influenza B virus lineage shifts and reassortants circulating in Thailand in 2014-2016. Infect Genet Evol. 2017;47:35-40.

97. Tewawong N, Suntronwong N, Vichiwattana P, Vongpunsawad S, Theamboonlers A, Poovorawan Y. Genetic and antigenic characterization of hemagglutinin of influenza A/H3N2 virus from the 2015 season in Thailand. Virus Genes. 2016;52(5):711-5.

98. Suntronwong N, Klinfueng S, Vichiwattana P, Korkong S, Thongmee T, Vongpunsawad S, et al. Genetic and antigenic divergence in the influenza A(H3N2) virus

circulating between 2016 and 2017 in Thailand. PLoS One. 2017;12(12):e0189511.

99. Imai K, Tamura K, Tanigaki T, Takizawa M, Nakayama E, Taniguchi T, et al. Whole Genome Sequencing of Influenza A and B Viruses With the MinION Sequencer in the Clinical Setting: A Pilot Study. Front Microbiol. 2018;9:2748.

100. Alnaji FG, Holmes JR, Rendon G, Vera JC, Fields CJ, Martin BE, et al. Sequencing Framework for the Sensitive Detection and Precise Mapping of Defective Interfering Particle-Associated Deletions across Influenza A and B Viruses. J Virol. 2019;93(11):e00354-19.

101. Suwannakarn K, Payungporn S, Chieochansin T, Samransamruajkit R, Amonsin A, Songserm T, et al. Typing (A/B) and subtyping (H1/H3/H5) of influenza A viruses by multiplex real-time RT-PCR assays. J Virol Methods. 2008;152(1-2):25-31.

102. Tewawong N, Chansaenroj J, Klinfueng S, Vichiwattana P, Korkong S, Thongmee T, et al. Lineage-specific detection of influenza B virus using real-time polymerase chain reaction with melting curve analysis. Arch Virol. 2016;161(6):1425-35.

103. World Health Organization. WHO manual on animal influenza diagnosis and surveillance 2002 [Available from: https://apps.who.int/iris/handle/10665/68026.

104. Ilyushina NA, Ikizler MR, Kawaoka Y, Rudenko LG, Treanor JJ, Subbarao K, et al. Comparative study of influenza virus replication in MDCK cells and in primary cells derived from adenoids and airway epithelium. J Virol. 2012;86(21):11725-34.

105. Xue J, Chambers BS, Hensley SE, López CB. Propagation and Characterization of Influenza Virus Stocks That Lack High Levels of Defective Viral Genomes and Hemagglutinin Mutations. Front Microbiol. 2016;7:326.

106. Zhou B, Donnelly ME, Scholes DT, St George K, Hatta M, Kawaoka Y, et al. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. J Virol. 2009;83(19):10309-13.

107. Korsun NS, Angelova SG, Trifonova IT, Georgieva IL, Tzotcheva IS, Mileva SD, et al. Predominance of influenza B/Yamagata lineage viruses in Bulgaria during the 2017/2018 season. Epidemiol Infect. 2019;147:e76.

108. Hall T, Biosciences I, Carlsbad C. BioEdit: an important software for molecular biology. GERF Bull Biosci. 2011;2(1):60-1.

109. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary

Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35(6):1547-9.

110. Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, et al. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One. 2011;6(1):e15925.

111. Thongpan I, Suntronwong N, Vichaiwattana P, Wanlapakorn N, Vongpunsawad S, Poovorawan Y. Respiratory syncytial virus, human metapneumovirus, and influenza virus infection in Bangkok, 2016-2017. PeerJ. 2019;7:e6748.

112. Höper D, Hoffmann B, Beer M. A comprehensive deep sequencing strategy for full-length genomes of influenza A. PLoS One. 2011;6(4):e19075.

113. Zhao J, Liu J, Vemula SV, Lin C, Tan J, Ragupathy V, et al. Sensitive Detection and Simultaneous Discrimination of Influenza A and B Viruses in Nasopharyngeal Swabs in a Single Assay Using Next-Generation Sequencing-Based Diagnostics. PLoS One. 2016;11(9):e0163175.

114. Kampmann ML, Fordyce SL, Avila-Arcos MC, Rasmussen M, Willerslev E, Nielsen LP, et al. A simple method for the parallel deep sequencing of full influenza A genomes. J Virol Methods. 2011;178(1-2):243-8.

115. Zhao XN, Zhang HJ, Li D, Zhou JN, Chen YY, Sun YH, et al. Whole-genome sequencing reveals origin and evolution of influenza A(H1N1)pdm09 viruses in Lincang, China, from 2014 to 2018. PLoS One. 2020;15(6):e0234869.

116. Meinel DM, Heinzinger S, Eberle U, Ackermann N, Schönberger K, Sing A. Whole genome sequencing identifies influenza A H3N2 transmission and offers superior resolution to classical typing methods. Infection. 2018;46(1):69-76.

117. Xue KS, Moncla LH, Bedford T, Bloom JD. Within-Host Evolution of Human Influenza Virus. Trends Microbiol. 2018;26(9):781-93.

118. Van Poelvoorde LAE, Saelens X, Thomas I, Roosens NH. Next-Generation Sequencing: An Eye-Opener for the Surveillance of Antiviral Resistance in Influenza. Trends Biotechnol. 2020;38(4):360-7.

119. Deng YM, Spirason N, Iannello P, Jelley L, Lau H, Barr IG. A simplified Sanger sequencing method for full genome sequencing of multiple subtypes of human influenza A viruses. J Clin Virol. 2015;68:43-8.

120. Cobbin JCA, Alfelali M, Barasheed O, Taylor J, Dwyer DE, Kok J, et al. Multiple

Sources of Genetic Diversity of Influenza A Viruses during the Hajj. J Virol. 2017;91(11):e00096-17.

121. Sobel Leonard A, McClain MT, Smith GJ, Wentworth DE, Halpin RA, Lin X, et al. Deep Sequencing of Influenza A Virus from a Human Challenge Study Reveals a Selective Bottleneck and Only Limited Intrahost Genetic Diversification. J Virol. 2016;90(24):11247-58.

122. Rondy M, Kissling E, Emborg HD, Gherasim A, Pebody R, Trebbien R, et al. Interim 2017/18 influenza seasonal vaccine effectiveness: combined results from five European studies. Euro Surveill. 2018;23(9):18-00086.

123. Flannery B, Chung JR, Belongia EA, McLean HQ, Gaglani M, Murthy K, et al. Interim Estimates of 2017-18 Seasonal Influenza Vaccine Effectiveness - United States, February 2018. MMWR Morb Mortal Wkly Rep. 2018;67(6):180-5.

124. Centers for Disease Control and Prevention. Summary of the 2017-2018 Influenza Season. 2019 [Available from: https://www.cdc.gov/flu/about/season/flu-season-2017-2018.htm.

125. World Health Organization. Recommended composition of influenza virus vaccines for use in the 2018 southern hemisphere influenza season 2017 [Available from: https://www.who.int/influenza/vaccines/virus/recommendations/201709_recommendation.pdf?ua=1.

126. Suntronwong N, Klinfueng S, Korkong S, Vichaiwattana P, Thongmee T, Vongpunsawad S, et al. Characterizing genetic and antigenic divergence from vaccine strain of influenza A and B viruses circulating in Thailand, 2017-2020. Sci Rep. 2021;11(1):735.

127. Belanov SS, Bychkov D, Benner C, Ripatti S, Ojala T, Kankainen M, et al. Genome-Wide Analysis of Evolutionary Markers of Human Influenza A(H1N1)pdm09 and A(H3N2) Viruses May Guide Selection of Vaccine Strain Candidates. Genome Biol Evol. 2015;7(12):3472-83.

128. Elderfield RA, Watson SJ, Godlee A, Adamson WE, Thompson CI, Dunning J, et al. Accumulation of human-adapting mutations during circulation of A(H1N1)pdm09 influenza virus in humans in the United Kingdom. J Virol. 2014;88(22):13269-83.

129. Sarmah K, Borkakoty B, Sarma K, Hazarika R, Das PK, Jakharia A, et al. Genetic variations of the Hemagglutinin gene of Pandemic Influenza A (H1N1) viruses in Assam, India during 2016. 3 Biotech. 2018;8(10):408.

130. Ramos AP, Herrera BA, Ramírez OV, García AA, Jiménez MM, Valdés CS, et al. Molecular and phylogenetic analysis of influenza A H1N1 pandemic viruses in Cuba, May 2009 to August 2010. Int J Infect Dis. 2013;17(7):e565-7.

131. Roubidoux EK, Carreño JM, McMahon M, Jiang K, van Bakel H, Wilson P, et al. Mutations in the Hemagglutinin Stalk Domain Do Not Permit Escape from a Protective, Stalk-Based Vaccine-Induced Immune Response in the Mouse Model. mBio. 2021;12(1):e03617-20.

132. Jimenez-Jorge S, Pozo F, de Mateo S, Delgado-Sanz C, Casas I, Garcia-Cenoz M, et al. Influenza vaccine effectiveness in Spain 2013/14: subtype-specific early estimates using the cycEVA study. Euro Surveill. 2014;19(9):20727.

133. European Centre of Disease Prevention and Control (ECDC) (2016) Surveillance report. Influenza virus characterization. Summary Europe 2016 [Available from: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/influenza-virus-characterisation-may-2016.pdf.

134. Cheng AC, Subbarao K. Epidemiological Data on the Effectiveness of Influenza Vaccine-Another Piece of the Puzzle. J Infect Dis. 2018;218(2):176-8.

135. Jones S, Prasad R, Nair AS, Dharmaseelan S, Usha R, Nair RR, et al. Whole-Genome Sequences of Influenza A(H1N1)pdm09 Virus Isolates from Kerala, India. Genome Announc. 2017;5(28):e00598-17.

136. Al Khatib HA, Al Maslamani MA, Coyle PV, Thompson IR, Farag EA, Al Thani AA, et al. Inter- Versus Intra-Host Sequence Diversity of pH1N1 and Associated Clinical Outcomes. Microorganisms. 2020;8(1):133.

137. Clark AM, Nogales A, Martinez-Sobrido L, Topham DJ, DeDiego ML. Functional Evolution of Influenza Virus NS1 Protein in Currently Circulating Human 2009 Pandemic H1N1 Viruses. J Virol. 2017;91(17):e00721-17.

138. Krug RM. Functions of the influenza A virus NS1 protein in antiviral defense. Curr Opin Virol. 2015;12:1-6.

139. Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM. Influenza virus NS1 protein

interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol Cell. 1998;1(7):991-1000.

140. Komissarov A, Fadeev A, Sergeeva M, Petrov S, Sintsova K, Egorova A, et al. Rapid spread of influenza A(H1N1)pdm09 viruses with a new set of specific mutations in the internal genes in the beginning of 2015/2016 epidemic season in Moscow and Saint Petersburg (Russian Federation). Influenza Other Respir Viruses. 2016;10(4):247-53.

141. Mao H, Sun Y, Chen Y, Lou X, Yu Z, Wang X, et al. Co-circulation of influenza A(H1N1), A(H3N2), B(Yamagata) and B(Victoria) during the 2017-2018 influenza season in Zhejiang Province, China. Epidemiol Infect. 2020;148:e296.

142. Boonnak K, Mansanguan C, Schuerch D, Boonyuen U, Lerdsamran H, Jiamsomboon K, et al. Molecular Characterization of Seasonal Influenza A and B from Hospitalized Patients in Thailand in 2018-2019. Viruses. 2021;13(6).

143. Jagadesh A, Krishnan A, Nair S, Sivadas S, Arunkumar G. Genetic characterization of hemagglutinin (HA) gene of influenza A viruses circulating in Southwest India during 2017 season. Virus Genes. 2019;55(4):458-64.

144. Wan H, Gao J, Yang H, Yang S, Harvey R, Chen YQ, et al. The neuraminidase of A(H3N2) influenza viruses circulating since 2016 is antigenically distinct from the A/Hong Kong/4801/2014 vaccine strain. Nat Microbiol. 2019;4(12):2216-25.

145. Potter BI, Kondor R, Hadfield J, Huddleston J, Barnes J, Rowe T, et al. Evolution and rapid spread of a reassortant A(H3N2) virus that predominated the 2017-2018 influenza season. Virus Evol. 2019;5(2):vez046.

146. Takashita E, Daniels RS, Fujisaki S, Gregory V, Gubareva LV, Huang W, et al. Global update on the susceptibilities of human influenza viruses to neuraminidase inhibitors and the cap-dependent endonuclease inhibitor baloxavir, 2017-2018. Antiviral Res. 2020;175:104718.

147. de Steenhuijsen Piters WA, Sanders EA, Bogaert D. The role of the local microbial ecosystem in respiratory health and disease. Philos Trans R Soc Lond B Biol Sci. 2015;370(1675).

148. Kong WH, Li Y, Peng MW, Kong DG, Yang XB, Wang L, et al. SARS-CoV-2 detection in patients with influenza-like illness. Nat Microbiol. 2020;5(5):675-8.

149. Korten I, Mika M, Klenja S, Kieninger E, Mack I, Barbani MT, et al. Interactions of

respiratory viruses and the nasal microbiota during the first year of life in healthy infants. Msphere. 2016;1(6).

150. Mayuramart O, Nimsamer P, Rattanaburi S, Chantaravisoot N, Khongnomnan K, Chansaenroj J, et al. Detection of severe acute respiratory syndrome coronavirus 2 and influenza viruses based on CRISPR-Cas12a. Exp Biol Med (Maywood). 2021;246(4):400-5.

151. Farfour E, Lesprit P, Visseaux B, Pascreau T, Jolly E, Houhou N, et al. The Allplex 2019-nCoV (Seegene) assay: which performances are for SARS-CoV-2 infection diagnosis? Eur J Clin Microbiol Infect Dis. 2020;39(10):1997-2000.

152. Sawaswong V, Praianantathavorn K, Chanchaem P, Khamwut A, Kemthong T, Hamada Y, et al. Comparative analysis of oral-gut microbiota between captive and wild long-tailed macaque in Thailand. Sci Rep. 2021;11(1):14280.

153. Sirivongrangson P, Kulvichit W, Payungporn S, Pisitkun T, Chindamporn A, Peerapornratana S, et al. Endotoxemia and circulating bacteriome in severe COVID-19 patients. Intensive Care Med Exp. 2020;8(1):72.

154. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-7.

155. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581-3.

156. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72(7):5069-72.

157. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. Isme j. 2011;5(2):169-72.

158. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12(6):R60.

159. Li Y, Ding J, Xiao Y, Xu B, He W, Yang Y, et al. 16S rDNA sequencing analysis of upper respiratory tract flora in patients with influenza H1N1 virus infection. Frontiers in laboratory medicine. 2017;1(1):16-26.

160. Zhong H, Wang Y, Shi Z, Zhang L, Ren H, He W, et al. Characterization of

respiratory microbial dysbiosis in hospitalized COVID-19 patients. Cell Discov. 2021;7(1):23.

161. Xu R, Liu P, Zhang T, Wu Q, Zeng M, Ma Y, et al. Progressive deterioration of the upper respiratory tract and the gut microbiomes in children during the early infection stages of COVID-19. J Genet Genomics. 2021;48(9):803-14.

162. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C. The microbiome of the upper respiratory tract in health and disease. BMC Biol. 2019;17(1):87.

163. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med. 2011;184(8):957-63.

164. Schenck LP, Surette MG, Bowdish DM. Composition and immunological significance of the upper respiratory tract microbiota. FEBS Lett. 2016;590(21):3705-20.

165. De Boeck I, Wittouck S, Wuyts S, Oerlemans EFM, van den Broek MFL, Vandenheuvel D, et al. Comparing the Healthy Nose and Nasopharynx Microbiota Reveals Continuity As Well As Niche-Specificity. Front Microbiol. 2017;8:2372.

166. Lemon KP, Klepac-Ceraj V, Schiffer HK, Brodie EL, Lynch SV, Kolter R. Comparative analyses of the bacterial microbiota of the human nostril and oropharynx. mBio. 2010;1(3).

167. Lysholm F, Wetterbom A, Lindau C, Darban H, Bjerkner A, Fahlander K, et al. Characterization of the viral microbiome in patients with severe lower respiratory tract infections, using metagenomic sequencing. PLoS One. 2012;7(2):e30875.

168. Aljabr W, Alruwaili M, Penrice-Randal R, Alrezaihi A, Harrison AJ, Ryan Y, et al. Amplicon and Metagenomic Analysis of Middle East Respiratory Syndrome (MERS) Coronavirus and the Microbiome in Patients with Severe MERS. mSphere. 2021;6(4):e0021921.

169. Dandagi GL. Nosocomial pneumonia in critically ill patients. Lung India. 2010;27(3):149-53.

170. Cusumano JA, Dupper AC, Malik Y, Gavioli EM, Banga J, Berbel Caban A, et al. Staphylococcus aureus Bacteremia in Patients Infected With COVID-19: A Case Series. Open Forum Infect Dis. 2020;7(11):ofaa518.

VITA

NAME	Miss Somruthai Rattanaburi
DATE OF BIRTH	16 July 1993
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Biomedical Sciences (Interdisciplinary Program), Graduate School, Chulalongkorn University
HOME ADDRESS	9 Ladplakao 76, Ladplakao road, Anusaowaree, Bangkhen, Bangkok 10220
PUBLICATION	Rattanaburi S, Sawaswong V, Chitcharoen S, Sivapornnukul P, Nimsamer P, Suntronwong N, Puenpa J, Poovorawan Y, Payungporn S. Bacterial microbiota in upper respiratory tract of COVID-19 and influenza patients. Experimental Biology and Medicine, 2021 Nov 13:15353702211057473
	Nimsamer P, Mayuramart O, Rattanaburi S, Chantaravisoot N, Saengchoowong S, Puenpa J, Poovorawan Y, Payungporn S. Comparative performance of CRISPR-Cas12a assays for SARS-CoV-2 detection tested with RNA extracted from clinical specimens. Journal of Virological Methods, 2021 Apr;290:114092.
	Mayuramart O, Nimsamer P, Rattanaburi S, Chantaravisoot N, Khongnomnan K, Chansaenroj J, Puenpa J, Suntronwong N, Vichaiwattana P, Poovorawan Y, Payungporn S. Detection of severe acute respiratory syndrome coronavirus 2 and influenza viruses based on CRISPR-Cas12a. . Experimental Biology and Medicine, 2021 Feb;246(4):400-405.