

**MOLECULAR CHARACTERIZATION OF HAEMAGGLUTININ
GENES OF INFLUENZA B VIRUSES CIRCULATING IN GHANA
DURING 2016 AND 2017**

BY

ALHASSAN MOHAMMED YAKUBU

(10598419)

**THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA,
LEGON IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR
THE AWARD OF
MPHIL MEDICAL BIOCHEMISTRY DEGREE**

JULY, 2018

DECLARATION

I, Alhassan Mohammed Yakubu, of the Department of Medical Biochemistry, School of Biomedical and Allied Health Sciences, do hereby declare that, with the exception of quoted articles and references, this work was duly carried out by me and the results obtained herein are a true reflection of the work done under the supervision of Dr. Nii Ayite Aryee at the Department of Medical Biochemistry and Prof. William Kwabena Ampofo, Department of Virology at Noguchi Memorial Institute for Medical Research, University of Ghana.

Signature

Date

Alhassan Mohammed Yakubu

(Student)

Signature

Date

Dr. Nii Ayite Aryee
(Supervisor)

Signature

Date

Prof. William Kwabena Ampofo

(Supervisor)

ABSTRACT

Background: Influenza B viruses are receiving attention lately perhaps due to the degree of haemagglutinin (HA) antigen mismatching between vaccine composition strains and circulating strains. Neuraminidase inhibitor (NI) antiviral has also been implicated in prophylaxis and treatment of severe cases of influenza diseases. However, mutations in the two major surface glycoproteins, HA and neuraminidase (NA) could cause the virus to escape host defence mechanisms leading to failure of the host immune system to recognize the viruses as well as failure to antiviral therapy, data seems scanty hence the focus of this research.

Main Objective: To molecularly characterize the lineages of influenza B virus strains that circulated in Ghana between 2016 and 2017.

Methods: This was a retrospective cross-sectional study that used selected Ghanaian archived influenza B clinical specimens: representatives of both lineages. Viral RNAs were extracted and amplified using real time reverse transcriptase assays (rtRT-PCR) and subsequently sequenced and analyzed.

Results: A total of eleven and six amino acids substitutions were detected in the recent Ghanaian influenza B strains-influenza B Victoria and Yamagata lineages, respectively. Both Influenza B Victoria and Yamagata lineages were also closely related to Influenza B/Brisbane/60/2008 and Influenza B/Phuket/3073/2013 respectively.

Conclusions: Three main amino acid substitutions (P31S, I117V and R151K) were found in influenza B Victoria lineages circulating between 2016 and 2017, with one strain possessing a unique glycosylation site at amino acid position 51 in the HA2 subunit. Two main substitutions (L172Q and M251V) were also detected in the HA gene of influenza B Yamagata. The recent deletion sub-group in influenza B virus reported by the US CDC was

not identified among analysed specimens. Monitoring of the patterns of influenza B evolution would aid in efficient selection of representative viruses for use in the both the design and formulation of influenza vaccines.

DEDICATION

I dedicate this project to my wife, Jamilatu Baba, my children (Naasiratu Nasara Alhassan, Muhsinatu Maltiti Alhassan and Yakubu Tipagya Alhassan) and all my siblings for their spiritual guidance.

ACKNOWLEDGEMENT

All that I am and all that I will be is by Allah. I could not have done anything unless it had been given to me from the Most High. My tremendous appreciation goes to the Almighty God for seeing me throughout the two-year's course.

Secondly, I wish to acknowledge the efforts of my supervisors Dr. Nii Ayite Aryee and Prof. William Kwabena Ampofo for their guidance throughout this research. Sincere appreciation to US Naval Medical Research Unit 3, Ghana Detachment for provision of all reagents and consumables.

Immense appreciation to fellows and colleagues of the National Influenza Centre (NIC), Noguchi Memorial Institute for Medical Research, for granting me the space and equipment for the work. Many thanks to Dr. Kofi Bonney, Dr. Mrs. Evelyn Yayra Bonney and Mr. Erasmus Kotey for their technical assistance, guidance and warm reception.

Due appreciation to Dr. Grace Korkor Ababio of the Medical Biochemistry Department, University of Ghana, Michael R. Wiley and Catherine B. Pratt of the Center for Genome Sciences, United States Army Medical Research Institute of Infectious Diseases (USAMRIID) for their technical assistance and warm reception.

Many thanks to the Department of Medical Biochemistry, School of Biomedical and Allied Health Sciences and Virology Department of Noguchi Memorial Institute for Medical Research, College of Health Sciences, University of Ghana for training me as an M. Phil student.

My final thanks go to the Ghana Armed Forces High Command, Chief of Army Staff, Director General Medical Services, 37 Military Hospital Command and the management of Pathology Division Laboratory for granting me the study leave.

TABLE OF CONTENTS

CONTENTS	PAGE
DECLARATION	ii
ABSTRACT.....	iii
DEDICATION.....	v
ACKNOWLEDGEMENT	vi
TABLE OF CONTENTS.....	vii
LIST OF TABLES.....	xi
LIST OF FIGURES	xii
LIST OF APPENDICES.....	xiii
LIST OF ABBREVIATIONS.....	xiv
CHAPTER ONE.....	1
1.0 INTRODUCTION	1
1.1 Background of Research	1
1.2 Problem Statement	3
1.3 Justification	4
1.4 Research Hypothesis	5
1.5 Overall Objective	5
1.6 Specific Objectives.....	5
CHAPTER TWO	6

2.0	LITERATURE REVIEW	6
2.1	INFLUENZA VIRUSES.....	6
2.1.1	Classification.....	6
2.1.2	Lineage-specific markers of Influenza B Viruses.....	6
2.1.3	Morphology.....	7
2.1.4	Replication cycle.....	8
2.2	GENOME STRUCTURE OF INFLUENZA VIRUSES	9
2.2.1	Internal proteins	10
a.	Polymerase Basic Protein 2.....	10
b.	Polymerase Basic Protein 1	10
c.	Polymerase Acid Protein (PA).....	11
d.	Nucleoprotein.....	11
e.	Matrix proteins	11
f.	Non-Structural proteins	12
2.2.2	External Proteins.....	13
a.	Haemagglutinin	13
b.	Neuraminidase and NB	16
2.3	THE INFLUENZA VIRAL DISEASE.....	17
2.3.1	Clinical presentation of Influenza and its determinants.....	18
2.3.2	Epidemiology.....	18

2.4	THE INFLUENZA SURVEILLANCE IN GHANA.....	19
CHAPTER THREE		23
3.0	METHODS	23
3.1	Research Design.....	23
3.2	Study Sites and Sample Collection	23
3.3.	Inclusion Criteria.....	23
3.4	Exclusion Criteria.....	23
3.5	Sample Selection.....	24
3.6	Laboratory Procedure.....	25
3.6.1	Genomic Detection of Influenza Viruses.....	25
3.6.1.1	Viral Ribonucleic Acid (RNA) Extraction.....	25
3.6.1.2	Conventional RT-PCR protocol for Influenza B HA gene	26
3.6.1.3	Agarose Gel electrophoresis of the amplicons.....	27
3.6.1.4	Post PCR purification.....	27
3.6.1.5	Cycle sequencing.....	28
3.6.1.6	Post cycle sequencing purification.....	28
3.7	Gene assembling, multiple sequence analyses and phylogeny	29
3.8	Data handling	30
3.9	Ethical Consideration	30
CHAPTER FOUR.....		31

4.0	RESULTS	31
4.1	Study Participants.....	31
4.2	Fragments Amplification.....	32
4.3	Nucleotide changes distinguishing Influenza B Victoria from Yamagata lineages..	34
4.4	Analysis of Antigenic drift on Influenza B Victoria HA gene.....	34
4.5	Analysis of Antigenic drift on Influenza B Yamagata HA gene.....	36
4.6	Glycosylation analysis.....	38
4.7	Phylogenetic analysis	41
4.7.1	Phylogenetic analysis of Influenza B Victoria lineage.....	41
4.7.2	Phylogenetic analysis of Influenza B Yamagata lineage.....	42
	CHAPTER FIVE	43
5.0	DISCUSSION AND CONCLUSIONS	43
5.1	Discussion	43
5.2	Conclusions	46
	RECOMMENDATIONS	47
	REFERENCES	48
	APPENDICES	56

LIST OF TABLES

TABLE	PAGE
Table 1: Cycle threshold (Ct) values obtained from real time RT-PCR.....	24
Table 2: Demographics and clinical symptoms of study participants	32
Table 3: Summary of amplification	33
Table 4: Amino acid substitutions in the HA genes of Influenza B Victoria lineage.....	36
Table 5: Amino acid substitutions in the HA genes of Influenza B Yamagata lineages	38
Table 6: Potential Glycosylation sites of Influenza B Victoria HA genes	39
Table 7: Potential Glycosylation sites of Influenza B Yamagata HA genes	40

LIST OF FIGURES

FIGURE	PAGE
Figure 1: Structure of Influenza B virus	8
Figure 2: Influenza B virus replication cycle.....	9
Figure 3: Epitopes of Influenza B virus HA	15
Figure 4: A map of Ghana showing the locations of sentinel sites as at December 2017	21
Figure 5: Occurrence of Influenza B real time RT-PCR positives from 2010 to 2018	22
Figure 6: A schematic of primer locations on the Influenza B HA gene.....	27
Figure 7: Lineage specific markers of Influenza B HA gene	34
Figure 8: Phylogenetic analysis of Influenza B Victoria lineage using HA genes	41
Figure 9: Phylogenetic analysis of Influenza B Yamagata lineage using the HA genes	42

LIST OF APPENDICES

APPENDIX	PAGE
Appendix I: Scanned copy of influenza case investigation form-Ghana.....	56
Appendix II: Selection of specimens from the NIC-Ghana confirmed influenza B positives samples during 2016 and 2017	58
Appendix III: Summary of sequencing out come	59
Appendix IV: Fragments assembling using Geneious 10.2.3.....	60
Appendix V: A sample chromatogram of influenza B HA gene fragment.....	61
Appendix VI: HA amino acid alignment for influenza B Victoria lineage viruses.....	62
Appendix VII: HA amino acid alignment for influenza B Yamagata lineage.....	65
Appendix VIII: Ethical clearance certificate from the College of Health Sciences	68
Appendix IX: Ethical clearance certificate from the Noguchi Memorial Institute for Medical Research.....	69
Appendix X: Detailed of primers used for this research.....	70
Appendix XI: A sample gel image.....	71

LIST OF ABBREVIATIONS

A	Alanine
ARI	Acute Respiratory Infection
B/Vic	Influenza B Victoria
bp	Base pairs
CDC	Centres for Disease Control and Prevention
cRNA	Complementary RNA
D	Aspartic acid
E	Glutamic acid
ECDC	European Centre for Disease Prevention and Control
EDTA	Ethylene Diamine Tetra-acetic Acid
G	Glycine
H1N1pdm09	Influenza A (H1N1) Pandemic 2009 strain
HA	Haemagglutinin
I	Isoleucine
ILI	Influenza-Like Illness
K	Lysine
L	Leucine

M	Methionine
M1	Matrix protein 1
M2	Matrix protein 2
mRNA	Messenger RNA
N	Asparagine
NA	Neuraminidase
NI	Neuraminidase Inhibitor
NP	Nucleoprotein
NS1	Nonstructural protein 1
nts	Nucleotides
ORF	Open Reading Frame
P	Proline
PA	Polymerase acid protein
PB1	Polymerase basic protein 1
PCR	Polymerase Chain Reaction
PKR	Protein Kinase R
Q	Glutamine
R	Arginine

RIG-1	Retinoic acid-inducible gene 1
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RT-PCR	Reverse transcriptase polymerase chain reaction
S	Serine
SARI	Severe Acute Respiratory Infection
ssRNA	Single Stranded RNA
T	Threonine
TLR7	Toll-like receptor 7
US	United States
USAMRIID	United States Army Medical Research Institute of Infectious Diseases
US-NAMRU 3	United States Naval Medical Research Unit 3
V	Valine
vRNA	Viral RNA
WHO CC	World Health Organization Collaborating Centre
WHO	World Health Organization

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background of Research

Influenza virus is known to cause a major respiratory infections leading to epidemics worldwide, nearly every year, with an estimated annual global attack rate of 5-10 % and 20-30 % in adults and children, respectively (WHO, 2012). Presently, seasonal trivalent influenza vaccines are designed to protect against three influenza viruses namely, influenza A/H1N1pdm09, A/H3N2 and B viruses. Since each type/subtype of influenza virus continually evolve, the specific strains to be included in seasonal influenza vaccines are chosen based on a prediction of the strains likely to circulate in the upcoming influenza season.

Despite the fact that there are two antigenically distinct lineages of influenza B viruses co-circulating globally since 1985 (Rota et al., 1990), only one lineage can be selected for inclusion in current trivalent influenza vaccines. The trivalent vaccines however, have been reported to provide limited immunity against strains of the other lineage (Belshe, Coelingh, Ambrose, Woo, & Wu, 2010). Studies have also shown that, in 5 out of 10 influenza seasons (2001 to 2011), the predominant circulating influenza B lineage was different from that chosen for the vaccine (Ambrose & Levin, 2012). As a result, influenza vaccination campaigns have had limited effects against influenza B epidemics during seasons in which a significant proportion of the disease was caused by opposite-lineage influenza B strains (Ambrose & Levin, 2012).

A recent World Health Organization (WHO) update has it that 39 Influenza B Victoria (B/Vic) deletion variant viruses have been detected in the United States (Lenee Blanton et al., 2017). As there are two kinds of influenza B, there are also vaccine preparations with an additional B virus. These are known as quadrivalent vaccines which further improves the effectiveness of vaccine formulation (Ambrose & Levin, 2012).

In sub-Saharan Africa, seasonal influenza epidemiology reports revealed that 10% of outpatient acute respiratory cases and 7% of children hospitalized with acute respiratory infection (ARI) tested positive for influenza (Gessner, Shindo, & Briand, 2011). Influenza disease has a correlation with economic burden arising from health-care costs, lost days of work or education, and general social disruption across all age groups (WHO, 2017a).

In Ghana, it is estimated that the incidence of influenza accounted for 9% and 18% of medically attended severe acute respiratory illness (SARI) and Influenza-like illness (ILI), respectively. These rates were substantially high among children (Ntiri et al., 2016).

Averagely, influenza B is responsible for about 25% of laboratory documented influenza. Morbidity is highest in school children, but all age groups are at risk (Glezen, 2014). The burden of Influenza B virus infection is reported to be substantial for children and adolescents (Glezen, 2014). Also, 22–44 % of pediatric influenza related deaths in the United States are caused by influenza B (Glezen, 2014; Thompson et al., 2003).

However, there is incomplete documentation on Influenza B virus infection in Ghana. Reports have recently indicated vaccine mismatch with respect to the influenza B viruses in circulation in parts of United States and China. These variants currently known as the deletion sub-group poses amino acid deletion at position 162, 163 and/or 164. The 162 and 163 deletions were detected in Georgia, United State, while that of 162, 163 and 164 deletions detected in Hong-Kong, China (Lenee Blanton et al., 2017). Influenza virus

neuraminidase (NA) and the haemagglutinin (HA) proteins are the two main glycoproteins on the surface of the virus particle which are involved in the interaction between the host cells and the virus (Gloria, 2011).

Influenza B viruses have only one HA and one NA subtype (Gloria, 2011) but with two distinct lineages, the Victoria and Yamagata lineages (R. Chen & Holmes, 2008). Influenza A viruses have received much more attention than B due to having several HA and NA proteins and hence their virulence and many hosts. Notwithstanding this fact, influenza B has the ability to cause fulminant diseases, precipitate Reye's syndrome, and result in deadly illness. Available data suggest that, influenza B caused 38% of pediatric influenza deaths in the US between 2010 and 2011 (CDC, 2011) giving the potential for pediatric hospitalization and deaths. Thus, influenza B poses a significant burden on the healthcare system, planning and preparedness (CDC, 2011) as it has strong potential to cause a scourge.

1.2 Problem Statement

In general, there is limited research data on influenza B as compared to influenza A (Jackson, Elderfield, & Barclay, 2011). Influenza B is capable of causing fulminant disease such as Reye's syndrome, which leads in fatal illness; available data suggest that 38% of pediatric influenza deaths in the US between 2010 and 2011 were associated with Influenza B. The potential for pediatric hospitalizations and deaths caused by influenza B poses a serious risk on the healthcare systems and for influenza pandemic planning and preparedness. The sporadic emergence of mutant viruses compromises effectiveness of vaccination and as a result the population immunological defense will continuously be manipulated by this virus (Glezen, 2014). Therefore, this requires continues monitoring by regular characterization of the circulating strains.

Between 2016 and 2017, influenza B Victoria lineage deletion variant viruses have been

detected in some parts the United States and China. These variants were antigenically distinct from influenza B/Brisbane/60/2008-like virus that was recommended for use in the 2016-2017 trivalent influenza vaccine (L. Blanton et al., 2017; ECDC, 2017). Meanwhile, there are no reports that indicate an influenza vaccine composition that caters for these variants, perhaps due to less surveillance data about these variants. There is therefore the need to constantly monitor and characterize such variants in different populations, as this would inform on the subsequent influenza vaccine formulation. The characterization of globally representative influenza strains is conducted by WHO Collaborating centers (CC). In a bid to improve virological surveillance in Ghana, this molecular epidemiological study therefore sought to establish the characteristics of recent Influenza B virus strains in Ghana.

1.3 Justification

Reports of 2017-2018 influenza season reveal influenza B dominance over influenza A in Europe (Tan, Arunkumar, & Krammer, 2018). In Ghana, the 2016 influenza surveillance data also suggests influenza B dominance especially during the dry season as compared to Influenza type A (WHO, 2017b). Noticing the capacity of influenza B virus to cause hospitalizations or deaths, there is the need to acquire information on the characteristics of the circulating influenza B strains. This research could become the model for subsequent characterization of influenza B viruses from clinical cases identified in Ghana and lead to a long-term study on the influenza B viruses in Ghana. Data from this study will help inform policy makers on the appropriate influenza vaccine should the need for influenza immunization be adopted in Ghana.

The project is consistent with the mission of the University of Ghana to develop world class human resources and builds capacity for medical research in influenza which is of both national and global public health importance.

1.4 Research Hypothesis

- a. There are no differences in the lineage-specific haemagglutinins of influenza B Victoria or Yamagata that circulated in Ghana in 2016 and 2017.
- b. The influenza B virus lineages in Ghana are not different from those circulating globally.

1.5 Overall Objective

To molecularly characterize the lineages of influenza B virus strains that circulated in Ghana between 2016 and 2017.

1.6 Specific Objectives

- a. To determine the genetic variations within haemagglutinin (HA) genes of influenza B viruses in Ghana.
- b. To establish the phylogeny of these Ghanaian influenza B viruses by comparing them to recent vaccine virus strains.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 INFLUENZA VIRUSES

2.1.1 Classification

Influenza virus infection, commonly called “the Flu”, is a contagious respiratory illness, causing seasonal influenza epidemics every year and is an important disease responsible for school and job absenteeism, hospitalization and deaths (Zaman et al., 2009) and this is of public health concern. Influenza viruses of medical importance are from the family of *Orthomyxoviridae* (Hause et al., 2014). The *Orthomyxoviridae* family is divided into six genera; Influenza virus A, Influenza virus B, Influenza virus C, Influenza virus D, Isavirus and Thogotovirus, based on antigenic differences in two of the major structural proteins of the virus, the nucleoprotein (NP) and the matrix protein (M) (Gloria, 2011). Orthomyxoviruses are capable of removing chemical side chains (sialic acids) from mucoproteins thereby facilitating infection of cells of the mucous membranes in the respiratory tract (Gloria, 2011). Human influenza A and B viruses are known to cause seasonal epidemics of disease almost every winter in the United States. Influenza C infections generally results in mild respiratory illness and are not thought to cause epidemics. However, influenza D primarily affects swine and cattle but are not known to infect or cause illness in humans (CDC, 2016; Chiapponi et al., 2016).

2.1.2 Lineage-specific markers of Influenza B Viruses

Influenza B Victoria (B/Victoria/2/1987) and Yamagata (B/Yamagata/16/1988) lineages have co-circulated since the late 1980s (Burton, Poirnard, Stanfield, & Wilson, 2012). Several

lineage-specific molecular signatures have been identified on the HA gene as markers that distinguish Influenza B Victoria from Influenza B Yamagata lineages. These variations of nucleotide sequences can be found at positions 522, 538–540, 548, 555, 558, 568 as well as two amino acids deletion at residue 162 and 163 of HA1 of influenza B Victoria (Arvia, Corcioli, Pierucci, & Azzi, 2014).

2.1.3 Morphology

Influenza viruses are mostly spherical, with about 500 projecting spikes covering a lipid envelope. These spikes are actually glycoproteins, because they consist of protein linked to sugars known as haemagglutinin (HA) and neuraminidase (NA) (Gloria, 2011). These HA and NA are proteins that determine the subtype of influenza virus. The HA helps in the attachment of the viral strain on the host cell surface while the NA is responsible for initiation of viral infection (Gloria, 2011). The gene coding assignments in influenza B virus are similar to influenza A virus except that PB1-F2 is absent and the NA gene segment codes for both NA and NB protein (Shaw, Choppin, & Lamb, 1983).

There are eight discrete fragments of the negative-sense ssRNA genome, approximately 17 kb in size that complexed with proteins (NP, PA, PB1, and PB2) to form a ribonucleoprotein (RNP) arranged in a helix (Fig. 1). The 5' and 3' ends of each of the genome segments contain sequences that are complementary and each genome segments bind through these complementary ends to form a pan handle (Gloria, 2011).

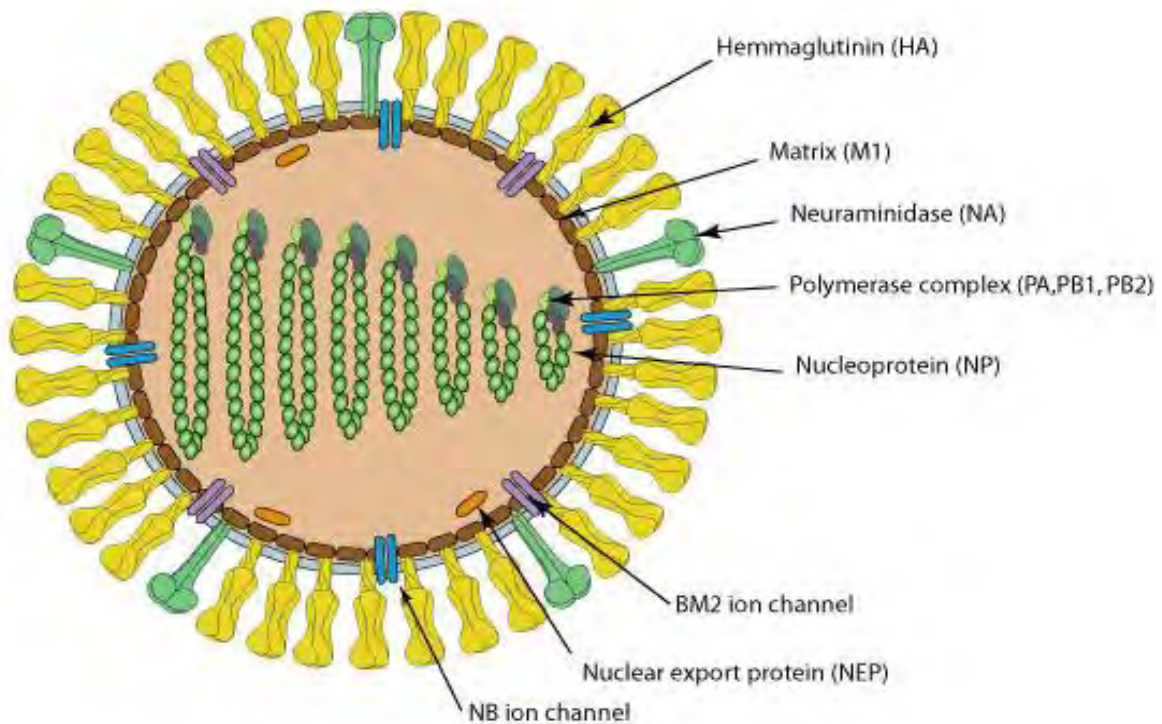


Figure 1: Structure of Influenza B virus

Source: (SIB, 2012).

2.1.4 Replication cycle

Interactions between host and pathogen are mediated by HA antigen which serve as a ligand on the side of the virus and sialic acid on the side of the host cell as a receptor (Figure 2). Successful attachment of the virus to the receptor results in membrane fusion and the formation of an endosome through endocytosis (Hyland, Webby, Sandbulte, Clarke, & Hou, 2006). Each segmented genome codes for a particular protein and is transcribed by RNA polymerase to the positive sense with mRNA strand before it is translated into proteins and then assembled into a new virion. Each virus leaves the cell by simple budding out using the NA which cleaves sialic acid sugars in the membrane releasing a newly created virus from

the cell. Apoptosis of the host cell occurs after the release of newly replicated virions (Krejcova, Michalek, Hynek, Adam, & Kizek, 2015).

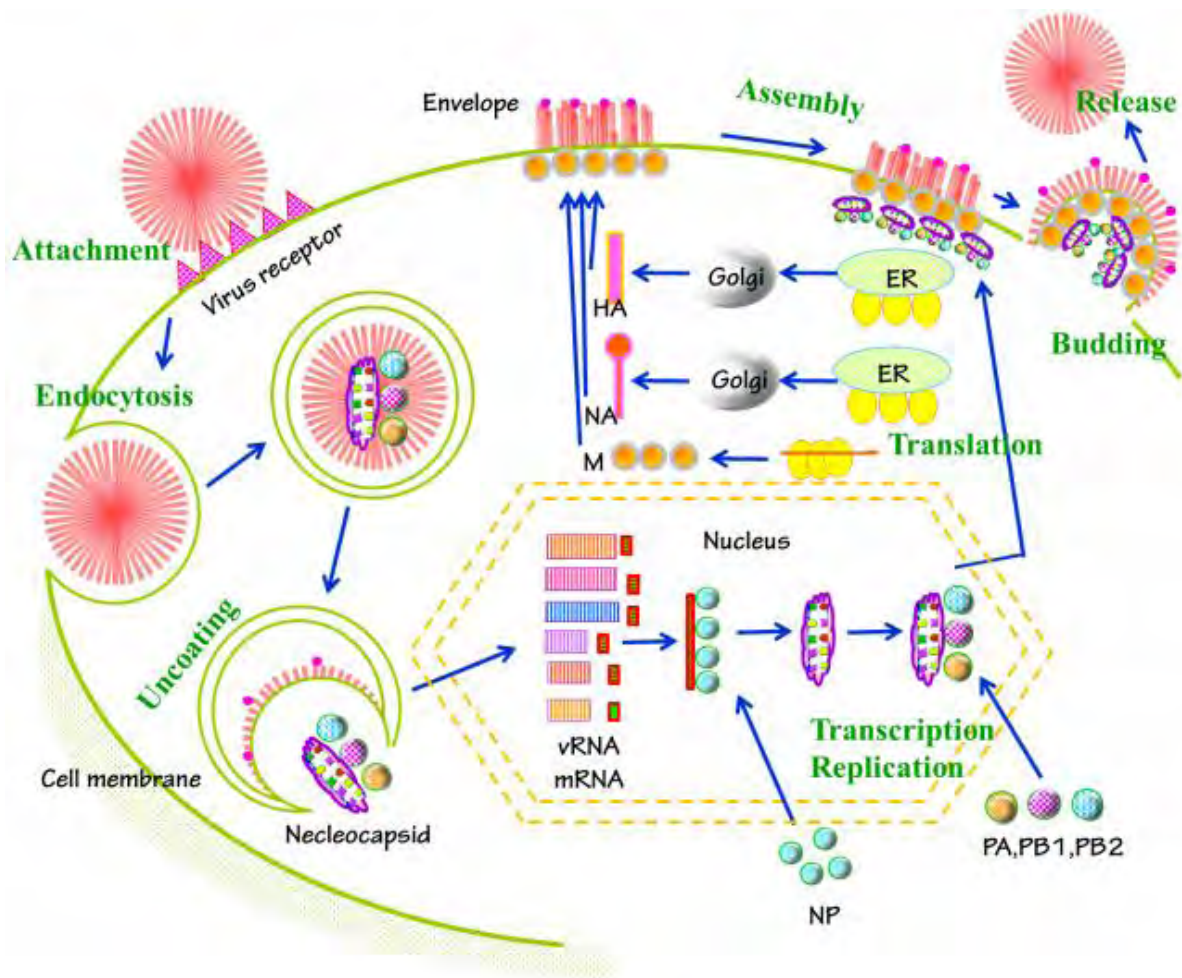


Figure 2: Influenza B virus replication cycle

Source: (Kim, 2012)

2.2 GENOME STRUCTURE OF INFLUENZA VIRUSES

Influenza A and B viruses have eight genomic segments while type C has seven segments (Palese, 1977). The first three segments encode the subunits of the viral polymerase complex. These segments are polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and

polymerase acidic protein (PA). These three are the largest gene segments and are known to transcribe messenger RNAs and to synthesize positive sense antigenomic template RNAs (cRNAs) which are transcribed into genomic segments (vRNAs). These gene segments are expressed in two types of proteins, the internal and external proteins (Palese, 1977).

2.2.1 Internal proteins

a. Polymerase Basic Protein 2

The first segment of influenza B viruses encodes the PB2 protein which is 2.4 kb in size. This segment is part of the polymerase complex which carries out virus transcription and replication. The PB2 contains binding sites for both the PB1 subunit and the nucleoprotein (NP) and is also responsible for cap-binding during viral mRNA synthesis. The cap-binding properties of PB2 leads to creation of primers for viral mRNA synthesis by binding to the 5' methylated cap of the host cell pre-mRNAs before they are cleaved (Shi, Galarza, & Summers, 1996).

b. Polymerase Basic Protein 1

The second segment encodes the PB1 protein whose function is for polymerase activity. The PB1 subunit contains binding sites for both the PA and PB2 thereby playing a key role in both the assembly of the three polymerase protein subunits and serves the catalytic function of RNA polymerization. It has been proposed that the catalytic specificity of the PB1 subunit is modulated by transcriptase by binding PB2 or the replicase by interaction with PA (Gloria, 2011). The PB1 segment encodes an additional protein, PB1-F2, which has been implicated in regulation of polymerase activity, immunopathology, susceptibility to secondary bacterial infection, and induction of apoptosis (McAuley et al., 2007).

c. Polymerase Acid Protein (PA)

The third segment encodes the acidic polymerase protein (PA). The PA contains nuclear localization signals (Nieto, de la Luna, Barcena, Portela, & Ortin, 1994) required for transport into the nucleus. Studies have shown that the PA subunit contains the endonuclease active site which synthesizes the viral messenger RNAs. In addition, the PA contains residues important for protein stability, promoter binding and cap-binding. PA-X is a fusion protein incorporating the N-terminal endonuclease domain of the PA protein with a short C-terminal domain encoded by an overlapping ORF („X“) in segment 3 (Jagger et al., 2012). This protein functions to inhibit cellular gene expression especially those genes involved in regulating the initiation of the cellular immune response which results in host cell shut off (Weber, Gruber, Haller, & Kochs, 1999).

d. Nucleoprotein

The product of the fifth segment is the nucleoprotein (NP) which is 1.5 kb in size (Gloria, 2011). The NP mediates the transport of the viral ribonucleoproteins (RNP) from the viral particles to the nucleus and also plays a role in assembly and budding of the influenza virus. It encapsidates the negative strand viral RNA for it to be recognized as templates for the viral polymerase (Portela & Digard, 2002).

e. Matrix proteins

The seventh segment encodes two gene products, the matrix 1 (M1) and matrix 2 (M2) proteins (BM2 in influenza B viruses) which is 1.027 kb in length. The M1 interacts with both the viral RNP and surface glycoproteins. The M1 mRNA product after transcription is responsible for virus assembly and budding after viral replication while the M2 is an integral membrane protein which forms an ion channel on the viral envelope (Gloria, 2011). The M2

protein is only present in influenza A viruses and form a drug target for the anti-influenza drug called adamantanes (Gloria, 2011). This class of drugs belongs to the M2 ion channel blocker group and includes rimantidine and amantadine. Amantadine is more commonly used and has more side effects than rimantadine. The adamantanes are ineffective against influenza B virus since these viruses lack the M2 protein in their viral structure but have a substitute called the NB that is not affected by amantadine (Brassard, Leser, & Lamb, 1996).

f. Non-Structural proteins

The eighth segment is 890 bases long and encodes two proteins: nonstructural protein 1 (NS1) and nonstructural protein 2 (NS2). The NS1 protein is a multifunctional protein involved in nuclear exportation of mRNA, posttranscriptional regulation, and inhibition of cellular interferon response (Gloria, 2011) while the NS2 protein mediates the nuclear export of virion RNAs by acting as an adaptor between viral ribonucleoprotein complexes and the nuclear export machinery of the cell (Steinhauer & Skehel, 2002b).

The NS1 has been shown to block the innate host immune response by interfering with the signaling pathway of retinoic acid-inducible gene-1 (RIG-1) which together with the toll-like receptor 7 (TLR 7) activate antiviral host responses and lead to production of type 1 interferons (Kumagai, Takeuchi, & Akira, 2008). The NS1 also attaches to the antiviral protein, protein kinase R (PKR), whose function is to suppress the translation of viral mRNA in the host cell. When the NS1 binds to the PKR, there is inhibition of antiviral function of PKR (Min, Li, Sen, & Krug, 2007).

2.2.2 External Proteins

a. Haemagglutinin

The haemagglutinin (HA) glycoprotein is the fourth gene segment of influenza B virus and 1.7 kb in size. The HA is regarded as the binding and fusion protein which is responsible for binding to the monosaccharide sialic acid present on the host cells surface. The virus is then engulfed by the host cell through endocytosis forming an endosome (Wang, 2010). The HA is synthesized as a precursor polypeptide, HA0, which is post-translated and then cleaved into two disulphide-linked subunits, HA1 and HA2 (figure 3) (Gloria, 2011). The cleavage of the HA0 into HA1 and HA2 is essential for viral infectivity. The cleavage also allows the native HA molecule to undergo a conformational change, triggered by an acidic environment and is essential for membrane fusion (Gloria, 2011). The extracellular HA0 cleavage is believed to be activated by trypsin-like proteases (Gloria, 2011). However, the presence of multiple basic amino acid residues within the cleavage site allows the protein to be cleaved by intracellular proteases e.g. furin (J. Chen et al., 1998).

There are various strategies employed by Influenza virus to ensure immune evasion from the host (Alcami & Koszinowski, 2000). Adaptive immune response includes generation of quasi viral species formed from accumulating amino acid substitutions in the antigenic sites of HA that are recognized by virus-neutralizing antibodies. Commonly known as antigenic drift, this phenomenon has been shown to allow the virus to evade recognition by antibodies and to cause recurrent influenza epidemics annually (Steinhauer & Skehel, 2002a). There are various strategies used by viruses to evade recognition by virus-specific T cells. For example, encoding proteins that interfere with various steps in the antigen processing and presentation pathways is a mechanism commonly employed by viruses with large genomes. Most RNA

viruses with smaller genomes and limited coding capacity, employ high mutation rates and subsequent selective pressure as a means of evading recognition by T cells (Reaney, 1982).

Antigenic variation in influenza B virus is mostly caused by amino acid substitutions at four major antigenic epitopes (120-loop, 150-loop, 160 loop and 190- helix). These epitopes have been identified in previous studies (Pechirra et al., 2005). Thus, mutations in the HA antigenic region (positive selection) can result in vaccine failure due to the virus not being recognized by the immune system (S. Shen et al., 2008).

Influenza B/Brisbane/60/2008, the current quadrivalent candidate vaccine component for the southern hemisphere is defined by the amino acid substitutions N75K, N165K and S172P on the HA1 coding region of the HA gene of Influenza B Victoria lineage (Byarugaba et al., 2013).

The 120 loop and its surrounding regions; The 120 loop is located at amino acid positions 116-137 on the HA structure. This loop appears to be one of the most frequently mutated regions in field isolates (Q. Wang, F. Cheng, M. Lu, X. Tian, & J. Ma, 2008). Influenza B Victoria lineage with specific substitutions at HA1 residues 129 and 137, produced by using reverse genetics, were found to cause altered antigenicity (Qinghua Wang, Feng Cheng, Mingyang Lu, Xia Tian, & Jianpeng Ma, 2008).

ii. The 150 loop; This loop is located at position 141-150 on the HA structure. It has an unusually long protruding loop, the tip, which code for Threonine at position 147 is pointed away from the main body of the structure by 9 Å. This site is one of the reported antibody-binding sites on B/HongKong/8/73 HA (Qinghua Wang et al., 2008).

iii. The 160 loop; The 160 loop is located at position 162-167 on the HA structure. It is the only region in influenza B virus HAs where insertion and deletion have frequently been

detected in field isolates from different epidemic seasons. The protruding nature of the 160 loop may make it easy to accommodate more multiple-residue for insertions or deletions (Qinghua Wang et al., 2008).

iv. The 190 helix and its surrounding regions; This loop is located at position 194-202 on the HA structure. All of the residues at the external face of the 190 helix have important antigenic roles. The location on the 190 helix HA1 (194–196), is a potential glycosylation site. A single mutation of A196→T, might potentially creates a new glycosylation site at HA1 (194–196), rendering the virus epidemic (Krystal et al., 1983).

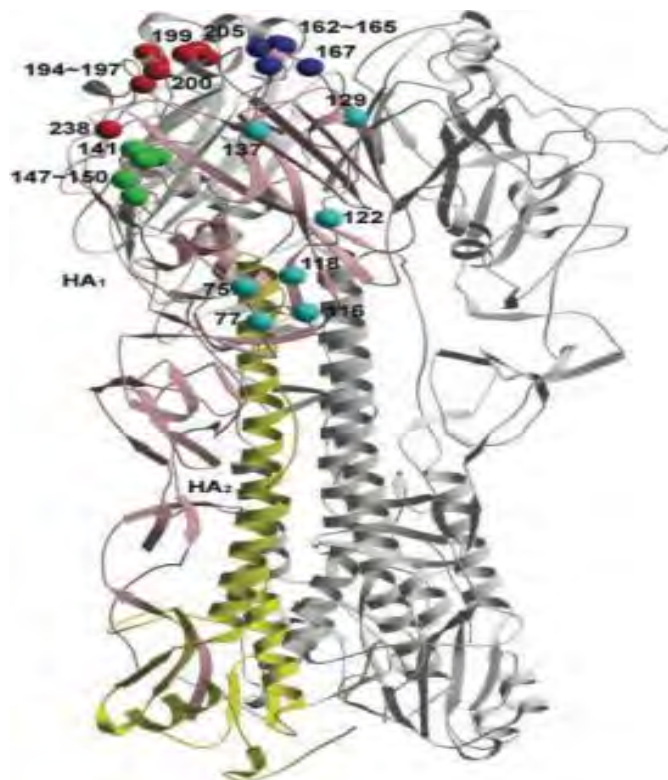


Figure 3: Epitopes of Influenza B virus HA

The trimeric HA is shown with one monomer highlighted in colour: Pink for HA1 and yellow for HA2. Mutations in these four regions, the 120-loop (cyan), 150-loop (green), 160-loop (blue), and 190-helix (red), have been found to cause antigenicity in influenza B virus lineages.

Source: (J. Shen, Kirk, Ma, & Wang, 2009).

b. Neuraminidase and NB

The sixth segment which is 1.4 kb encodes neuraminidase (NA) which destroys sialic acid containing inhibitors for the virus in the mucus secretions of the respiratory tract thus allowing mature virions to be released. The NA is a homotetramer and consists of a globular head, a thin stalk, a transmembrane domain and a cytoplasmic domain (Neumann & Kawaoka, 2011). The NB protein is only found in Influenza B virus which is encoded by segment 6. The NB protein possesses ion channel activity and has been shown to initiate efficient replication in vivo but not in cell culture (Hatta & Kawaoka, 2003).

The neuraminidase active site is a major drug target against influenza viruses. The neuraminidase is a viral enzyme that has three functions use in initiating virus spread (Gubareva, Kaiser, & Hayden, 2000). Firstly, neuraminidase digests neuraminic (sialic) acid in the hemagglutinin receptors thus releasing the virus particles that bud off the host cell membrane. Secondly, the virus particles released from the host cell have haemagglutinin receptors from the host's cell membrane coating them. These newly released virus particles bind to the hemagglutinins of other newly released viruses causing them to clump. The NA cleaves these residues allowing the virus to disperse and infect other cells. Thirdly, NA also digests neuraminic acid in respiratory mucus.

Currently, two classes of influenza antivirals namely adamantanes and neuraminidase inhibitors (NAI) are available to manage influenza. The NAI drugs block the activity of the neuraminidase by binding to the NA active site. However, a mutation at the NA active site leads to resistance to the anti-NA drugs (Samson, Pizzorno, Abed, & Boivin, 2013).

The two neuraminidase inhibitors that are in current use in the treatment of influenza are Oseltamivir (Tamiflu) and zanamivir (Relenza). Zanamivir is used by inhalation while oseltamivir is administered orally. The two antivirals reduce the duration of illness by about

1.5 to 2.5 days and also the severity of disease is modified (Stiver, 2003). In neuraminidase inhibition by antivirals, inhibitor molecules mimic NA's natural substrate and bind to the active site thus preventing NA from cleaving host cell receptors and releasing progeny virus. Rearrangement of amino acids in the active site is essential to accommodate oseltamivir's hydrophobic side chain; mutations that prevent this rearrangement may lead to resistance to oseltamivir. Zanamivir is more structurally similar to the natural substrate of NA enabling it to fit directly into the active site. Mutations that prevent the rearrangement would therefore not bring about resistance to zanamivir (Moscona, 2008).

An amino acid substitution at the conserved NA residues decreases NA enzymatic activity (Yen et al., 2006). The drug will thus no longer be able to bind to the NA active site becoming less effective. The sites where these mutations have been known to occur in influenza B virus NA are E119G, E119A, D198N, D198Y, D198E, I222V/I, I222T, H274Y, N294S, R371K, G402S and R152K (Yen et al., 2006). Another two amino acid deletion at residue 162 and 163 of HAI of influenza B Victoria lineage have been reported by CDC (Lenee Blanton et al., 2017). This site is prone to insertion/deletions and it is the site at which the deletion occurred that differentiated viruses of the B Victoria lineage from those of the B Yamagata lineage whose NA gene fell into a group that shared a K373Q substitution (The Francis Crick Institute, 2016).

2.3 THE INFLUENZA VIRAL DISEASE

The influenza virus is transmitted primarily by droplets or respiratory secretions of infected persons and has incubation period ranging from 2-3 days and after which there is an onset of shivering, malaise, headache, aching in the limbs and back and or other complications of pneumonia (Van Elden et al., 2001). Influenza pneumonia can occur due to viral replication in the epithelial cells of the alveoli leading to rupture of walls of alveoli and bronchioles

causing exudation into the air sacs. Secondary bacterial infection often leads to influenza pneumonia leading to respiratory distress, cyanosis and collapse within 2-3 days of the onset of infection (Rello & Pop-Vicas, 2009). The most common co-infecting bacterial species are *Streptococcus pneumoniae* and *Staphylococcus aureus* (McCullers, 2014).

2.3.1 Clinical presentation of Influenza and its determinants

The most common symptoms of influenza in humans are fever, headache, fatigue and /or weakness, chest discomfort, bronchitis, pneumonia, non-pulmonary complications [myositis, cardiac complications and encephalopathy (Walker et al., 1994)], severe pains or aches, stuffy nose, sore throat, and cough. Unlike common cold infections, influenza is not characterized by runny noses or sore throat at the beginning. Body temperature rises rapidly to about 39 degrees Celsius. Influenza is more severe among the very young (under five years old) and the elderly (above sixty five years of age) (Cox & Subbarao, 2000).

2.3.2 Epidemiology

Seasonally, Influenza disease burden is reported to affect the elderly, very young, and people with high-risk chronic illnesses such as chronic obstructive pulmonary disease (COPD) and diabetes (Cox & Subbarao, 2000; Kaji, Watanabe, & Aizawa, 2003). While this is true in relation to the likelihood of hospitalization and death, from another perspective this is totally disputed. Large cohort studies of respiratory virus infections in families between 1960 and 1970 in the USA have established that the highest annual attack rates for Influenza occur in children and teenagers (Leung et al., 2014). This is especially true for Influenza B, which affects adults far less frequently than children (Aymard, Valette, & Luciani, 2003). While attack rates in the young are highest, infection generally carries fewer consequences in children over 2 years old, teenagers and healthy adults, with correspondingly low levels of morbidity, mortality and hospitalization (Thompson et al., 2003). Reports from the European

Centre for Disease Prevention and Control (ECDC) and World Health Organization (WHO) revealed that there are some Influenza B Victoria deletion variants circulating in China and United States, respectively. These deletion variant viruses currently referred to “deletion subgroup”, possess the following amino acid (AA) changes; I180V, R498K, and 162-164 deletions. These viruses are antigenically distinct from B/Brisbane/60/2008-like viruses (L. Blanton et al., 2017). Therefore, seasonal influenza is somewhat paradoxical with the highest attack rates in the young but the greatest public health impact in the elderly.

In temperate climates, the burden of influenza disease is believed to exist at a low level throughout the year but exhibits a marked seasonal increase, typically during the colder months. Influenza peak seasons happen between November-March in the Northern Hemisphere and during April-October in the Southern Hemisphere (Freedman & Leder, 2005). In these regions, influenza-related deaths contribute approximately 5% of all winter mortality in persons over 65 years of age (Shaman & Kohn, 2009). Influenza epidemics and outbreaks occur in tropical areas as well, although the timing and impact is not as well defined (Shek & Lee, 2003). Most studies suggest that Influenza in tropical regions circulate throughout the year with slight increases during the colder months (Freedman & Leder, 2005).

2.4 THE INFLUENZA SURVEILLANCE IN GHANA

Surveillance of influenza in Ghana started in 2007 following the avian influenza A (H5N1) virus outbreak in May 2007 with the initial coverage of three regions namely Greater Accra, Volta and Brong-Ahafo regions all near the Military establishments (NIC-Ghana, unpublished data, 2014). The establishment of the surveillance system was through the collaborative effort of the Ghana Health Service (GHS), Noguchi Memorial Institute for Medical Research (NMIMR) and the World Health Organization (WHO). Data from the

surveillance system informed policy and so to gain a nation-wide coverage of influenza virus activity, the surveillance system was expanded to all the ten regions. Currently, there are twenty-seven (27) sentinel sites comprising of sixteen (16) GHS sentinel sites and eleven (11) Military reception stations (Figure 4). The aim of the surveillance is to monitor and characterize circulating influenza viruses for policy formulation. The surveillance system was instrumental during the 2009 influenza A (H1N1) pandemic detecting an estimated 1150 influenza viruses with 964 (84%) influenza A (H1N1) pdm09 viruses (NIC-Ghana, unpublished data). Bonney et al., 2012, analyzed about 2810 respiratory specimen from children less than 11years from 2008 through to 2010 and found that 636 were positive for influenza virus by real-time polymerase chain reaction (RT-PCR) (Bonney et al., 2012). Data from the National Influenza Centre (NIC) shows that, from 2007 through to 2017 over 2100 cases of influenza-like-illness have been detected with 294 (14%) being positive for influenza virus. Preliminary data from the NIC has also shown the co-circulation of Influenza B variants since 2010 and until 2013 when the sub-type began, both lineages of Influenza B have indicated (Figure 5) an alternating occurrence over the period of 2013 to second quarter of 2018 (NIC-Ghana, unpublished data, 2018).



Figure 4: A map of Ghana showing the locations of sentinel sites as at December 2017

The regions representing the northern zone for the purpose of this study are Upper West, Upper East, Northern, and Brong Ahafo regions of Ghana and that of the southern zone are Ashanti, Western, Central, Eastern Volta and Greater Accra regions of Ghana.

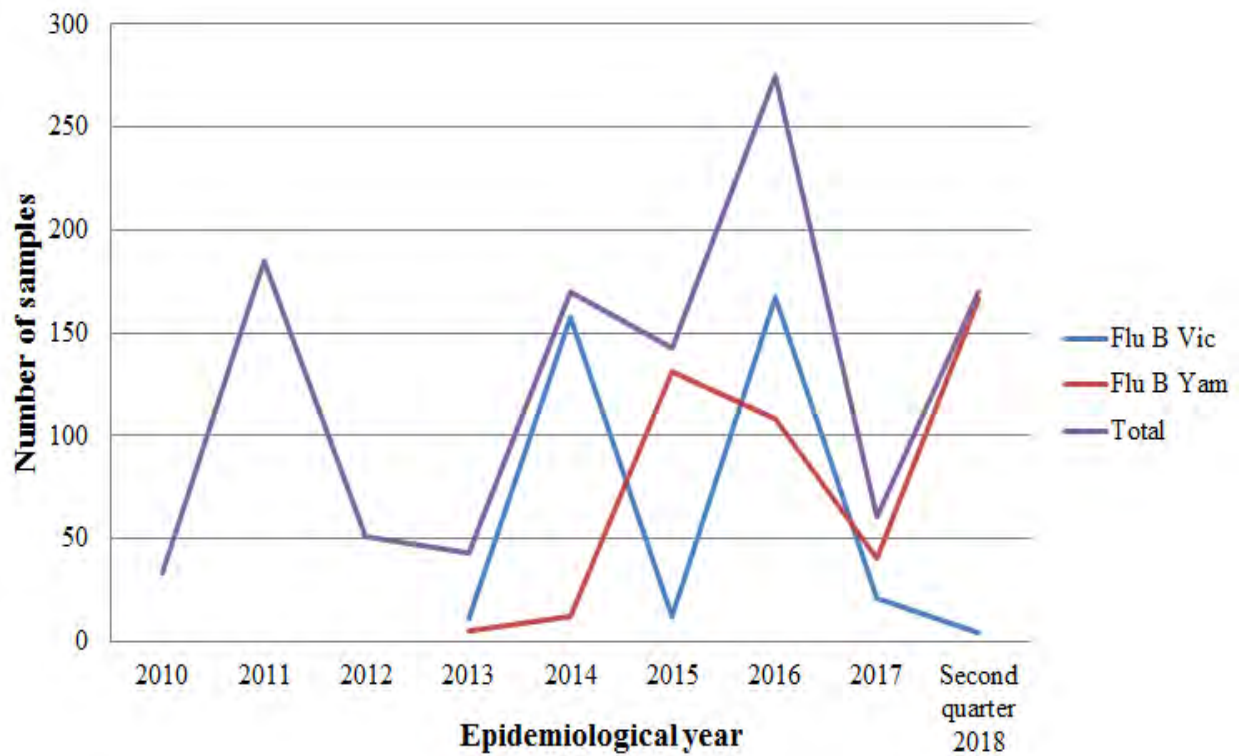


Figure 5: Occurrence of Influenza B real time RT-PCR positives from 2010 to 2018

The figure shows the trend of Influenza B virus circulation in Ghana since 2010. Influenza B Victoria lineage positives are shown in blue while Influenza B Yamagata shown in red with the sum of both lineages in violet.

Source: NIC-Ghana, 2018.

CHAPTER THREE

3.0 METHODS

3.1 Research Design

This was a retrospective cross-sectional study that utilized archived Influenza B positive clinical samples collected from 2016-2017 as part of the national Influenza surveillance across Ghana.

3.2 Study Sites and Sample Collection

Samples were obtained from males and females who tested positive for influenza B during 2016 and 2017 influenza seasons. Secondary data such as age, sex, cough, body temperature (fever), sore throat, myalgia, headache and difficulty in breathing were retrieved from the case based forms for the samples (Appendix I). The sites in which specimen were selected from are shown in table 2.

3.3. Inclusion Criteria

Influenza B virus confirmed samples collected during 2016-2017 and recording Ct values ≤ 29 .

3.4 Exclusion Criteria

Influenza B virus confirmed samples collected outside the period of 2016-2017, and / or those collected during 2016-2017 but with Ct values > 29 .

3.5 Sample Selection

Samples were conveniently selected based on their concentrations (Cycle threshold, Ct values) from the real-time RT-PCR assay. These samples had Ct values ≤ 29 . In total, 8 specimens were selected; 4 from the south and the other 4 from the North, all representing 2016 and 2017 samples (Table 1 and Appendix II).

Table 1: Cycle threshold (Ct) values obtained from real time RT-PCR

Sample ID	Strain Type	Ct value
B/Ghana/FS/1688/2016	B/Vic	17.8
B/Ghana/FS/1980/2016	B/Vic	19.0
B/Ghana/ARI/0005/2017	B/Vic	21.9
B/Ghana/ARI/0090/2017	B/Vic	21.9
B/Ghana/FS/0730/2016	B/Yam	27.0
B/Ghana/FS/1912/2016	B/Yam	20.1
B/Ghana/FS/0747/2017	B/Yam	23.9
B/Ghana/FS/0009/2017	B/Yam	21.2

Threshold cycle (CT) is the calculated cycle number at which the PCR product crosses a threshold of detection. The CV value is inversely proportional to the amount of nucleic acid present in real time PCR assay. Therefore, the lower CT values are indicative of abundant target viral nucleic acid for sequencing.

3.6 Laboratory Procedure

3.6.1 Genomic Detection of Influenza Viruses

Viral RNA was extracted using the QIAmp viral RNA kit (Qiagen) according to the manufacturer's instructions. RNA extracts were amplified by RT-PCR targeting the Influenza B HA genes. Agarose gel electrophoresis was used to confirm the amplicons. The amplicons were then purified and cycle sequenced. The cycle sequencing products were purified and loaded onto the genetic analyzer.

3.6.1.1 Viral Ribonucleic Acid (RNA) Extraction

Extraction of viral RNA was done using QIAmp Viral RNA extraction mini kit (Qiagen, Germany) following the manufacturer's protocol. Each frozen aliquot of oropharyngeal or nasopharyngeal swab was gently thawed at room temperature (22⁰C – 28⁰C) and vortexed. One hundred and forty microliters (140 µl) of the sample was pipetted and added to 560 µl lysis buffer in an appropriately labelled microcentrifuge tube. The mixture was then vortexed for 15 seconds and incubated for 10 minutes. Five hundred and sixty (560 microliters) of cold absolute ethanol was added to the tube and mixed by pulse-vortexing for 15 seconds. Six hundred and thirty microliters (630 µl) of the mixture was transferred on to the spin column and centrifuged at 8000 rpm for 1 minute in an Eppendorf 5415D centrifuge (Eppendorf AG 22331 Hamburg). The collection tube was discarded and the spin column placed into a new collection tube. The procedure was repeated for the remaining mixture in the microcentrifuge tube. Five hundred microliters (500 µl) of wash buffer 1 (AW1) was added to the spin column and centrifuged at 8000 rpm for 1 minute. The collection tube was discarded and the spin column placed into a new collection tube. Five hundred microliters (500 µl) of wash buffer 2 (AW2) was added to the spin column and centrifuged at 14000 rpm for 3 minutes. The spin column was spun again in a new collection tube at 14000 rpm for 1 minute. The collection

tube was discarded and the spin column placed in an appropriately labeled Eppendorf tube. Sixty microliters (60 μ l) of elution buffer was pipetted on to the spin column and incubated for 1 minute. The tube was centrifuged at 8000 rpm for 1 minute. The procedure was repeated for other clinical samples. The RNA extracts were stored at 2⁰C to 8⁰C.

3.6.1.2 Conventional RT-PCR protocol for Influenza B HA gene

The haemagglutinin genes from the selected samples were amplified using the QIAGEN OneStep RT-PCR kit (QIAGEN, USA). The whole HA gene (\approx 1800 bp) was amplified using 8 pairs of primers. Primers (Appendix X) and primer targets spanning the whole HA gene are displayed in figure 6. The reagent formula and protocol for amplification conditions were in line with the CDC protocol adopted by the NIC. The reaction mix was prepared by mixing 5.0 μ l of the 5x QIAGEN OneStep RT-PCR Buffer, 0.5 μ l of the forward primer (10 μ M), 0.5 μ l of the reverse primer (10 μ M), 1.0 μ l of dNTP Mix, 0.5 μ l QIAGEN OneStep RT-PCR Enzyme Mix and 12.5 μ l of nuclease-free water to make a total of 20 μ l. An amount of 5 μ l RNA extract (template) was added to make up a total reaction volume of 25 μ l. Amplification was carried out on an ESCO Thermal Cycler (ESCO Microplate Ltd, Singapore) by the following conditions; 1 cycle of reverse transcription at 50⁰C for 30 minutes follow by an initial denaturation at 95⁰C for 5 minutes, 45 cycles of 94⁰C for 35 seconds, 55⁰C for 30 seconds and 72⁰C for 30 seconds. The amplicons were then stored at 4⁰C for further analysis.

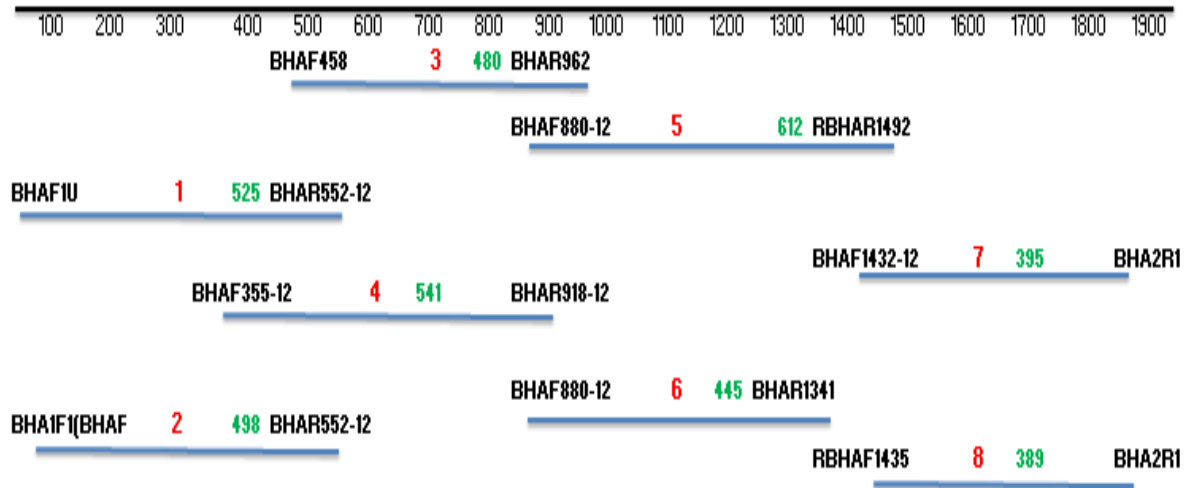


Figure 6: A schematic of primer locations on the Influenza B HA gene

A schematic representation of Influenza B HA gene showing overlapping fragments and their expected sizes (in green) in base pairs. All sixteen primers and their corresponding 8 fragments are shown. Figures written in red indicate fragment numbers.

3.6.1.3 Agarose Gel electrophoresis of the amplicons

An agarose gel of concentration 1.5% stained with ethidium-bromide was prepared in 1X Tris-acetate EDTA (TAE) buffer. A 100 bp DNA ladder (Promega., Switzerland) was loaded alongside amplicons to aid in the confirmation of the expected sizes. The gel was visualized using UV-Transilluminator (BioDoc-it™ 220 imaging system; Upland, CA, USA). A representative gel image is shown in appendix XI.

3.6.1.4 Post PCR purification

Purification of PCR products was done using QIAquick PCR purification kit following the manufacturer's protocol. Both binding buffer (PB) and wash buffer (PE) were prepared and five (5) volumes of buffer PB was added to volume of the PCR amplicon. For each specimen, the content was mixed and transferred into a spin-column fitted into a collection tube (flow-

through tube). The sample in the column with the collection tube was centrifuged for 60 seconds at 13,200 rpm in a microcentrifuge (Eppendorf AG 22331 Hamburg). The flow through was discarded and 750 ul PE buffer added to the column and centrifuged at 8000 rpm for 60 secs. The column was placed in a 1.5 milliliter collection tube and 50 ul elution buffer (EB) added to elute the DNA for onward cycle sequencing.

3.6.1.5 Cycle sequencing

Cycle sequencing was done using the BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, USA). Eight fragments of the HA gene were sequenced for each sample making a total of 64 fragments and 128 sequencing reads. Master mix for various primers (Appendix X) were prepared using the manufacturer's instruction. In brief, a 10 µl reaction volume was prepared consisting 2.0 ul of nuclease free water (NFW) 2.0 ul of 5x Sequencing Buffer, 2.0 ul of BigDye terminator, 2.0 ul of forward (F) or reverse (R) primer (same used for the amplification) and 2.0 ul of the purified DNA (Appendix 8). The mixture was then loaded onto the ESCO Thermal Cycler (ESCO Microplate Ltd, Singapore). The cycling conditions were at 94⁰C 2 minutes / (94⁰C/ 30 seconds; 50⁰C/ 15 seconds; 60⁰C/ 4minutes)x 25 cycles/ 4⁰C hold. The cycle-sequenced products were purified to remove excess primers, dNTPs and other unwanted products using Agencourt CleanSeq Dye Terminator Removal Kit and procedure (Beckman Coulter, Inc. 250 S. Kraemer Blvd. Brea, CA 92821 U.S.A.).

3.6.1.6 Post cycle sequencing purification

Eight microliters (8.0 ul) of magnetic beads was added to the cycle sequenced products followed by 42 ul of 85% ethanol and kept on the magnetic field for 3 minutes. With the PCR tube still sitting in the magnetic field after 3 minutes, 100 ul of 85% ethanol was added to each tube and kept on the magnetic field for another 3 minutes. All the ethanol was aspirated

from the tube and the tube left to dry for 10 minutes. A total of 50 ul) nuclease-free water was added and the tube was removed from the magnetic field onto a plastic rack for 5 minutes. This was to ensure that the magnetic beads fell off the walls of the PCR tube and sediment at the bottom of the tube. The tubes were placed back onto the magnetic field for another 2 minutes after which 40 ul of the clear solution (water and DNA) was pipetted into a 96 well plate. The plate was loaded onto the 3130xl Genetic analyzer (ABI 3130xl, Applied Biosystems, USA) to read the sequences.

3.7 Gene assembling, multiple sequence analyses and phylogeny

Eight fragments (Appendix III) of Influenza B HA gene of the eight specimens were sequenced using the ABI3130XL Genetic Analyzer. Fragments of the real-time RT-PCR identified influenza B Victoria and Yamagata viruses were assembled (Appendix IV) to the reference sequences of B/Brisbane/60/2008 and B/Phuket/3073/2013, respectively, using the Geneious software version 10.2.3 (Kearse et al., 2012). Prior to the assembly, the sequences were viewed in Chromas software, (version 2.6.2) to detect background noise of the base calls by the genetic analyser (Appendix V). Corresponding sequences were edited on the Bioedit platform, version 7.2.5. Assembled sequences were aligned with at least one sequence from each continent available on the National Center for Biotechnology Information (NCBI) or the Global Initiative on Sharing All influenza Data (GISAID) databases. Multiple sequence alignment was carried out using ClustalW in BioEdit with a bootstrap replicates of 1000 in line with Edgar (2004b) (appendices VI and VII). The analysis of antigenic drift was done for both lineages using the WHO vaccine candidate virus sequences for 2018-2019 influenza season (B/Brisbane/60/2008, the Victoria lineage and B/Phuket/3073/2013, the Yamagata lineage).

Phylogenetic analyses were done using Molecular Evolutionary Genetic Analysis version seven (MEGA7) (Kumar, Stecher, & Tamura, 2016). The phylogenetic tree was then visualized using FigTree, version 1.4.3 (Rambaut & Drummond, 2007). Sequences generated from this study are accessible in GenBank under accession numbers MH748708 – MH748715.

3.8 Data handling

To ensure confidentiality, all patient identifiers were removed and the samples were coded with laboratory IDs. All data obtained from this study would be made available by formal request to the investigator or the supervisors.

3.9 Ethical Consideration

This study was approved by College of Health Sciences Ethical and Protocol Review Committee (CHS-Et/M.3-P 2.7/2017-2018) and Noguchi Memorial Institute for Medical Research Institutional Review Board (NMIMR-IRB CPN 062/17-18). Copies of the approval documents are available in Appendix VIII and IX.

CHAPTER FOUR

4.0 RESULTS

4.1 Study Participants

The study population included both males and females, whose oropharyngeal and/or nasopharyngeal specimen was positive for Influenza B with real time RT-PCR at the national influenza centre in 2016 and 2017. Seven (7) of specimens selected were from males (Table 2). The clinical presentation included fever (88%), cough (75%), sore throat (25%) and breathing difficulty (25%) (Table 2). None of the patients whose specimen were selected showed symptom of myalgia.

Table 2: Demographics and clinical symptoms of study participants

Sample ID	Strain Type	Location/ Region	Age	Sex	Fever	Cough	Sore throat	Breathing difficulty	Myalgia
B/Ghana/FS/1688/2016	B/Vic	Ashanti	7	M	Y	Y	Y	N	N
B/Ghana/FS/1980/2016	B/Vic	Ashanti	2	M	Y	Y	N	U	N
B/Ghana/ARI/0005/2017	B/Vic	Upper East	7	M	Y	N	N	Y	N
B/Ghana/ARI/0090/2017	B/Vic	Upper East	2	M	Y	Y	N	N	N
B/Ghana/FS/0730/2016	B/Yam	Greater Accra	26	M	Y	N	N	Y	N
B/Ghana/FS/1912/2016	B/Yam	Ashanti	1	M	Y	Y	N	N	N
B/Ghana/FS/0747/2017	B/Yam	Upper West	2	M	Y	Y	N	N	N
B/Ghana/FS/0009/2017	B/Yam	Brong Ahafo	26	F	N	Y	Y	N	N
%					87.50%	75%	25%	25%	0%

Key: M=male, F=female, Y=yes, N=no, U=unknown.

4.2 Fragments Amplification

All the eight fragments amplified successfully for all the specimen with the exception of specimen 7 that had five fragments amplified. In all sixty-one (61) fragments were amplified (Table 3).

Table 3: Summary of amplification

		Sample IDs							
		T1	T2	T3	T4	T5	T6	T7	T8
Primer pair	1	X	X	X	X	X	X	X	X
	2	X	X	X	X	X	X	O	X
	3	X	X	X	X	X	X	X	X
	4	X	X	X	X	X	X	O	X
	5	X	X	X	X	X	X	X	X
	6	X	X	X	X	X	X	X	X
	7	X	X	X	X	X	X	O	X
	8	X	X	X	X	X	X	X	X
Total fragment amplified		8	8	8	8	8	8	5	8

Note: This table indicates the combinations of primers (1-8) that were used in the amplification of each of the specimens (T1 to T8). Where as X indicates a successful amplification, O indicates no amplification.

4.3 Nucleotide changes distinguishing Influenza B Victoria from Yamagata lineages

The figure below highlight the nucleotide changes that distinguish influenza B Victoria from Yamagata.

	500	510	520	530	540
B/Brisbane/60/2008	GGATT	TTTCGCAACA	ATGGCTTGGG	CCGTCCCAA	AAACGACAAA AACAA
B/Ghana/FS/1688/2016	GGATT	CTTCGCAACA	ATGGCTTGGG	CCGTCCCAA	AAACGACAAA AACAA
B/Ghana/FS/1980/2016	GGATT	CTTCGCAACA	ATGGCTTGGG	CCGTCCCAA	AAACGACAAA AACAA
B/Ghana/ARI/0005/2017	GGATT	CTTCGCAACA	ATGGCTTGGG	CCGTCCCAA	AAACGACAAA AACAA
B/Ghana/ARI/0090/2017	GGATT	CTTCGCAACA	ATGGCTTGGG	CCGTCCCAA	AAACGACAAA AACAA
B/Wisconsin/01/2010	GGATT	TTTTGCAACA	ATGGCTTGGG	CTGTCCCAA	GGACAACATA --CAA
B/Massachusetts/02/2012	GGATT	TTTCGCAACA	ATGGCTTGGG	CTGTCCCAA	GGACAACATA --CAA
B/Ghana/FS/0730/2016	GGATT	TTYGCAACA	ATGGCTTGGG	CTGTTCCAAA	GGACAACATA --CAA
B/Ghana/FS/1912/2016	GGATT	TTTCGCAACA	ATGGCTTGGG	CTGTTCCAAA	GGACAACATA --CAA
B/Ghana/FS/0747/2017	GGATT	TTTCGCAACA	ATGGCTTGGG	CTGTTCCAAA	GGACAACATA --CAA
B/Ghana/FS/0009/2017	GGATT	TTYGCAACA	ATGGCTTGGG	CTGTTCCAAA	GGACAACATA --CAA

	550	560	570	580	590
B/Brisbane/60/2008	AACAG	CAACAAATCC	ATTAACAATA	GAAGTACCAT	ACATTTGTAC AGAAG
B/Ghana/FS/1688/2016	AACAG	CAACAAATCC	ATTAACAATA	GAAGTACCAT	ACATTTGCAC AGAAG
B/Ghana/FS/1980/2016	AACAG	CAACAAATCC	ATTAACAATA	GAAGTACCAT	ACATTTGCAC AGAAG
B/Ghana/ARI/0005/2017	AACAG	CAACAAATCC	ATTAACAATA	GAAGTACCAT	ACATTTGCAC AGAAG
B/Ghana/ARI/0090/2017	AACAG	CAACAAATCC	ATTAACAATA	GAAGTACCAT	ACATTTGCAC AGAAG
B/Wisconsin/01/2010	AAATG	CAACGAACCC	ACTAACAGTA	GAAGTACCAT	ACATTTGTAC AGAAG
B/Massachusetts/02/2012	AAATG	CAACGAACCC	ATTAACAGTA	GAAGTACCAT	ACATTTGTAC AGAAG
B/Ghana/FS/0730/2016	AAATG	CAACGAACCC	ACAAACAGTG	GAAGTACCAT	ACATTTGTAC AGAAG
B/Ghana/FS/1912/2016	AAATG	CAACGAACCC	ACAAACAGTG	GAAGTACCAT	ACATTTGTAC AGAAG
B/Ghana/FS/0747/2017	AAATG	CAACGAACCC	ACAAACAGTG	GAAGTACCAT	ACATTTGTAC AGAAG
B/Ghana/FS/0009/2017	AAATG	CAACGAACCC	ACAAACAGTG	GAAGTACCAT	ACATTTGTAC AGAAG

Figure 7: Lineage specific markers of Influenza B HA gene

Multiple sequence alignment was carried out using “ClustalW” in BioEdit with a bootstrap replicates of 1000 in line with Edgar (2004b). Influenza B Victoria lineages from this study were referenced to B/Brisbane/60/2008 while that of Yamagata lineages were referenced to B/Wisconsin/1/2010, Clade 3 and B/Massachusetts/2/2012, Clade 2. Nucleotide changes distinguishing the influenza B Victoria from that of Influenza B Yamagata lineages are shown in yellow while clade specific markers are shown in green. About 68 nts long are shown in the figure with lineages specific markers (nts 522, 540-542, 548, 549, 555, 558 and 568) are shown in yellow. The clade specific (nts 538, 562 and 589) however are shown in green.

4.4 Analysis of Antigenic drift on Influenza B Victoria HA gene

After comparing the recently circulating Ghanaian strains from this study with that of the 2016-2017 WHO recommended trivalent vaccine candidate virus sequences

(B/Brisbane/60/2008), a total of eleven (11) amino acid substitutions were detected in the HA gene. Out of these, five (5) amino acids substitutions and six (6) amino acids substitutions were detected in the HA1 and HA2 respectively (Table 4). All the Ghanaian specimens had P31S and I117V mutations in HA1 and R151K mutation in HA2. Two previously sequenced Ghanaian Influenza B Victoria specimens retrieved from GISAID, and included in the analysis, also showed the above mutations. Influenza B/Ghana/FS/1688/2016 and B/Ghana/FS/1980/2016 both from the Ashanti Region have the highest amino acid substitution of six while Influenza B/Ghana/ARI/0090/2017 from the Upper East Region had the least amino acid substitution of three. The two Ghanaian reference sequences added have four amino acid substitutions at the same position as compared to the study sequences. Details of the mutations detected in the individual specimens are shown in (Table 4).

For the HA2, apart from all of the sequences showing the R151K amino acid substitution, three samples each had N53T, E82D, I92K, E97K and S103A amino acid substitutions respectively. None of the sequences from this study were found to be among the deletion subgroup (162, 163 and 164 deletions).

Table 4: Amino acid substitutions in the HA genes of Influenza B Victoria lineage

Virus Strains	HA1 Amino acid at position					HA2 Amino acid at position					
	31	117	129	278	283	53	82	92	97	103	151
B/Brisbane/60/2008	P	I	N	G	I	N	E	I	E	S	R
*B/Ghana/FS/1688/2016	S	V	-	A	-	-	-	K	-	A	K
*B/Ghana/FS/1980/2016	S	V	D	-	-	-	D	-	R	-	K
*B/Ghana/ARI/0005/2017	S	V	-	-	R	T	-	-	-	-	K
*B/Ghana/ARI/0090/2017	S	V	-	-	-	-	-	-	-	-	K
B/Ghana/DILI-16-1091/2016	S	V	D	-	-	-	-	-	-	-	K
B/Ghana/FS-16-1620/2016	S	V	D	-	-	-	-	-	-	-	K

Key: S (Serine), P (Proline), I (Isoleucine), V (Valine), N (Asparagine), D (Aspartic acid), G (Glycine), A (Alanine), K (Lysine), R (Arginine), T (Threonine), E (Glutamic acid), - (Consensus to references sequence), * Sequences obtained from this study.

4.5 Analysis of Antigenic drift on Influenza B Yamagata HA gene

The Influenza B Yamagata lineage HA gene from this study had less amino acid substitutions compared to Influenza B Victoria lineage. When the sequences from this study were compared with that of the 2016-2017 WHO quadrivalent vaccine candidate virus sequences (B/Phuket/3073/2013), it was detected that, Influenza B/Ghana/FS/1912/2016 from Ashanti Region and Influenza B/Ghana/FS/0747/2017 from Upper West region had the highest amino acid substitutions of four, while majority of the sequences had three amino acid substitutions. All the sequences of this study including the two reference Ghanaian sequences have L172Q and M251V amino acid substitutions within the HA1 gene. Only one specimen, Influenza

B/Ghana/FS/0747/2017 from Upper West region had an I150S amino acid substitution within the HA1 gene. The WHO vaccine candidate virus of Influenza B/Phuket/3073/2013, the Yamagata lineage, showed an amino acid degeneracy at position I/T198T within the HA1 gene to all the specimens from this study and the Ghanaian reference sequences. Majority (three out of four) of the sequences from this study, including the two Ghanaian reference sequences had N158D amino acid substitution within the HA2 gene. In addition, there was a detection of two different amino acids substitutions at the same position on the HA2 gene. Influenza B/Ghana/FS/0730/2016 had an E72K amino acid substitution while Influenza B/Ghana/FS/1912/2016 had an E72Q amino acid substitution. There was no detection of 162, 163 and 164 deletions among the sequences from this study. The details of mutations in the study samples in relation to the reference sequence B/Phuket/3073/2013 are shown in Table 5.

Table 5: Amino acid substitutions in the HA genes of Influenza B Yamagata lineages

Virus Strains	HA1 Amino acid at position				HA2 Amino acid at position	
	150	172	198	251	76	158
B/Phuket/3073/2013	I	L	TI	M	E	N
*B/Ghana/FS/0730/2016	-	Q	-	V	K	-
*B/Ghana/FS/1912/2016	-	Q	-	V	Q	D
*B/Ghana/FS/0747/2017	S	Q	-	V	-	D
*B/Ghana/FS/0009/2017	-	Q	-	V	-	D
B/Ghana/DILI-16-11149-2016	-	Q	-	V	-	D
B/Ghana/532/2017	-	Q	-	V	-	D

Key: S (Serine), I (Isoleucine), V (Valine), N (Asparagine), D (Aspartic acid), K (Lysine), T (Threonine), E (Glutamic acid), Q (Glutamine), L (Leucine), M (Methionine), - (Consensus to reference sequence). At position 198, the study and the Ghanaian reference sequences have T, * (sequences obtained from this study).

4.6 Glycosylation analysis

The analysis of glycosylation of the Influenza B Victoria (Table 6) revealed that one of the specimens (B/Ghana/ARI/0005/2017) had gained a glycosylation site at position 51 of the HA2 gene relative to the vaccine sequences (B/Brisbane/60/2008). Another gain of glycosylation site at position 196 of the HA1 gene of Influenza B Yamagata lineage was observed for all the Ghanaian sequences (both study and reference) but absent in the vaccine candidate virus sequence due to amino acid degeneracy (Table 7). Detailed presentations of amino acid alignment showing the potential N-glycosylation sites identified for both Influenza B Victoria and Influenza B Yamagata lineages are attached at appendices VI and VII.

Table 6: Potential Glycosylation sites of Influenza B Victoria HA genes

Strain	Number of glycosylation sites	HA1 Amino acid position								HA2 Amino acid position				
		25	59	145	166	197	233	304	333	51	145	171	184	216
B/Brisbane/60/2008	12	NVT	NCT	NVT	NKT	NET	NQT	NKS	NCT	NLN	NQT	NIT	NHT	NVS
*B/Ghana/FS/1688/2016	12	-	-	-	-	-	-	-	-	-	-	-	-	-
*B/Ghana/FS/1980/2016	12	-	-	-	-	-	-	-	-	-	-	-	-	-
*B/Ghana/ARI/0005/2017	13	-	-	-	-	-	-	-	-	NLT	-	-	-	-
*B/Ghana/ARI/0090/2017	12	-	-	-	-	-	-	-	-	-	-	-	-	-
B/Ghana/DILI-16-1091/2016	12	-	-	-	-	-	-	-	-	-	-	-	-	-
B/Ghana/FS-16-1620/2016	12	-	-	-	-	-	-	-	-	-	-	-	-	-

Key: N (Asparagine), V (Valine), T (Threonine), C (Cysteine), K (Lysine), E (Glutamic acid), Q (Glutamine), S (Serine), L (Leucine), I (Isoleucine), H (Histidine), - (Consensus), * (sequences obtained from this study).

Table 7: Potential Glycosylation sites of Influenza B Yamagata HA genes

Strain	Number of glycosylation sites	HA1 Amino acid position							HA2 amino acid position			
		25	59	145	167	196	303	332	145	171	184	216
B/Phuket/3073/2013	10/11	NVT	NCT	NAT	NAT	NKT (I)	NKS	NCT	NQT	NIT	NHT	NVS
*B/Ghana/FS/0730/2016	11	-	-	-	-	NKT	-	-	-	-	-	-
*B/Ghana/FS/1912/2016	11	-	-	-	-	NKT	-	-	-	-	-	-
*B/Ghana/FS/0747/2017	11	-	-	-	-	NKT	-	-	-	-	-	-
*B/Ghana/FS/0009/2017	11	-	-	-	-	NKT	-	-	-	-	-	-
B/Ghana/DILI-16-11149-2016	11	-	-	-	-	NKT	-	-	-	-	-	-
B/Ghana/532/2017	11	-	-	-	-	NKT	-	-	-	-	-	-

Key: N (Asparagine), V (Valine), T (Threonine), C (Cysteine), A (Alanine), K (Lysine), I (Isoleucine), S (Serine), Q (Glutamine), H (Histidine), - (Consensus). The vaccine sequences had amino acid degeneracy (T/I) at position 198. If T is present, glycosylation will be possible otherwise with I, no glycosylation, * (sequences obtained from this study).

4.7 Phylogenetic analysis

4.7.1 Phylogenetic analysis of Influenza B Victoria lineage

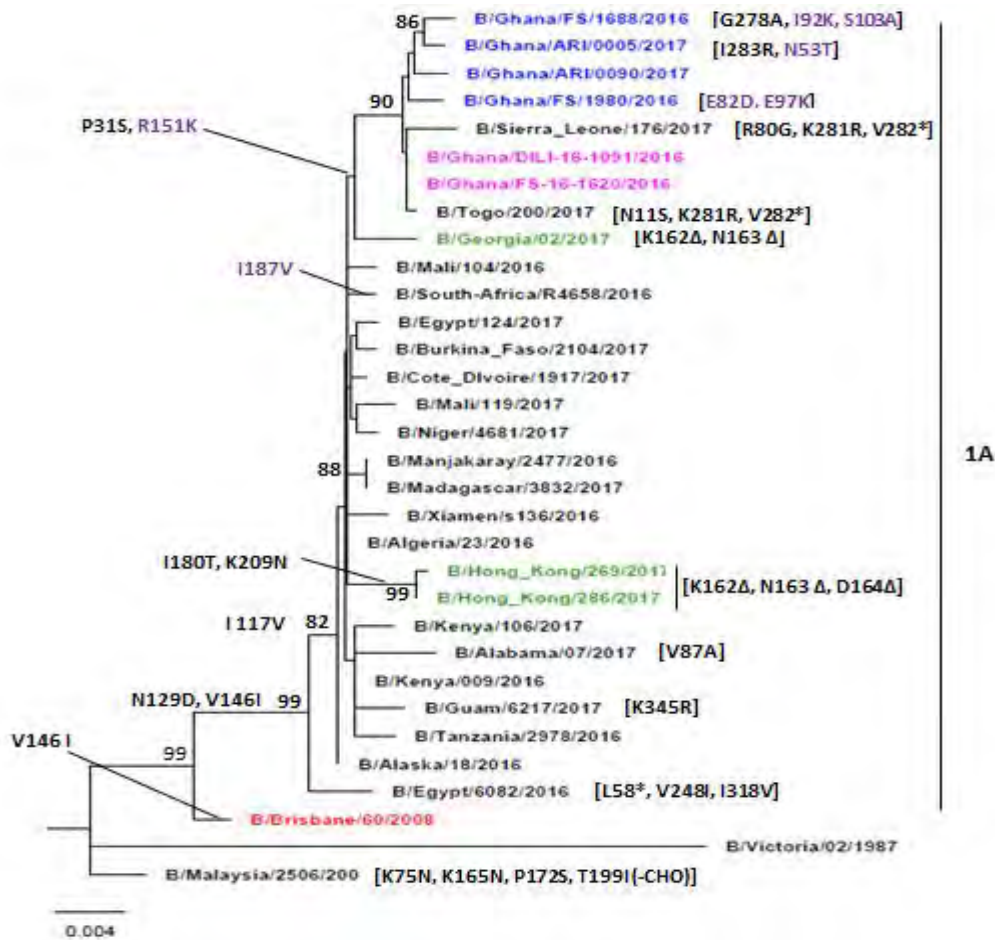


Figure 8: Phylogenetic analysis of Influenza B Victoria lineage using HA genes

The phylogenetic analysis was done using MEGA (Kumar et al., 2016). The percentage (greater than 80%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985).

The sequence identified in red font represented the WHO vaccine candidate virus genome, those in the pink font represents reference Ghanaian specimens sequenced at the Francis Crick Institute, the sequence identified in blue font represent the study sequences, green font sequences represent the deletion sub-group, violet font sequences represent the HA2 numbering.

All the HA genes analyzed generally fell into genetic group 1A, the B/Brisbane/60/2008 genetic group (the WHO recommended candidate vaccine), which is defined by V146I amino acid substitution. However, sequences from this study and the reference Ghanaian sequences clustered together into a unique genetic group defined by two HA1 amino acid substitutions, P31S and I117V,

compared with B/Brisbane/60/2008: R151K in the HA2 gene. One of the sequences from this study, including the two Ghanaian reference sequences also had a N129D amino acid substitution.

4.7.2 Phylogenetic analysis of Influenza B Yamagata lineage

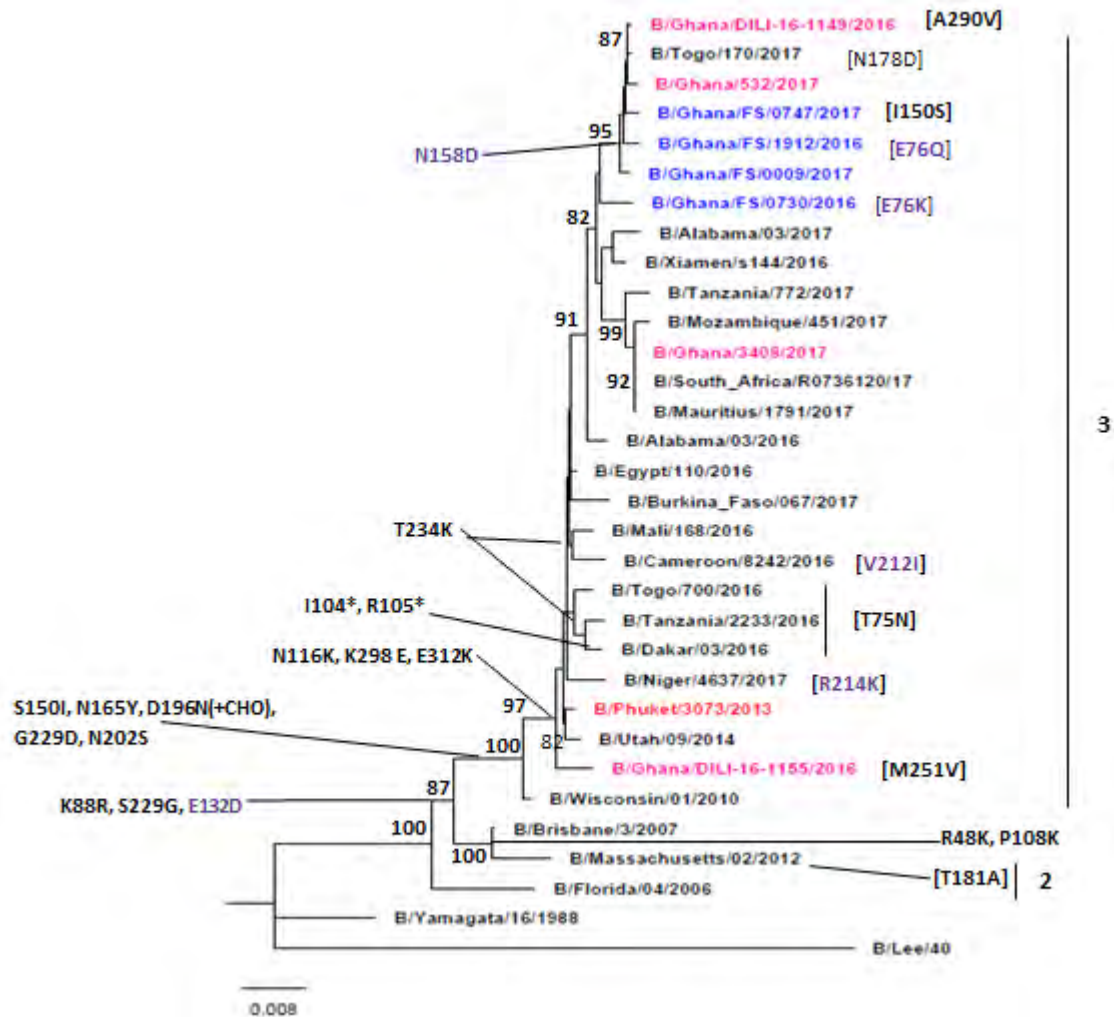


Figure 9: Phylogenetic analysis of Influenza B Yamagata lineage using the HA genes

The phylogenetic analysis was done using MEGA (Kumar et al., 2016). The percentage (greater than 80%) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein, 1985).

The red font represented the WHO Vaccine candidate virus genome, pink font represents reference Ghanaian specimens sequenced at the Francis Crick Institute, Blue font represents sequences from this study, and violet font represents the HA2 numbering.

All the HA genes analyzed generally fell into genetic clade 3, the B/Phuket/3073/2013 clade, with three clusters into another genetic group defined by N158D of the HA2 gene and one falling into E76K amino acid substitutions.

CHAPTER FIVE

5.0 DISCUSSION AND CONCLUSIONS

5.1 Discussion

This study sought to describe the Influenza B Victoria and Yamagata lineages circulating in Ghana between 2016 and 2017 by sequencing their HA genes. Characterization by sequencing was done using eight selected specimens representative of both lineages of Influenza B Victoria and Yamagata across the country. Molecular characterization of Influenza B is vital in the assessment of similarities and or differences between reference viruses and circulating viruses. This characterization is also useful to WHO in formulating vaccine components that will ensure vaccine effectiveness in the upcoming influenza season. However, due to mutations within the influenza virus genome, there is always a challenge in achieving this goal.

All the nucleotide changes that distinguished Influenza B Victoria from that of Influenza B Yamagata lineages from this study are in line with reports by Arvia et al (2014), except for nucleotide positions 540-542 and 538, which are at variance. Antigenic variation in influenza B virus is mostly caused by amino acid substitutions at four major antigenic epitopes (120-loop, 150-loop, 160 loop and 190- helix). These epitopes have been identified in previous studies (Pechirra et al., 2005). Six (6) amino acids substitutions were observed in the HA1 region of Influenza B Victoria lineages of the Ghanaian strains (study and reference) relative to Influenza B/Brisbane/60/2008, the 2016-2017 trivalent and now 2017-2018 quadrivalent influenza vaccine component defined by the amino acid substitutions N75K, N165K and S172P on the HA1 coding region of the HA gene of Influenza B Victoria lineage (Byarugaba et al., 2013). These amino acid substitutions are P31S, 1117V, N129D, G278A, K281R, and

I283R. However, only two of these substitutions (I117V and N129D) could be found in the antigenic sites around loop120 of HA1 coding region. All the sequences from this study, including the reference sequences from Ghana showed the P31S and I117V substitutions. The 129 residue is reported to have altered antigenicity which might affect vaccine efficacy (Qinghua Wang et al., 2008). The amino acid substitutions P31S, I117V and N129D of sequences from this study were similar to the Ghanaian reference sequences included in these analyses. It was also observed that at least each sequence from this study has a unique amino acid substitution. For example, only influenza B/Ghana/FS/1688/2016 has G278A amino acid substitution relative to the vaccine component; whereas influenza B/Ghana/ARI/0005/2017 has an I283R amino acid substitution.

Another amino acid substitution R151K of HA2 (R498K) coding region was observed similar to that observed in the reference Ghanaian sequences. In all, five (5) amino acid substitutions were observed in all four sequences. Four (4) other amino acid substitutions were identified to be unique to only three sequences from this study. Influenza B/Ghana/ARI/0005/2017 has a N53T in the HA2 (N400T), influenza B/Ghana/FS/1980/2016 was detected to have two amino acid substitutions at residues E82D and E97R in the HA2 (E429D) and (E444R) and Influenza B/Ghana/ FS/1688/2016 also detected to have S103A amino acid substitution in the HA2 (S450A). These substitutions may have an effect on the fusion ability of Influenza B HA2 domain due to change from a non-charged polar amino acid (serine) to a non-charged non-polar amino acid (alanine)(Wang, 2010).

For the influenza B Yamagata sequences, all the four (4) viruses sequenced, in addition to the Ghanaian reference viruses sequenced by the WHO CC fell into a group defined by the HA substitutions, L172Q and M251V in HA1 and N158D in HA2 (N504D) from B/Phuket/3073/2013 in clade 3 (the B/Wisconsin/1/2010-B/Phuket/3073/2013 clade) except Influenza B/Ghana/FS/0730/2016 which did not show any amino acid change at N158 in

HA2 (N504). Apart from the above, there were parallel amino acid substitutions observed in three (3) of the sequences from this study which were not found in the Ghanaian reference sequences. These amino acid substitutions can be found in Influenza B/Ghana/FS/0747/2017 at I150S of the HAA1 domain, while the other two are Influenza B/Ghana/FS/0730/2016 and Influenza B/Ghana/FS/1912/2016 at residues E76K and E76Q in the HA2 (E422K and E422Q) respectively. Amino acid degeneracy was also observed at I/T198T relative to the WHO vaccine candidate (Influenza B/Phuket/3073/2013).

The presence of glycosylation site acts as a natural barrier to defending immune cells and antibodies such that it is difficult for the natural immune system to recognize and target the influenza virus for elimination. Glycosylation plays a critical role in determining protein structure, function and stability. Structurally, glycosylation affects the three dimensional configuration of proteins (Arey, 2012).

Glycosylation is a post-translational modification involving addition of sugar moieties to proteins. Influenza virus glycoproteins are thus a classic example of such modifications. In influenza viruses, carbohydrate moieties are mostly found on the R-amino group of asparagine, (N) in a close succession with another amino acid except proline, (P) and serine, (S) or threonine, (T), thus N-X-S/T. This is known to affect the antigenic character of proteins (Wilson & Cox, 1990).

In this study, the glycosylation pattern revealed by influenza B Victoria lineages of the Ghanaian origin and the other Ghanaian reference sequences were similar to Influenza B/Brisbane/60/2008 except for Influenza B/Ghana/ARI/0005/2017 which showed a gain in glycosylation site at position 51 of the HA2 (398) domain. This gain of a glycosylation site is not likely to unmask new antigenic sites because this site is not reported to be an epitope on the HA protein. However, evidence for the importance of carbohydrates in modulating antigenicity was presented by selection of a mutant HA (D197N) which resulted in a new

glycosylation site that prevented antibody binding and viral neutralization (Wilson & Cox, 1990). The glycosylation pattern of influenza B Yamagata lineages from this study was similar to the reference Ghanaian sequences, but with a gain at position 196 as compared with the vaccine virus sequences (Influenza B/Phuket/3073/2013). This gain of a glycosylation site of the Yamagata lineage is likely to unmask new antigenic site because this site, is located at the epitope 190 helix of the HA1 gene. A single mutation of A196T leads to a creation of a new potential glycosylation site at HA1 (194 – 196) domain rendering the virus epidemic (Qinghua Wang et al., 2008). It is worthy of note that none of the deletion variants reported by the WHO CC was identified among the Ghanaian sequences from this study (Lenee Blanton et al., 2017).

Phylogenetic analyses of both influenza B Victoria and Yamagata lineages (Fig 2 and 3 respectively) have shown that, the influenza B Victoria sequences from this study clustered together and belong to the vaccine virus sequences (B/Brisbane/60/2008). Three of the influenza B Yamagata sequences from this study clustered together with the reference Ghanaian sequences while the fourth sequence clustered separately. These clustering patterns are similar to previous reports of Ghanaian sequences between 2016 and 2017 flu season (WHO, 2017c).

5.2 Conclusions

Influenza B Victoria lineages that circulated in Ghana between 2016 and 2017 were found to have 3 main substitutions (P31S, I117V and R151K). One of the influenza B Victoria sequences (B/Ghana/ARI/0005/2017) analysed had gained a unique glycosylation site at amino acid position 51 in the HA2 subunit, this was absent from all other sequences reported. The influenza B Yamagata sequences analysed, had 2 main substitutions, namely L172Q and M251V, representing the nature of this lineage of viruses in circulation during the period for which sampling was done.

Yamagata and Victoria sequences analysed showed that there were only minor drifts, but were genetically similar to the vaccine reference strains, B/Phuket/3073/2013 and B/Brisbane/60/2008, respectively. Furthermore, the recent deletion sub-group in Influenza B virus reported by the US CDC was not identified among the specimens analysed.

RECOMMENDATIONS

The following are recommended based on the results of the current study:

1. The National Influenza Centre (NIC-Ghana) should consider conducting this study on a larger sample size to provide a detailed representation of influenza B viruses in Ghana.
2. Whole genome sequencing should be done by the NIC to provide information regarding circulating reassortant variants in Ghana.
3. The Disease Surveillance Unit of Ghana Health Service must focus more on the influenza surveillance at the National level.

REFERENCES

- Alcami, A., & Koszinowski, U. H. (2000). Viral mechanisms of immune evasion. *Immunol Today*, 21(9), 447-455.
- Ambrose, C. S., & Levin, M. J. (2012). The rationale for quadrivalent influenza vaccines. *Human vaccines & immunotherapeutics*, 8(1), 81-88.
- Arey, B. J. (2012). The role of glycosylation in receptor signaling *Glycosylation: InTech*.
- Arvia, R., Corcioli, F., Pierucci, F., & Azzi, A. (2014). Molecular markers of influenza B lineages and clades. *Viruses*, 6(11), 4437-4446.
- Aymard, M., Valette, M., & Luciani, J. (2003). Burden of influenza in children: preliminary data from a pilot survey network on community diseases. *Pediatr Infect Dis J*, 22(10 Suppl), 211-214. doi: 10.1097/01.inf.0000092189.42748.cc
- Belshe, R. B., Coelingh, K., Ambrose, C. S., Woo, J. C., & Wu, X. (2010). Efficacy of live attenuated influenza vaccine in children against influenza B viruses by lineage and antigenic similarity. *Vaccine*, 28(9), 2149-2156.
- Blanton, L., Alabi, N., Mustaquim, D., Taylor, C., Kniss, K., Kramer, N., . . . Chung, J. (2017). Update: Influenza Activity in the United States During the 2016-17 Season and Composition of the 2017-18 Influenza Vaccine. *MMWR. Morbidity and mortality weekly report*, 66(25), 668-676.
- Blanton, L., Alabi, N., Mustaquim, D., Taylor, C., Kniss, K., Kramer, N., . . . Brammer, L. (2017). Update: Influenza Activity in the United States During the 2016-17 Season and Composition of the 2017-18 Influenza Vaccine. *MMWR Morb Mortal Wkly Rep*, 66(25), 668-676. doi: 10.15585/mmwr.mm6625a3
- Bonney, J. H., Kronmann, K. C., Lindan, C. P., Asante, I. A., Parbie, P., Aboagye, J., . . . Nzussouo, N. T. (2012). Virological surveillance of influenza-like illness among

- children in Ghana, 2008–2010. *The Journal of infectious diseases*, 206(suppl_1), S108-S113.
- Brassard, D. L., Leser, G. P., & Lamb, R. A. (1996). Influenza B Virus NB Glycoprotein Is a Component of the Virion. *Virology*(220), 350 – 360.
- Burton, D. R., Poignard, P., Stanfield, R. L., & Wilson, I. A. (2012). Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. *Science*, 337(6091), 183-186.
- Byarugaba, D. K., Erima, B., Millard, M., Kibuuka, H., Lukwago, L., Bwogi, J., . . . Krauss, S. (2013). Genetic analysis of influenza B viruses isolated in Uganda during the 2009–2010 seasons. *Virology journal*, 10(1), 11.
- CDC. (2011). Influenza-associated pediatric deaths--United States, September 2010–August 2011. *MMWR. Morbidity and mortality weekly report*, 60(36), 1233.
- Chen, J., Lee, K. H., Steinhauer, D. A., Stevens, D. J., Skehel, J. J., & Wiley, D. C. (1998). Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation. *Cell*, 95(3), 409-417.
- Chen, R., & Holmes, E. C. (2008). The evolutionary dynamics of human influenza B virus. *Journal of molecular evolution*, 66(6), 655.
- Cox, N. J. a., & Subbarao, K. (2000). Global epidemiology of influenza: past and present. *Annu Rev Med*, 51, 407-421. doi: 10.1146/annurev.med.51.1.407
- ECDC. (2017). Influenza virus characterisation.
- Edgar, R. C. (2004b). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 32(32), 1792-1797.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783-791.
- Freedman, D. O., & Leder, K. (2005). Influenza: changing approaches to prevention and treatment in travelers. *Journal of travel medicine*, 12(1), 36-44.

- Gessner, B. D., Shindo, N., & Briand, S. (2011). Seasonal influenza epidemiology in sub-Saharan Africa: a systematic review. *Lancet Infect Dis*, *11*(3), 223-235.
- Glezen, W. P. (2014). Editorial commentary: Changing epidemiology of influenza B virus. *Clin Infect Dis*, *59*(11), 1525-1526. doi: 10.1093/cid/ciu668
- Gloria, J. G. (2011). *Molecular Characterization of the Surface Glycoproteins of Influenza B Viruses Isolated in Kenya from 2011-2012*. University of Nairobi.
- Gubareva, L. V., Kaiser, L., & Hayden, F. G. (2000). Influenza virus neuraminidase inhibitors. *The Lancet*, *355*, 827-835.
- Hatta, M., & Kawaoka, Y. (2003). The NB protein of influenza B virus is not necessary for virus replication in vitro. *J Virol*, *77*, 6050-6054.
- Hause, B. M., Collin, E. A., Liu, R., Huang, B., Sheng, Z., Lu, W., . . . Li, F. (2014). Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *MBio*, *5*(2), e00031-00014.
- Hyland, L., Webby, R., Sandbulte, M. R., Clarke, B., & Hou, S. (2006). Influenza virus NS1 protein protects against lymphohematopoietic pathogenesis in an in vivo mouse model. *Virology*, *349*(1), 156-163.
- Jackson, D., Elderfield, R. A., & Barclay, W. S. (2011). Molecular studies of influenza B virus in the reverse genetics era. *Journal of General Virology*, *92*(1), 1-17.
- Jagger, B. W., Wise, H. M., Kash, J. C., Walters, K. A., Wills, N. M., Xiao, Y. L., . . . Digard, P. (2012). An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science*, *337*(6091), 199-204. doi: 10.1126/science.1222213
- Kaji, M., Watanabe, A., & Aizawa, H. (2003). Differences in clinical features between influenza A H1N1, A H3N2, and B in adult patients. *Respirology*, *8*(2), 231-233.

- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649.
- Kim, H. (2012). Mechanisms for the enhancement of influenza B virus vaccine antigen yield.
- Krejcová, L., Michálek, P., Hynek, D., Adam, V. a., & Kizek, R. (2015). Structure of influenza viruses, connected with influenza life cycle. *Journal of Metallomics and Nanotechnologies*.
- Krystal, M., Young, J. F., Palese, P., Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1983). Sequential mutations in hemagglutinins of influenza B virus isolates: definition of antigenic domains. *Proceedings of the National Academy of Sciences*, 80, 4527-4531.
- Kumagai, Y., Takeuchi, O., & Akira, S. (2008). Pathogen recognition by innate receptors. *J Infect Chemother*, 14(2), 86-92. doi: 10.1007/s10156-008-0596-1
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, 33(7), 1870-1874.
- Leung, C. H., Tseng, H. K., Wang, W. S., Chiang, H. T., Wu, A. Y., & Liu, C. P. (2014). Clinical characteristics of children and adults hospitalized for influenza virus infection. *J Microbiol Immunol Infect*, 47(6), 518-525. doi: 10.1016/j.jmii.2013.06.002
- McAuley, J. L., Hornung, F., Boyd, K. L., Smith, A. M., McKeon, R., Bennink, J., . . . McCullers, J. A. (2007). Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell host & microbe*, 2(4), 240-249.
- McCullers, J. A. (2014). The co-pathogenesis of influenza viruses with bacteria in the lung. *Nature Reviews Microbiology*, 12(4), nrmicro3231.

- Min, J.-Y., Li, S., Sen, G. C., & Krug, R. M. (2007). A site on the influenza A virus NS1 protein mediates both inhibition of PKR activation and temporal regulation of viral RNA synthesis. *Virology*, *363*(1), 236-243.
- Moscona, A. (2008). Medical management of influenza infection. *Annu Rev Med*, *59*, 397-413. doi: 10.1146/annurev.med.59.061506.213121
- Neumann, G., & Kawaoka, Y. (2011). Influenza viruses: molecular virology. *Wiley Online Library*.
- Nieto, A., de la Luna, S., Barcena, J., Portela, A., & Ortin, J. (1994). Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. *J Gen Virol*, *75*(1), 29-36. doi: 10.1099/0022-1317-75-1-29
- Ntiri, M. P., Duque, J., McMorrow, M. L., Frimpong, A. J., Parbie, P., Badji, E., . . . Ampofo, W. (2016). Incidence of medically attended influenza among residents of Shai-Osudoku and Ningo-Prampram Districts, Ghana, May 2013 – April 2015. *BMC Infectious Diseases*, *16*(1), 757.
- Palese, P. (1977). The genes of influenza virus. *Cell*, *10*(1), 1-10.
- Pechirra, P., Nunes, B., Coelho, A., Ribeiro, C., Goncalves, P., Pedro, S., . . . Rebelo-de-Andrade, H. (2005). Molecular characterization of the HA gene of influenza type B viruses. *J Med Virol*, *77*(4), 541-549. doi: 10.1002/jmv.20490
- Portela, A., & Digard, P. (2002). The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J Gen Virol*, *83*(Pt 4), 723-734. doi: 10.1099/0022-1317-83-4-723
- Rambaut, A., & Drummond, A. (2007). FigTree. Molecular evolution, phylogenetics and epidemiology. Dostupné z <http://tree.bio.ed.ac.uk>.
- Reaney, D. (1982). The evolution of RNA viruses. *Annual Reviews in Microbiology*, *36*, 47-73.

- Rello, J., & Pop-Vicas, A. (2009). Clinical review: primary influenza viral pneumonia. *Crit Care*, 13(6), 235. doi: 10.1186/cc8183
- Rota, P. A., Wallis, T. R., Harmon, M. W., Rota, J. S., Kendal, A. P., & Nerome, K. (1990). Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. *Virology*, 175(1), 59-68.
- Samson, M., Pizzorno, A., Abed, Y., & Boivin, G. (2013). Influenza virus resistance to neuraminidase inhibitors. *Antiviral Res*, 98(2), 174-185. doi: 10.1016/j.antiviral.2013.03.014
- Shaman, J., & Kohn, M. (2009). Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci U S A*, 106(9), 3243-3248. doi: 10.1073/pnas.0806852106
- Shaw, M. W., Choppin, P. W., & Lamb, R. A. (1983). A previously unrecognized influenza B virus glycoprotein from a bicistronic mRNA that also encodes the viral neuraminidase. *Proceedings of the National Academy of Sciences*, 80(16), 4879-4883.
- Shek, L. P., & Lee, B. W. (2003). Epidemiology and seasonality of respiratory tract virus infections in the tropics. *Paediatr Respir Rev*, 4(2), 105-111.
- Shen, J., Kirk, B. D., Ma, J., & Wang, Q. (2009). Diversifying selective pressure on influenza B virus hemagglutinin. *Journal of medical virology*, 81, 114-124.
- Shen, S., Mahadevappa, G., Oh, H. L., Wee, B. Y., Choi, Y. W., Hwang, L. A., . . . Tan, Y. J. (2008). Comparing the antibody responses against recombinant hemagglutinin proteins of avian influenza A (H5N1) virus expressed in insect cells and bacteria. *J Med Virol*, 80(11), 1972-1983. doi: 10.1002/jmv.21298
- Shi, L., Galarza, J. M., & Summers, D. F. (1996). Recombinant-baculovirus-expressed PB2 subunit of the influenza A virus RNA polymerase binds cap groups as an isolated subunit. *Virus Res*, 42(1-9).

- SIB. (2012). Influenza B virus. *ViralZone*. Retrieved 31 July, 2017
- Steinhauer, D. A., & Skehel, J. J. (2002a). Genetics of influenza viruses. *Annual Review of Genetics*, 36, 305-332.
- Steinhauer, D. A., & Skehel, J. J. (2002b). Genetics of influenza viruses. *Annual Review of Genetics*. 36, 305-332.
- Stiver, G. (2003). The treatment of influenza with antiviral drugs. *CMAJ*, 168(1), 49-56.
- Tan, J., Arunkumar, G. A., & Krammer, F. (2018). Universal influenza virus vaccines and therapeutics: where do we stand with influenza B virus? *Current opinion in immunology*, 53, 45-50.
- Thompson, W. W., Shay, D. K., Weintraub, E., Brammer, L., Cox, N., Anderson, L. J., & Fukuda, K. (2003). Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA*, 289(2)(2), 179-186.
- Van Elden, L., Van Essen, G., Boucher, C., Van Loon, A., Nijhuis, M., Schipper, P., . . . Hoepelman, I. (2001). Clinical diagnosis of influenza virus infection: evaluation of diagnostic tools in general practice. *Br J Gen Pract*, 51(469), 630-634.
- Walker, J. A., Molloy, S. S., Thomas, G., Sakaguchi, T., Yoshida, T., Chambers, T. M., & Kawaoka, Y. (1994). Sequence specificity of furin, a proprotein-processing endoprotease, for the hemagglutinin of a virulent avian influenza virus. *J Virol*, 68(2), 1213-1218.
- Wang, Q. (2010). Influenza type B virus haemagglutinin: antigenicity, receptor binding and membrane fusion. *Influenza: Molecular Virology*, 29-52.
- Wang, Q., Cheng, F., Lu, M., Tian, X., & Ma, J. (2008). Crystal structure of unliganded influenza B virus hemagglutinin. *J Virol*, 82(6), 3011-3020. doi: 10.1128/JVI.02477-07
- Wang, Q., Cheng, F., Lu, M., Tian, X., & Ma, J. (2008). Crystal structure of unliganded influenza B virus hemagglutinin. *Journal of virology*, 82(6), 3011-3020.

- Watzinger, F., Suda, M., Preuner, S., Baumgartinger, R., Ebner, K., Baskova, L., . . . Lion, T. (2004). Real-time quantitative PCR assays for detection and monitoring of pathogenic human viruses in immunosuppressed pediatric patients. *J Clin Microbiol*, *42*(11), pp.5189-5198.
- Weber, F., Gruber, S., Haller, O., & Kochs, G. (1999). PB2 polymerase subunit of Thogoto virus (Orthomyxoviridae family). *Arch Virol*, *144*(8), 1601-1609.
- WHO. (2011). WHO information for molecular diagnosis of influenza virus in humans - update
- WHO. (2012). Vaccines against influenza. WHO position paper–November 2012. *Wkly Epidemiol Rec*, *87*, 461-476.
- WHO. (2017a). Biologicals/Vaccines/Influenza. Retrieved 16 July 2017
- WHO. (2017b). Influenza Laboratory Surveillance Information. Retrieved 16 July, 2017
- WHO. (2017c). Influenza Laboratory Surveillance Information. Retrieved 16 July, 2017
- Wilson, I. A., & Cox, N. J. (1990). Structural basis of immune recognition of influenza virus hemagglutinin. *Annual review of immunology*, *8*(1), 737-787.
- Yen, H. L., Hoffmann, E., Taylor, G., Scholtissek, C., Monto, A. S., Webster, R. G., & Govorkova, E. A. (2006). Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. *J Virol*, *80*(17), 8787-8795. doi: 10.1128/JVI.00477-06
- Zaman, R. U., Alamgir, A., Rahman, M., Azziz-Baumgartner, E., Gurley, E. S., Sharker, M. A. Y., . . . Lindstrom, S. (2009). Influenza in outpatient ILI case-patients in national hospital-based surveillance, Bangladesh, 2007–2008. *PloS one*, *4*(12), e8452.

APPENDICES

Appendix I: Scanned copy of influenza case investigation form-Ghana

Fill in the Blank Space or Tick the box as appropriate

Case Id Number: GHA - _____ Date Received by National level
 Region District Year Onset Case Number (Disease Surveillance Department): ____/____/____

1. Reporting Details

Region _____ District _____
 Sub-district _____ Health Facility _____
 Date Notified ____/____/____ Date Investigated ____/____/____

2. Demographic Details

Name of patient _____ Sex Male, Female
 Date of birth ____/____/____ Age Years _____ OR Months _____
 (If DOB is unknown) (If <1 year)
 Name of Village/Suburb Town/ City _____ Contact tel. number _____
 Address (Location) _____ Residence Urban Rural
 Occupation _____ Name of School/Workplace _____

3. Signs and Symptoms

Date of onset ____/____/____ (dd/mm/yyyy) Fever/Body temp >38°C Yes No Unknown
 Cough Yes No Unknown Sore throat Yes No Unknown Breathing Difficulty Yes No Unknown
 Others (specify) _____

4. History of Admission

Date first seen at a health facility for this illness ____/____/____
 Was patient admitted? Yes No Unknown
 Name of Health Facility _____ District where Health Facility located: _____
 Date of admission (in-patient) ____/____/____ (dd/mm/yyyy)
 Date person discharged from hospital ____/____/____ (dd/mm/yyyy)
 Termination date of hospital stay ____/____/____ (dd/mm/yyyy)
 Admitted to intensive care unit Yes No Unknown

5. Exposure to Risk Factors

Previously vaccinated against Influenza? Yes No If Yes, specify month/year of vaccination _____
 List places visited during the past 7 days _____

Visited places with known lab confirmed pandemic influenza cases within 7 days prior to the onset of disease? Yes No
 Contact with suspected or confirmed Influenza patient(s)? Yes No
 Contact with sick or dead animals (wild or domestic)? Yes No
 Risk factors for severe disease: Pregnant Diabetic Immuno suppressed Other(s),
 specify _____

6. Laboratory Investigations

Type of specimen taken Oropharyngeal Nasopharyngeal Serum Plasma Date Sample taken ____/____/____
 Date Sample received at National Influenza Centre (Noguchi, Legon) ____/____/____ Date Sample tested ____/____/____
 Positive rapid diagnostic test for influenza virus Flu A Yes Flu B Yes No Unknown
 Positive real time RT-PCR for influenza virus Yes No Unknown
 Positive viral culture for influenza virus Yes No Unknown
 4-fold rise in influenza virus-specific antibody titre in paired serum samples Yes No Unknown
 Subtype of influenza virus:
 A(H1N1) A(H3N2) A(H5N1) A(H3N2) Bvic Byam p(H1N1₂₀₀₉) Typing Pending Un-subtypable,
 Date Results sent to reporting facility ____/____/____ Mode of Initial Communication of results Phone call Email
 Date Results received by reporting facility ____/____/____

Disease Surveillance Department, GHS, Tel: 0302 675142/3; Fax: 0302 660131; E-mail: surveillance_ghana@yahoo.com

7. Treatment

State treatment administered _____

8. Final Outcome

Final Outcome Recovered, Deceased Date of death ___/___/___

Lost to follow-up Transferred Out Still under treatment

Date final outcome established ___/___/___

9. Final classification Confirmed Probable Suspected or under investigation Not a case

Other comments and remarks:

Contact details of Investigator

Name _____

Title _____

Institution/Unit _____

Address _____

Telephone _____

E-mail _____

Date of last update (dd/mm/yyyy): _____

Appendix II: Selection of specimens from the NIC-Ghana confirmed influenza B positives samples during 2016 and 2017

Region	Year	Flu B positive specimens	No. of Flu B Victoria (V)	No. of Flu B Yamagata (Y)	No. of specimens selected at Ct ≤ 29 (V:Y)
Southern	2016-2017	91	56	35	4 (2:2)
Northern	2016-2017	77	50	27	4 (2:2)
Total		168	106	62	8 (4:4)

Note: Eight (8) influenza B positive specimens conveniently selected from the North and South at concentrations recording threshold cycle, Ct ≤ 29 , to increase the chances of efficient sequencing.

Appendix III: Summary of sequencing out come

	Fragment 1		Fragment 2		Fragment 3		Fragment 4		Fragment 5		Fragment 6		Fragment 7		Fragment 8	
	F	R	F	R	F	R	F	R	F	R	F	R	F	R	F	R
T1	0	X	X	X	X	X	X	0	X	X	X	X	X	X	X	X
T2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T3	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T4	X	X	X	X	X	X	X	0	X	X	X	X	0	X	X	X
T5	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T6	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
T7	X	X	N	N	X	X	N	N	X	X	X	X	N	N	X	X
T8	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

NOTE: T = Template/Sample ID, F = Sense, R = Antisense, X = Sequence successful, 0 = Sequence not successful, N = No amplification.

Appendix IV: Fragments assembling using Geneious 10.2.3

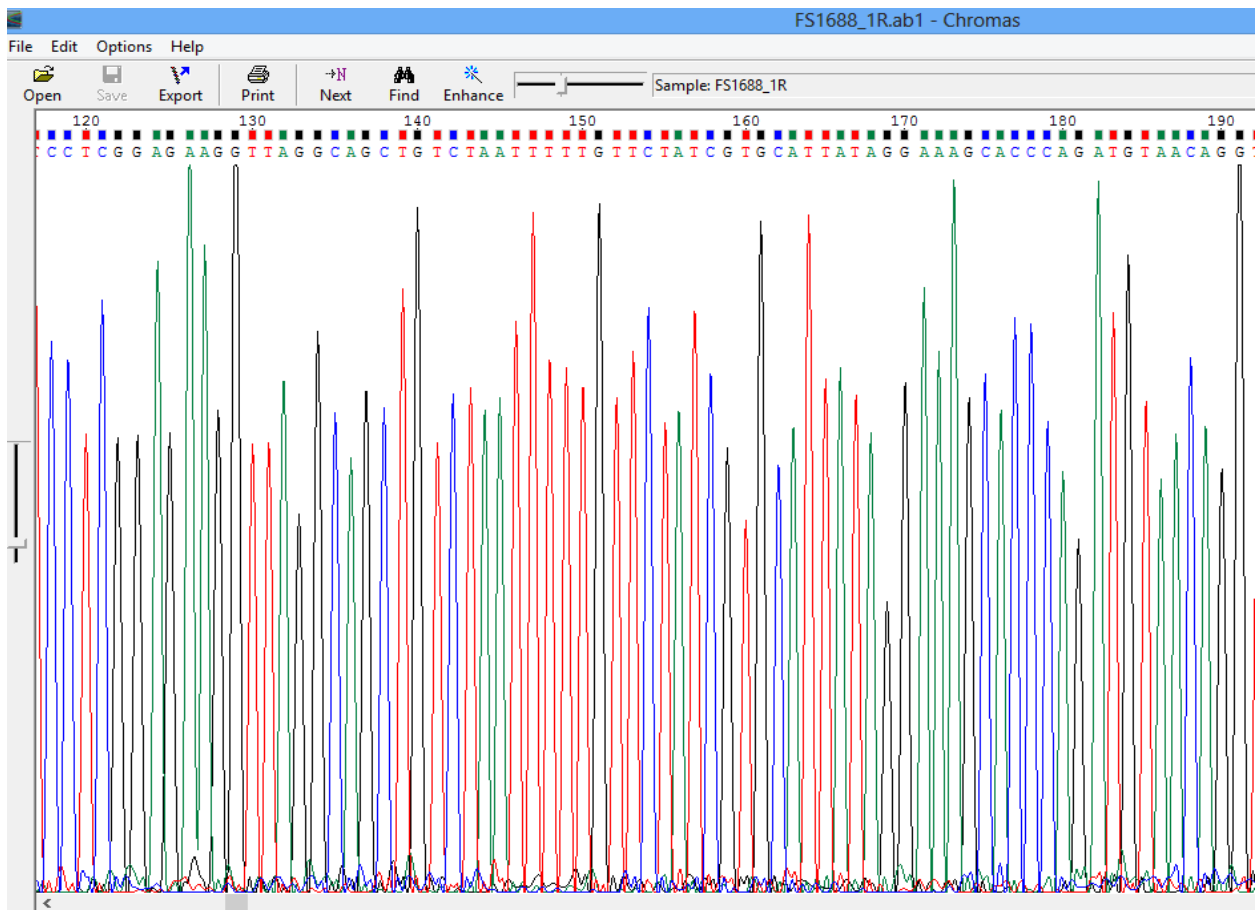
The screenshot displays the Geneious 10.2.3 interface. The top menu bar includes File, Edit, View, Tools, Sequence, Annotate & Predict, and Help. Below the menu is a toolbar with icons for navigation and analysis. The 'Sources' panel on the left shows a project tree with folders like 'Local (0)', 'Sample Documents (0)', 'Alignments (8)', 'Cloning (12)', 'Contig Assembly (7)', 'Genomes (233)', 'PlasMapper Features (314)', 'Plasmids from NEB (27)', 'Primers (12)', 'Protein Documents (6)', 'Tree Documents (4)', 'Test file (5)', 'FS 1980 (15)', 'T1 (20)', 'T2 (17)', 'T3 (18)', 'T4 (18)', 'T5 (18)', 'T6 (18)', 'T7 (12)', 'T8 (18)', 'Deleted Items (179)', 'Shared Databases', 'Operations', 'NCBI', and 'UniProt'.

The main window shows a table of sequence sources with the following columns: Name, Description, HQ%, Sequence Le..., Post-Trim, # Sequences, %GC, Min Sequen..., Max Sequen..., and Path. The table lists several FS1980 sequences and one assembly result for FJ766840.1.

Name	Description	HQ%	Sequence Le...	Post-Trim	# Sequences	%GC	Min Sequen...	Max Sequen...	Path
FS1980_SF.ab1	-	38.5%	628	628	-	43.8%	-	-	C:\Us...
FS1980_SR.ab1 (reversed)	-	2.5%	1,023	1,023	-	45.2%	-	-	C:\Us...
FS1980_6F.ab1	-	88.5%	443	443	-	44.7%	-	-	C:\Us...
FS1980_7F.ab1	-	0.0%	164	164	-	43.3%	-	-	C:\Us...
FS1980_7R.ab1 (reversed)	-	92.2%	387	387	-	40.6%	-	-	C:\Us...
FS1980_8F.ab1	-	33.7%	829	829	-	43.1%	-	-	C:\Us...
FS1980_8R.ab1 (reversed)	-	75.8%	455	455	-	42.2%	-	-	C:\Us...
FS1980 assembled to FJ766... FJ766840.1	11 reads from FS1980 mapped to FJ766840.1 us... Influenza B virus (B/Brisbane/60/2008) segment ...	71.2%	2,329	-	12	43.6%	164	829	C:\Us...

Below the table, the 'Contig View' shows a sequence alignment. The 'Consensus' sequence is: 4TATCAC TGAC AAEAC ACC CACCA AATCT CATTTT GCA AATCTCAAAGGAA CAGAAAC CAGGGGGAAACTATGCC CAAAATGCC TCAACTGCACAGATCTGGAA. The 'Coverage' is shown as a blue bar. The alignment tracks show the original reads (REV and FWD) and their alignment to the consensus. The 'Annotations and Tracks' panel on the right is empty, displaying the message: 'These sequences have no annotations.'

Appendix V: A sample chromatogram of influenza B HA gene fragment



	60	70	80	90	100	110	120	130	140	150
B/Malaysia/2506/2004	N N S L S E L E V K N L Q R L S G A M D E L H N E I L E L D E K V D D L R A D T I S S Q I E L A V L L S N E G I T N S E D E H L L A L E R K L K K M L G P S A V E I G N G C F E T K H K C N Q T C L D									
B/Brisbane/60/2008									
B/Ghana/FS/1688/2016 A									
B/Ghana/FS/1980/2016 D K									
B/Ghana/ARI/0005/2017 T									
B/Ghana/ARI/0090/2017									
B/Ghana/DILI-16-1091/2016									
B/Ghana/FS-16-1620/2016									
B/Victoria/02/1987 K									
B/Georgia/02/2017									
B/Hong Kong/269/2017									
B/Hong Kong/286/2017									
B/Alaska/18/2016									
B/Alabama/07/2017									
B/Kenya/009/2016									
B/Guam/6217/2017									
B/Xiamen/s136/2016 I									
B/Manjalaray/2477/2016									
B/Mali/104/2016									
B/South-Africa/R4658/2016									
B/Tanzania/2978/2016									
B/Egypt/6082/2016									
B/Algeria/23/2016									
B/Egypt/124/2017									
B/Togo/200/2017									
B/Mali/119/2017									
B/Burkina Faso/2104/2017									
B/Niger/4681/2017									
B/Kenya/106/2017									
B/Madagascar/3832/2017									
B/Sierra Leone/176/2017									
B/Cote D'Ivoire/1917/2017									

	160	170	180	190	200	210	220
B/Malaysia/2506/2004	R I A G T F D A C E F S L P T F D S L N I T A A S L N D D G L D N H T I L L Y Y S T A A S S L A V T L M I A I F V V Y M V S R D N V S C S I C L						
B/Brisbane/60/2008						
B/Ghana/FS/1688/2016	K						
B/Ghana/FS/1980/2016	K						
B/Ghana/ARI/0005/2017	K						
B/Ghana/ARI/0090/2017	K						
B/Ghana/DILI-16-1091/2016	K						
B/Ghana/FS-16-1620/2016	K						
B/Victoria/02/1987 N I						
B/Georgia/02/2017	K						
B/Hong Kong/269/2017						
B/Hong Kong/286/2017						
B/Alaska/18/2016						
B/Alabama/07/2017						
B/Kenya/009/2016						
B/Guam/6217/2017						
B/Xiamen/s136/2016						
B/Manjalaray/2477/2016						
B/Mali/104/2016						
B/South-Africa/R4658/2016 V						
B/Tanzania/2978/2016						
B/Egypt/6082/2016						
B/Algeria/23/2016						
B/Egypt/124/2017						
B/Togo/200/2017	K						
B/Mali/119/2017						
B/Burkina Faso/2104/2017						
B/Niger/4681/2017						
B/Kenya/106/2017						
B/Madagascar/3832/2017						
B/Sierra Leone/176/2017	K						
B/Cote D'Ivoire/1917/2017						

Note: The highlighted yellow shows the amino acid positions for potential N-linked glycosylation sites. The blue highlight indicates the HA1/HA2 boundary.

	210	220	230	240	250	260	270	280	290	300	
B/Florida/04/2006	INLYGDSNPK	FTSSANGVT	THVVSQIGS	FDPDQTE	DGGLPQSG	RIVVDYMM	QKPKGT	GTIVYQR	GVLLPQK	VWCASGRSK	VIKGSLPLICEADCLHEKYG
B/Phuket/3073/2013	S		D								E
B/Ghana/FS/0730/2016	S		D		V						E
B/Ghana/FS/1912/2016	S		D		V						E
B/Ghana/FS/0747/2017	S		D		V						E
B/Ghana/FS/0009/2017	S		D		V						E
B/Ghana/DILI-16-1149/2016	S		D		V						E
B/Ghana/DILI-16-1155/2016	S		D		V						E
B/Ghana/532/2017	S		D		V						E
B/Ghana/3408/2017	S		D		V						E
B/Yamagata/16/1988			D	N	V						E
B/Brisbane/3/2007			G								E
B/Wisconsin/01/2010	S		D								E
B/Massachusetts/02/2012			G								E
B/Lee/40	ER		G	N	E	K		V		I	
B/Utah/09/2014	S		D								E
B/Alabama/03/2016	S		D		V						E
B/Alabama/03/2017	S		D		V						E
B/Xiamen/s144/2016	S		D		V						E
B/Egypt/110/2016	S		D		V						E
B/Mali/168/2016	S		D	K	V						E
B/Tanzania/2233/2016	S		D	K	V						E
B/Cameroon/8242/2016	S		I	D	K	V					E
B/Dakar/03/2016	S		D	K							E
B/Togo/700/2016	S		D						T		E
B/Togo/170/2017	S		D		V						E
B/South Africa/R0736120/17	S		D		V						E
B/Burkina Faso/067/2017	S		D		V						E
B/Mauritius/1791/2017	S		D		V						E
B/Niger/4637/2017	S		D		I						E
B/Mozambique/451/2017	S		D		V						E
B/Tanzania/772/2017	S		N	A	V						E

	310	320	330	340	10	20	30	40	50	60										
B/Florida/04/2006	GLIKSKPY	YTGHAHAK	IGNCP	IVKTP	LKLANCT	KYRPPAK	LLKER	GFFGAI	AGFLECG	EGMIAG	HRGYT	SHGARG	VAVAA	ADLAK	STQ	EAINKI	TN	LNS	LSE	LEV
B/Phuket/3073/2013			K																	
B/Ghana/FS/0730/2016			K																	
B/Ghana/FS/1912/2016			K																	
B/Ghana/FS/0747/2017			K																	
B/Ghana/FS/0009/2017			K																	
B/Ghana/DILI-16-1149/2016			K																	
B/Ghana/DILI-16-1155/2016			K																	
B/Ghana/532/2017			K																	
B/Ghana/3408/2017			K																	
B/Yamagata/16/1988																				
B/Brisbane/3/2007																				
B/Wisconsin/01/2010																				
B/Massachusetts/02/2012																				
B/Lee/40																				
B/Utah/09/2014			K																	
B/Alabama/03/2016			K																	
B/Alabama/03/2017			K																	
B/Xiamen/s144/2016			K						V											
B/Egypt/110/2016			K																	
B/Mali/168/2016			K																	
B/Tanzania/2233/2016			K																	
B/Cameroon/8242/2016			K																	
B/Dakar/03/2016			K																	
B/Togo/700/2016			K																	
B/Togo/170/2017			K																	
B/South Africa/R0736120/17			K																	
B/Burkina Faso/067/2017			K																	
B/Mauritius/1791/2017			K																	
B/Niger/4637/2017			K																	
B/Mozambique/451/2017			K																	
B/Tanzania/772/2017			K																	

	70	80	90	100	110	120	130	140	150	160
B/Florida/04/2006	KNLQRLSGAMDELHNEI	LELDEKVDLDRADT	ISSQTE	LAVLLS	NEGIINSEDEHLL	LALERKLLKQMLGPS	AVEIGNGCFETKHK	QNTCL	LDRIAAGTFNAG	
B/Phuket/3073/2013							D			
B/Ghana/FS/0730/2016			K				D			
B/Ghana/FS/1912/2016			Q				D			D
B/Ghana/FS/0747/2017							D			D
B/Ghana/FS/0009/2017							D			D
B/Ghana/DILI-16-1149/2016							D			D
B/Ghana/DILI-16-1155/2016							D			D
B/Ghana/532/2017							D			D
B/Ghana/3408/2017							D			
B/Yamagata/16/1988							D			
B/Brisbane/3/2007							D			
B/Wisconsin/01/2010							D			
B/Massachusetts/02/2012							D			
B/Lee/40			NG	D						
B/Utah/09/2014							D			
B/Alabama/03/2016							D			
B/Alabama/03/2017							D			
B/Xiamen/s144/2016						S	D			
B/Egypt/110/2016							D			
B/Mali/168/2016							D			
B/Tanzania/2233/2016							D			
B/Cameroon/8242/2016							D			
B/Dakar/03/2016							D			
B/Togo/700/2016							D			
B/Togo/170/2017							D			D
B/South Africa/R0736120/17							D			
B/Burkina Faso/067/2017							D			
B/Mauritius/1791/2017							D			
B/Niger/4637/2017							D			
B/Mozambique/451/2017							D			
B/Tanzania/772/2017							D			

	170	180	190	200	210	220
B/Florida/04/2006	EFSLPTFDSLNIT	AASLNDDGLDNHT	ILLYYSTAASSLAVT	LMLAIFIVYMVSRDNV	SCSICL	
B/Phuket/3073/2013						
B/Ghana/FS/0730/2016						
B/Ghana/FS/1912/2016						
B/Ghana/FS/0747/2017						
B/Ghana/FS/0009/2017						
B/Ghana/DILI-16-1149/2016						
B/Ghana/DILI-16-1155/2016						
B/Ghana/532/2017						
B/Ghana/3408/2017						
B/Yamagata/16/1988					I	
B/Brisbane/3/2007						
B/Wisconsin/01/2010						
B/Massachusetts/02/2012						
B/Lee/40	D		D		I	
B/Utah/09/2014						
B/Alabama/03/2016						
B/Alabama/03/2017						
B/Xiamen/s144/2016						
B/Egypt/110/2016						
B/Mali/168/2016						
B/Tanzania/2233/2016						
B/Cameroon/8242/2016					I	
B/Dakar/03/2016						
B/Togo/700/2016						
B/Togo/170/2017						
B/South Africa/R0736120/17						
B/Burkina Faso/067/2017						
B/Mauritius/1791/2017						
B/Niger/4637/2017						K
B/Mozambique/451/2017						
B/Tanzania/772/2017						

Note: The highlighted yellow shows the amino acid positions for potential N-linked glycosylation sites. The blue highlight indicates the HA1/HA2 boundary.

Appendix VIII: Ethical clearance certificate from the College of Health Sciences



UNIVERSITY OF GHANA COLLEGE OF HEALTH SCIENCES

ETHICAL AND PROTOCOL REVIEW COMMITTEE

6th November, 2017.

Ref. No.:

Mr. Alahassan Mohammed Yakubu
Dept. of Medical Biochemistry
SBAHS
Korle-Bu, Accra

ETHICAL CLEARANCE

Protocol Identification Number: **CHS-Et/M.3- P 2.7/2017-2018**

The Ethical and Protocol Review Committee of the College of Health Sciences on the 5th of October, 2017 unanimously approved your research proposal.

TITLE OF PROTOCOL: **“Characterization of Influenza B Virus Variants in Ghana.”**

PRINCIPAL INVESTIGATOR: **Mr. Alahassan Mohammed Yakubu**

This approval requires that you submit six-monthly review reports of the protocol to the Committee and a final full review to the Ethical and Protocol Review Committee at the completion of the study. The Committee may observe, or cause to be observed, procedures and records of the study during and after implementation.

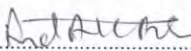
Please note that any significant modification of this project must be submitted to the Committee for review and approval before its implementation.

You are required to report all serious adverse events related to this study to the Ethical and Protocol Review Committee within seven (7) days verbally and fourteen (14) days in writing.

As part of the review process, it is the Committee's duty to review the ethical aspects of any manuscript that may be produced from this study. You will therefore be required to furnish the Committee with any manuscript for publication.

This ethical clearance is valid till 4th November, 2018.

Please always quote the protocol identification number in all future correspondence in relation to this protocol.

Signed: 

PROFESSOR ANDREW A. ADJEI
CHAIRPERSON, ETHICAL AND PROTOCOL REVIEW COMMITTEE

cc: Provost, CHS
Dean, SBAHS
Head of Department

Appendix IX: Ethical clearance certificate from the Noguchi Memorial Institute for Medical Research

NOGUCHI MEMORIAL INSTITUTE FOR MEDICAL RESEARCH

Established 1979A Constituent of the College of Health Sciences

Phone: +233-302-916438 (Direct)
+233-289-522574
Fax: +233-302-502182/513202
E-mail: nirb@noguchi.ug.edu.gh
Telex No: 2556 UGL GH

INSTITUTIONAL REVIEW BOARD



University of Ghana

Post Office Box LG 581
Legon, Accra
Ghana

My Ref. No: DF.22
Your Ref. No:

10th January, 2018

ETHICAL CLEARANCE

FEDERALWIDE ASSURANCE FWA 00001824

IRB 00001276

NMIMR-IRB CPN 062/17-18

IORG 0000908

On 10th January, 2018, the Noguchi Memorial Institute for Medical Research (NMIMR) Institutional Review Board (IRB) at a full board meeting reviewed and approved your protocol titled:

TITLE OF PROTOCOL : Characterization of Influenza B Virus Variants in Ghana

PRINCIPAL INVESTIGATOR : Alhassan Mohammed Yakubu, Mphil Cand.

Please note that a final review report must be submitted to the Board at the completion of the study. Your research records may be audited at any time during or after the implementation.

Any modification of this research project must be submitted to the IRB for review and approval prior to implementation.

Please report all serious adverse events related to this study to NMIMR-IRB within seven days verbally and fourteen days in writing.

This certificate is valid till 9th January, 2019. You are to submit annual reports for continuing review.

Signature of Chair:

Mrs. Chris Dadzie
(NMIMR – IRB, Chair)

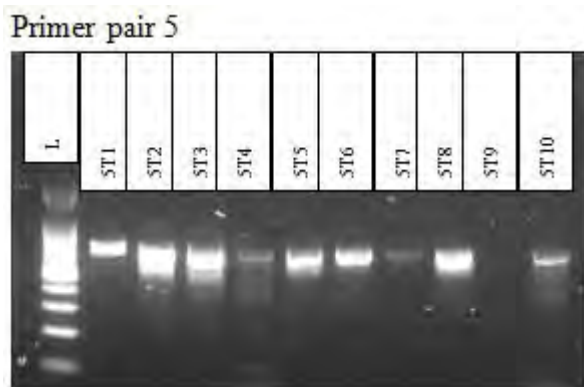
Appendix X: Detailed of primers used for this research

PRIMER ID	PRIMER SEQUENCES (5' - 3')	LOCATION ON HA GENE
BHAF1U*	TATTCGTCTCAGGGAGCAGAAGCAG	1-27
BHA1F1 (BHAF22)	AATATCCACAAAATGAAGGCAATA	22-45
BHAF355-12	GGGTGCTTYCCTATAATGCACGA	355-377
BHAF458	AGAAAAGGCACCAGGAGGACCCTA	459-482
BHAF880-12	TTRTTGCCTCAAAAGGTGTGGTG	880-902
BHAF1432-12	GAAGGAATAATAAACAGTGAAGA	1432-1454
RBHAF1435	GGAATAATAAACAGTGAAGAYGAGCA	1435-1460
BHAR552-12	CCAAGCCATTGTTGCGAARAATCC	552-529
BHAR918-12	GCTCCTGCCACTYGCGCACCA	918-898
BHAR962	TGGAGGCAATCTGCTTCACC	962-943
BHAR1341	TTCGTTGTGGAGTTCATCCAT	1347-1327
RBHAR1492	TCTACAGCAGAGGGRCCCAGGAT	1514-1492
BHA2R1	GTAATGGTAACAAGCAAACAAGCA	1849-1826

*-Primer sequences have been modified (previous length was 41bases).

Source: (Watzinger et al., 2004; WHO, 2011).

Appendix XI: A sample gel image



Legend: L = 100bp Ladder

5T1 – 5T8 = Primer 5, Sample/Template 1 – Primer pair 5, Sample/Template 8

5T9 = Primer pair 5, Negative Control

5T10 = Primer pair 5, Positive Control