USE OF SEMICONDUCTOR-BASED OLIGONUCLEOTIDE MICROARRAYS FOR IDENTIFICATION AND MONITORING OF H5N1 AVIAN INFLUENZA A VIRUS MUTATION

SUPAPORN KAEWPONGSRI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
(CLINICAL PATHOLOGY)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2008

Thesis

Entitled

USE OF SEMICONDUCTOR-BASED OLIGONUCLEOTIDE MICROARRAYS FOR IDENTIFICATION AND MONITORING OF H5N1 AVIAN INFLUENZA A VIRUS MUTATION

……..……
Miss. Supaporn Kaewpongsri
Candidate

……..……
Assoc. Prof. Wasun Chantratita,
Ph.D.
Major-advisor

……..……
Lect. Chonlaphat Sukasem,
Ph.D.
Co-advisor

…….
Assoc. Prof. Surangrat Srisurapanon,
Ph.D.
Co-advisor

……..……
Prof. Banchong Mahaisavariya,
M.D.
Dean
Faculty of Graduate Studies

……..……
Assoc. Prof. Punnee Butthep,
Ph.D.
Chair
Master of Science Programme
In Clinical Pathology
Faculty of Medicine
Ramathibodi Hospital

Thesis

Entitled

USE OF SEMICONDUCTOR-BASED OLIGONUCLEOTIDE MICROARRAYS FOR IDENTIFICATION AND MONITORING OF H5N1 AVIAN INFLUENZA A VIRUS MUTATION

was submitted to the Faculty of Graduate Studies, Mahidol University

for the degree of Master of Science (Clinical Pathology)
on

October 21, 2008

……
Miss. Supaporn Kaewpongsri
Candidate

……
Assoc. Prof. Viraphong Lulitanond,
Ph.D.
Chair

……..
Assoc. Prof. Surangrat Srisurapanon,
Ph.D.
Member

…….
Assoc. Prof Wasun Chantratita,
Ph.D.
Member

…….
Assoc.Prof. Pavapan Bhattarakosol,
Ph.D.
Member

……
Lect. Chonlaphat Sukasem,
Ph.D.
Member

…….…....
Prof. Banchong Mahaisavariya,
M.D.
Dean
Faculty of Graduate Studies
Mahidol University

..……....
Prof. Rajata Rajatanavin,
M.D., F.A.C.E.
Dean
Faculty of Medicine
Ramathibodi Hospital,
Mahidol University

ACKNOWLEDGEMENTS

There are a number of exceptional individuals who have had a hand in helping me reach this point in my academic life. The support during my time here in Ramathibodi Hospital came in either fantastically colossal proportions or unbelievably minuscule notches, but every amount is appreciated all the same. I would like to thank my supervisory committees, Associate Professor Wasun Chantratita for all of his guidance and generous advice through the last 2 years. He has continually encouraged me to think with a systematic mindset and to pursue scientific excellence no matter where the discipline leads me. I also would like to thank my-coadvisor, Associate Professor Surangrat Srisurapanond, who spend time to advise me, for her kindness, cheerfulness and providing valuable advice throughout the course of this thesis. I would like to thank Dr. Chonlaphat Sukasem for helpful suggestion and discussion. I would like to thank my committee member Associate Professor Veraphong Lulitanond and Pavapan Patarakosol for their indispensable participation and insight into this thesis.

Much gratitude goes to Mrs.Juthatip and Ms.Wantanit of Virology and Molecular Microbiology Unit, for technical assistance. I would like to special thank to Mr.Ekawat, for helping me in bioinformatics guidance. Thanks also to Faculty of veterinary, Mahidol University, Nakornprathom, Thailand, Faculty of Science, Mahidol University, Bangkok Thailand, and Department of Livestock Development, National Institute of Animal Health, Bangkok, Thailand for providing valuable H5N1 RNA sample for my thesis. I deeply thank Dr. Paul Turner from Microbiology Shoklo Malaria Research Unit for suggestions that have improved this thesis. I would like to thank Dr.Julian Zwag for helping me in statistic analysis. Thanks also to my all M.Sc. friends, best friends and family for their encouragement as well through the years. Thanks for your patience and understanding.

Supaporn Kaewpongsri

USE OF SEMICONDUCTOR-BASED OLIGONUCLEOTIDE MICROARRAYS FOR IDENTIFICATION AND MONITORING OF H5N1 AVIAN INFLUENZA A VIRUS MUTATION

SUPAPORN KAEWPONGSRI 4836324 RACP/M

M.Sc. (CLINICAL PATHOLOGY)

THESIS ADVISORS: WASUN CHANTRATITA, Ph.D.,
SURANGRAT SRISURAPANON, Ph.D., CHONLAPHAT SUKASEM, Ph.D.

ABSTRACT

Global surveillance of influenza A/H5N1 is critical for improvement in disease management, and is especially important for early detection, rapid intervention, and a possible reduction of the impact of an influenza pandemic. There is a great demand for new technological methods for viral discovery. We described a semiconductor-base oligonucleotide microarray for influenza A virus identification and monitoring. Central to this approach was microarray designed to identify influenza A virus hemagglutinin subtypes 1 through 16 and neuraminidase subtypes 1 through 9. To further characterize this influenza A virus, we proposed Bioedit program for accurately identifying viral sequences by discovering sequence signature, which was enough distinctive information for the sequence identification.

The 8 H5N1-RNA samples were collected and extracted from birds and mammal samples between 2003 to 2006, which were used to analyze H5N1 genotypic testing with semiconductor-base oligonucleotide microarray. One hundred percent (8/8) of the samples tested were subtyped correctly as H5N1. After absolutely correct subtyping from microarray signal intensity result, the method used microarray output combined with Bioinformatic tool for identification and monitoring of genetic variations of H5N1. The outstanding feature of this assay is its capability to distinguish different strains of H5N1. Ninety percent HA (4/5) and ninety percent NA (4/5) genes were sequenced correctly in accordance with previous examinations performed by classical diagnostic methods.

This assay was a convenient and practical tool for the study of avian influenza viruses, providing important HA and NA data more rapidly than conventional methods. However, the protocol monitoring of the new strain H5N1 is recommended for further study.

KEY WORDS: H5N1 VIRUS / SEMICONDUCTOR-BASED OLIGONUCLEOTIDE / MICROARRAY / MUTATION

101 pp.

การตรวจติดตาม และบ่งชี้การกลายพันธุ์ของเชื้อไวรัสไข้หวัดนก สายพันธุ์ เอช5เอ็น1 โดยเทคนิคของเซมิคอนดักเตอร์ เบสโอลิโกนิวคลีโอไทด์ ไมโครอะเรย์

(USE OF SEMICONDUCTOR-BASED OLIGONUCLEOTIDE MICROARRAYS FOR IDENTIFICATION AND MONITORING OF H5N1 AVIAN INFLUENZA A VIRUS MUTATION)

สุภาพร แก้วผ่องศรี 4836324 RACP/M

วท.ม. (พยาธิวิทยาคลินิก)

คณะกรรมการควบคุมวิทยานิพนธ์ : วสันต์ จันทราทิตย์, Ph.D., สุรางค์รัตน์ ศรีสุรภานนท์, Ph.D ชลภัทร สุขเกษม, Ph.D.

บทคัดย่อ

โรคไข้หวัดนกสายพันธุ์ เอช 5 เอ็น 1 เป็นโรคติดต่อที่ พบการระบาดทุกภูมิภาคทั่วโลก การเฝ้าระวังและการจัดการควบคุมการแพร่ระบาดที่ได้ผล ต้องอาศัยการตรวจวัดที่รวดเร็วและทันการ จึงจะสามารถควบคุมการระบาดที่รุนแรงของไวรัสไข้หวัดนก ดังนั้นเพื่อจัดการและควบคุมการแพร่ระบาดที่รุนแรงของไวรัสไข้หวัดนกอย่างได้ผล จะต้องอาศัยเทคโนโลยีในการตรวจจับที่รวดเร็ว การนำเทคโนโลยีเซมิคอนดักเตอร์ เบสโอลิโกนิวคลีโอไทด์ ไมโครอะเรย์ เป็นเทคโนโลยีร่วมของการผลิตชิปวงจรรวมและสารกึ่งตัวนำ มาใช้ระบุสายพันธุ์และครอบคลุมการติดตามการกลายพันธุ์ของไวรัสไข้หวัดนก โดยอาศัยเทคโนโลยีนี้สามารถระบุชนิดของไวรัสอินฟลูเอนซาตามคุณสมบัติแอนติเจนฮีมากลูตินินได้ 16 ชนิด และ นิวรามินิเดสได้ 9 ชนิด

โดยการศึกษานี้ได้นำตัวอย่าง อาร์เอ็นเอเชื้อไวรัสไข้หวัดนกสายพันธุ์ เอช 5 เอ็น 1 ที่สกัดจากสิ่งส่งตรวจจากนกและสัตว์เลี้ยงลูกด้วยนมที่ระบาดในปี 2003-2006 มาทำการตรวจวิเคราะห์ด้วยเทคโนโลยีดังกล่าว ซึ่งสามารถระบุชับไทป์ของตัวอย่างทั้ง 8 ได้อย่างถูกต้อง แต่เพื่อให้ครอบคลุมการตรวจติดตามการเกิดการกลายพันธุ์ของเชื้อไวรัสไข้หวัดนกระดับโมเลกุล จึงมีการใช้เทคโนโลยีชีวสารสนเทศ ร่วมกับผลการตรวจจับสัญญาณของเซมิคอนดักเตอร์ เบสโอลิโกนิวคลีโอไทด์ เพื่อที่จะหาลำดับเบสที่จะบ่งถึงความจำเพาะต่อสายพันธุ์เชื้อไวรัสไข้หวัดนกสายพันธุ์ เอช 5 เอ็น 1 โดยวิธีดังกล่าวสามารถถอดรหัสพันธุกรรมของฮีมากลูตินินและนิวรามินิเดสได้ถูกต้องและสอดคล้องกับผล จากวิธีการหาลำดับเบสด้วยวิธีดั้งเดิม จากตัวอย่าง 4 ใน 5 ตัวอย่าง

ถึงแม้การหาลำดับเบสด้วยวิธีการทางเทคโนโลยีชีวสารสนเทศ จะสะดวกกว่าวิธีการถอดรหัสพันธุกรรมด้วยวิธีดั้งเดิม แต่ข้อจำกัดของของวิธีนี้ คือ จะต้องปรับข้อมูลลำดับเบสในฐานข้อมูลให้ทันสมัยอย่างสม่ำเสมอ เพื่อรองรับการตรวจติดตามการกลายพันธุ์ไข้หวัดนกสายพันธุ์

101 หน้า

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

AIV	=	Avian influenza virus
BLAST	=	Basic Local Alignment Search Tool
°C	=	Degree of celcius
DNA	=	Deoxyribonucleic acid
dNTP	=	Deoxynucleotide triphosphate
EDTA	=	Ethylene diamine tetracetate
FLAN	=	Flu ANnotation
g	=	Gram
HA	=	Hemagglutinin
HPAI	=	High pathogenic avian influenza
hr	=	Hour (s)
L	=	Litres
LPAI	=	Low pathogenic avian influenza
m	=	Milli (10^{-3})
M	=	Matrix
min	=	Minutes
mg	=	milligram
$MgCl_2$	=	Magnesium chloride
NA	=	Neuraminidase
NaCl	=	Sodium chloride
NP	=	Neucleoprotein

LIST OF ABBREVIATIONS (Continued)

RNP	=	Ribonucleoprotein
p	=	pico (10^{-12})
PCR	=	Polymerase chain reaction
PHYLIP	=	Phylogenetic inference package
sec	=	Second (s)
TBE	=	Tris-borate-ethylene diamine tetraacetic acid
Tm	=	Melting temperature
TMB	=	Tetramethylbenzidine
U	=	Unit (s)
UV	=	Ultraviolet
μ	=	Micro (10^{-6})

CHAPTER I
INTRODUCTION

Bird flu is a complex and dangerous disease that recognizes no borders. It is being watched anxiously by the entire world. The most important fact about avian influenza is that it is used by the same virus that causes influenza among humans. Influenza A viruses infecting poultry can be divided into two distinct groups on the basis of their ability to cause disease highly pathogenic avian influenza (HPAI) and low pathogenic avian influenza (LPAI). HPAI are highly virulent and may result in mortality as high as 100%. These viruses have been restricted to subtypes H5 and H7, although not all viruses of these subtypes cause HPAI. LPAI viruses are generally lower virulence but these virus can serve as progenitors to HPAI viruses (1, 2).

Influenza viruses are segmented, negative-sense, RNA viruses that exhibit considerable antigenic diversity. They are divided into types, either A, B or C by variations in nucleoprotein (NP) and matrix protein (M) antigens. Type A viruses are further subtyped by the antigenic differences in two proteins, hemagglutinin (HA) and neuraminidase (NA), which are present on the virus particle surface. Sixteen HA subtypes (designated H1 to H16) and nine NA subtypes (designated N1 to N9) have been identified so this virus has many strain (3). Influenza B and C viruses are not divided into subtype. Although all HA and NA subtypes are found in aquatic birds, the number of subtypes that have crossed the species barrier and established stable lineages in mammals is limited. Only three HA and two NA subtypes (i.e. H1–3 and N1–2) have circulated in human since 1918. In horses, only two influenza A subtypes (H7N7 and H3N8) are found, while, despite susceptibility to all avian subtypes in experimental settings, the only subtypes recovered from pigs in nature are H1, H3, N1, and N2. The molecular, biological or ecological factors determining the apparent subtype-specific ability of viruses to cross species barriers and spread among a range of hosts remain largely unresolved (2). Different strains of virus are composed of different combinations of Hemagglutinin and Neuraminidase. Some strains lead to

severe symptoms. The strain that has spread through Thailand and other countries in the past two years causes severe symptoms. Its scientific name is H5N1 (4).

From a diagnostic standpoint, influenza is often described as a “moving target” due to the high mutation rate of HA and NA (5). But it is known that antigenic analysis of influenza A/H5N1 viruses from 2002-2003 have also shown significant antigenic drift compared to viruses from 1997 to 2001. In addition, the evolution of the HA gene of A/H5N1 since 2004 has resulted in several different lineages, and according to phylogenetic analyses (6). As this particular virus continues to mutate, it is important to have a diagnostic tool capable of detecting as many of the A/H5N1 variants as possible for reliable global strain surveillance efforts. Preparedness in diagnostic laboratories means recognition of the particular agent in good time, but obtaining information on the characteristics of the detected pathogen is also of increasing importance. The latter might help to trace the source of the infection, predict its spread and the potential epidemiological consequences by identifying virulence markers, receptor binding motifs, drug resistance traits, reassorting variants, etc., and elaborate the strategies to control it. The data regarding the above considerations is most accurately generated by in-depth nucleotide sequence analysis of the pathogen of concern (2, 7-9).

A number of diagnostic methods are available for the detection and subtype of influenza viruses. The current “gold standard” for complete influenza A subtyping (determination of both HA and NA) involves virus replication in egg or tissue culture followed by a hemagglutination inhibition test. This method is tedious and requires several days, with the analysis time often extended to several weeks for antigenically novel viruses (3, 10-12). Although these methods are excellent for determining subtypes, they do not give detailed genetic information when antigenic shifts occur. RT-PCR techniques depend on specific primers, which may fail when corresponding viral sequences mutate. By overcoming this limitation**,** microarrays have become valuable tools for viral discovery, detection, and genotyping (5, 12-17).

Recently, the molecular diagnostic methods have also gained ground and traditional as well as real-time PCR methods targeting the conserved region of the matrix protein gene are widely used for the detection of influenza A viruses. In positive cases the subtype of the detected virus is determined with other PCRs

targeting the haemagglutinin and neuraminidase genes. In case of H5 or H7 subtypes it is followed by the pathotyping, i.e. the characterization of the nucleotide composition of the region flanking the HA cleavage site. If the detected virus turns out to be HPAI (that is subsequently confirmed by biological assays as well), strictly regulated measures must be applied as detailed in the respective National Contingency Plans. Although the above steps carried out in National Reference laboratories accomplish the compulsory laboratory investigations, thorough analysis of every isolate has become a routine in the more sophisticated laboratories in order to be better prepared for a possible outbreak. In the case of the highly pathogenic H5N1 avian influenza, full-length sequencing protocols are being used in the diagnostic laboratories. These protocols can quickly provide comprehensive data on a detected H5N1 avian influenza virus to animal and public health authorities facilitating their actions (18, 19).

A Semiconductor-based oligonucleotide microarray platform can simultaneously detect and characterize many H1-H16 and N1-N9 subtypes of influenza A viruses. Oligonucleotide microarrays do this by simultaneously interrogating hundreds to thousands of immobilized oligonucleotide probes, where each probe provides a single query for a known sequence that is unique for each H gene and N gene. The micro-array chip can be read fluorescently or using a simple electrochemical detection (direct reading of the chip through an electrochemical sensing process). RNA samples are prepared, labeled and hybridized with arrays. Arrays are scanned and images are produced and analysed to obtain an intensity value for each probe. These intensities represent how much hybridization occurred for each oligonucleotide probe (12, 20).

This semiconductor-based microarray chip detected and accurately typed flu strains with hemagglutinin subtypes 1-16 and neuraminidase subtypes 1-9 but do not provide strain sequence information. It’s very difficult or ambiguous to tested genetic drift or shift in a flu virus within a chip. To simplifly the assay procedure for running the influenza assays, we demonstrated the bioinformatic’s ability to distinguish different strain of each isolates.

CHAPTER II
OBJECTIVES

1. To determine the performance of Semiconductor-Based Oligonucleotide Microarrays for detection and subtyping the Avian Influenza virus (H5N1) circulating in Thailand.

2. To describe the development of an algorithm to find potential oligonucleotide probe that would enable the nucleotide sequencing of different influenza A/H5N1 viruses on a microarray.

3. To monitor the antigenic diffrences of H5N1 during 2003-2006.

CHAPTER III
LITERATURE REVIEW

3.1 General Description of Avian Influenza

Influenza viruses hold generic status in the *Orthomyxoviridae* family and are classified into types A, B or C based on antigenic differences of their nucleo- and matrix proteins. Avian influenza viruses (AIV) belong to type A. Influenza A viruses are lipid-enveloped viruses containing a genome composed of eight strands of negative-sense RNA that encode ten viral proteins. These gene segments are encapsidated in a virally encoded nucleoprotein (NP), and the ribonucleoprotein (21) structures are associated with the three subunits of the viral polymerase (PB1, PB2, and PA). Virus particle formation occurs at the surface membrane of infected cells, where budding occurs from regions of the membrane at which the viral glycoproteins, hemagglutinin (HA) and neuraminidase (NA) have accumulated. The viral matrix protein (M1) is the most abundant component of the virion and is thought to play a pivotal role in the process of assembly and budding. Details on the mechanism of assembly are still forthcoming, but it is often assumed that M1 interacts with the RNPs and the cytoplasmic domains of HA, NA, and possibly the third integral membrane protein M2. The two other viral proteins, NS1 and NS2, were initially designated as nonstructural proteins, but there is now evidence for the presence of NS2 in virions (22, 23). Figure 1 is a schematic representation of Influenza A virion.

The main antigenic determinants of influenza A and B viruses are the haemagglutinin (H or HA) and the neuraminidase (N or NA) transmembrane glycoproteins, capable of eliciting subtype-specified and immune responses which are fully protective within, but only partially protective across, different subtypes. On the basis of the antigenicity of these glycoproteins, influenza A viruses currently cluster into sixteen H (H1 - H16) and nine N (N1 - N9) subtypes. These clusters are

substantiated when phylogenetically analysing the nucleotide and deduced amino acid sequences of the HA and NA genes, respectively

3.1.1 Description of the Gene Segments

Using the A/Puerto Rico/8/34 virus (PR8) as an example, the influenza A viral gene segments range in size from 890 to 2341 nucleotides and contain from 20 to 45 noncoding nucleotides at the 30 end and 23 to 61 nucleotides at the 50 end, depending on the segment.With the exception of position 4 at the 30 end of viral gene segments, which displays U/C heterogeneity, the 12 nucleotides at the 30 end and the 13 nucleotides at the 50 end are completely conserved for all segments in all strains of influenza A virus. These terminal RNA regions are partially complementary and viral promoter activity has been mapped to these domains, but with the exception of the polyadenylation signal, the functional significance of most of the noncoding sequences beyond the conserved domain remains unresolved. There is common speculation that these sequences are involved in binding of NP and/or polymerase complexes, or that they may contribute to packaging signals.

The three largest gene segments encode the subunits of the viral polymerase, PB2, PB1, and PA, which are so named because of their basic (PB2, PB1) or acidic (PA) properties on isoelectric focusing gels. These are responsible for transcribing messenger RNAs (mRNAs), for synthesizing positive-sense antigenomic template RNAs (cRNAs), and for transcribing the cRNAs into the gene segments (vRNAs) that are incorporated into progeny viruses. Segment 4 encodes the hemagglutinin glycoprotein , which is responsible for binding virus to sialic acid–containing cell-surface receptors and for membrane fusion during virus entry into host cells. It is also the principal target for neutralizing antibodies. The nucleoprotein is the product of the fifth gene segment. This is the protein that encapsidates cRNAs and vRNAs, which is necessary for them to be recognized as templates for the viral polymerase. Segment 6 encodes the neuraminidase , which cleaves sialic acid from virus and host cell glycoconjugates at the end of the virus life cycle to allow mature virions to be released. Segment 7 generates two gene products, the matrix protein, M1, and the M2

protein. M1 mRNA is a collinear transcript, and its product has a structural role in the virion and it is thought to play a fundamental role in virus assembly. The M2 is a small transmembrane protein derived from spliced mRNA. It has proton channel activity that aids in virus disassembly during the initial stages of infection. The eighth gene segment also encodes two proteins due to alternative splicing. These proteins were originally referred to as NS1 and NS2 because they were thought to be nonstructural, but NS2 has since been shown to be a component of virions. The NS1 has numerous functions. It is a regulator of both mRNA splicing and translation, and it also plays a critical role in the modulation of interferon responses to viral infection. The NS2 functions to mediate the export of newly synthesized RNPs from the nucleus and as such, it is also referred to as the nuclear export protein (NEP) (23).

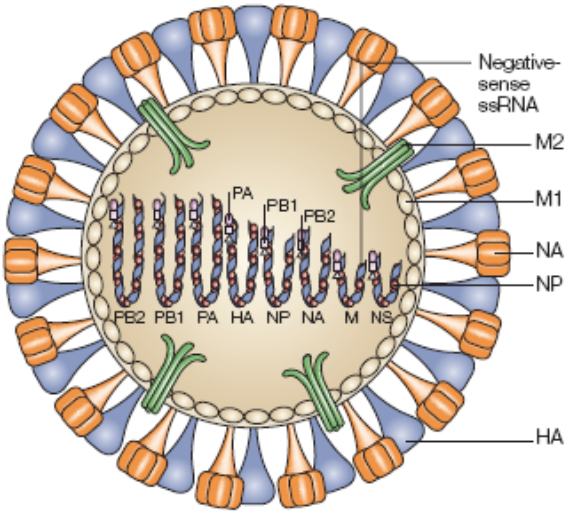


Figure 1 Schematic diagram of an influenza A virus virion. Two surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA), and the M2 ion-channel protein are embedded in the viral envelope, which is derived from the host plasma membrane. The ribonucleoprotein complex comprises a viral RNA segment associated with the nucleoprotein (NP) and three polymerase proteins (PA, PB1 and PB2). The matrix (M1) protein is associated with both ribonucleoprotein and the viral envelope. A small amount of non-structural protein 2 is also present, but its location within the virion is unknown (24).

3.1.2 Antigenic Influenza A virus

The main antigenic determinants of influenza A and B viruses are the haemagglutinin and the neuraminidase transmembrane glycoproteins. Hemagglutinin (H) is a glycoprotein containing either 2 of 3 glycosylation sites, with a molecular weight of approximately 76,000. It spans the lipid membrane so that the major part, which contains at least 5 antigenic domains, is presented at the outer surface. HA serves as a receptor by binding to sialic acid (N-acetylneuraminic acid) and induces penetration of the interior of the virus particle by membrane fusion. Haemagglutinin is the main influenza virus antigen; the antigenic sites being A, B (carrying the receptor binding site), C, D, and E. The antigenic sites are presented at the head of the molecule, while the feet are embedded in the lipid layer. The body of the HA molecule contains the stalk region and the fusiogenic domain which is needed for membrane fusion when the virus infects a new cell. At low pH, the fusion peptide is turned to an interior position. The HA forms trimers and several trimers form a fusion pore.

Neuraminidase is a glycoprotein, which is also found as projections on the surface of the virus. It forms a tetrameric structure with an average molecular weight of 220,000. The NA molecule presents its main part at the outer surface of the cell, spans the lipid layer, and has a small cytoplasmic tail. NA acts as an enzyme, cleaving sialic acid from the HA molecule, from other NA molecules and from glycoproteins and glycolipids at the cell surface. It also serves as an important antigenic site, and in addition, seems to be necessary for the penetration of the virus through the mucin layer of the respiratory epithelium (23, 25, 26).

3.1.3 Influenza virus replication cycle

The influenza virus binds to the cell surface by fixing the outer top of the HA to the sialic acid of a cell.s glycoproteins and glycolipids. The sialic acid linkage to the penultimate galactose, either alpha 2,3 (in birds) or alpha 2,6 (in humans), determines host specificity (27, 28). Since sialic acid-presenting carbohydrates are present on several cells of the organism, the binding capacity of the

HA explains why multiple cell types in an organism may be infected. After attachment, the virus is taken up by the cell via a clathrin-coated receptor mediated endocytosis process. When internalised, the clathrin molecules are liberated and the vesicle harbouring the whole virus fuses with endosomes. The contents of the vesicle are usually digested through a stepwise lowering of the pH within the phagosome. When a certain level is reached, the lowering of the pH is stopped by the action of the M2 protein which induces the partial liberation of the fusion peptide of the HA (23). This allows the fusion of the HA with the membrane of the vesicle and liberation of the ribonucleoproteins (RNPs) into the cytoplasm, as described above. The ion influx from the endosome to the virus particle leads to disconnection of the different viral proteins; M1-protein aggregation is disrupted and RNPs no longer adhere to the M1-protein complex. Uncoating is completed within 20-30 min of virus attachment. The RNPs are transported to the nucleus, where the polymerase complex binds to viral RNA, cleaves viral RNA by its endonuclease activity, and simultaneously leads to elongation. The production of viral RNA is limited by the NP in favour of mRNA. Both are transported to the cytoplasm, where viral proteins are generated at the ribosome. Part of the viral mRNA is spliced by cellular enzymes so that finally viral proteins, such as M1 and NS2, can be synthesized without any further cleavage. Some of the newly synthesized viral proteins are transported to the nucleus where they bind to viral RNA to form RNPs. Other newly synthesized viral proteins are processed in the endoplasmic reticulum and the Golgi apparatus where glycosylation occurs. These modified proteins are transported to the cell membrane where they stick in the lipid bilayer. When they reach a high enough concentration at the plasma membrane, RNPs and M1 proteins aggregate and condense to produce the viral particle. Finally, the particle is extruded from the membrane and will be liberated by the neuraminidase activity. The time from entry to production of new virus is on average 6 h (26, 29, 30).

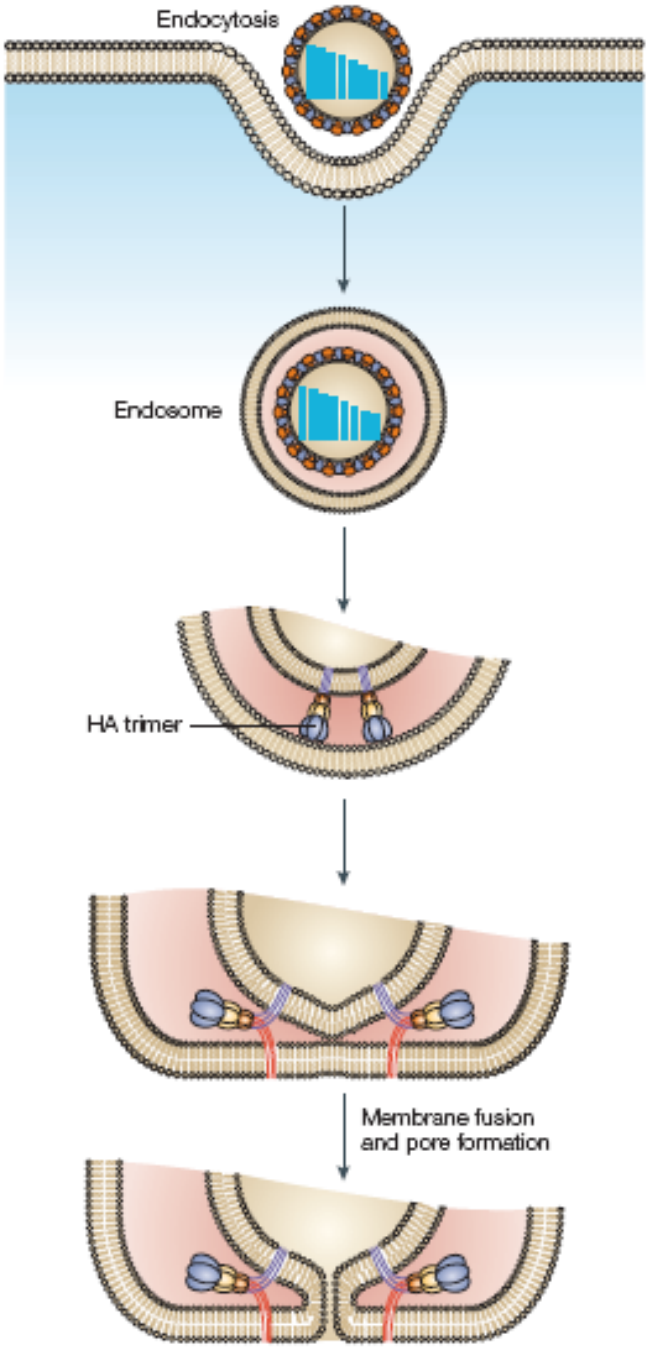


Figure 2 Schematic representation of mechanism of action of influenza A haemagglutinin (HA). HA mediates virus binding to sialic-acid-containing host-cell receptors. Following binding, the virus is internalized by endocytosis. Acidification of the endosome environment induces conformational changes in the HA trimer, mediating fusion between the viral envelope and the endosomal membrane, which allows delivery of the viral ribonucleoprotein into host cells (3).

3.1.4 Genetic variation

RNA viruses are extremely mutable and use very efficient strategies for generating viral diversity during evolution (Figure 3). RNA viruses have few or no proofreading mechanisms and many mutations are introduced during replication. As such, RNA viruses exist as a quasispecies comprising viruses of slightly different genetic composition. Thus, when selective conditions arise, such as the conditions present after jumping to a new host, variants with an advantage are quickly amplified. It can therefore take only limited replication under suboptimal conditions before an adapted virus emerges. Unfortunately, the evolution of viruses is not only confined to point mutations inserted during replication. Viruses can also undergo leaps in evolution through the processes of recombination and reassortment. These processes achieve population heterogeneity in viruses through the acquisition of large sections of genomic material from other viruses. Reassortment occurs in viruses with segmented genomes when sections of different viruses are exchanged on dual infection. Both the Asian (1957) and Hong Kong (1968) human influenza pandemics arose through reassortment of human and avian strains (31).

Global health authorities are concerned that the same may happen with the H5N1 viruses that are currently circulating in birds in southeastern Asia, with the result being a virus that is transmissible in humans. Unfortunately, we know little about the conditions for successful reassortment and it is difficult to assess the possibility of this happening. The available evidence from influenza A viruses is that there is a selective rather than random mechanism for packaging RNA segments into infectious virions36 and not all combinations of viral genes are compatible. Recent findings from vaccine development have shown that the surface proteins of the H5N1 influenza viruses are at least compatible with older human strains, suggesting that reassortants of H5N1 and human influenza virus may well be viable. Arenaviruses are also capable of reassortment, although reassortment again does not seem to be completely random and some combinations of genes do not seem to be compatible (38). Nevertheless, there has been some speculation that reassortment may have contributed to the emergence of new arenaviruses with disease associations (32). Recombination leads to the same end as reassortment—that is, the acquisition of

foreign genetic material—but through a different route. Rather than through the acquisition of gene segments, recombination occurs through 'splicing' of the foreign component into the virus genome. Recombination is often observed in positive-sense RNA viruses such as picornaviruses and coronaviruses (33, 34) although it also occurs to a limited extent in negative-sense viruses such as influenza A (35). As with reassortment, recombination has the capacity to change the phenotype of a virus markedly and has the potential to aid in the transmission of viruses between species (36).

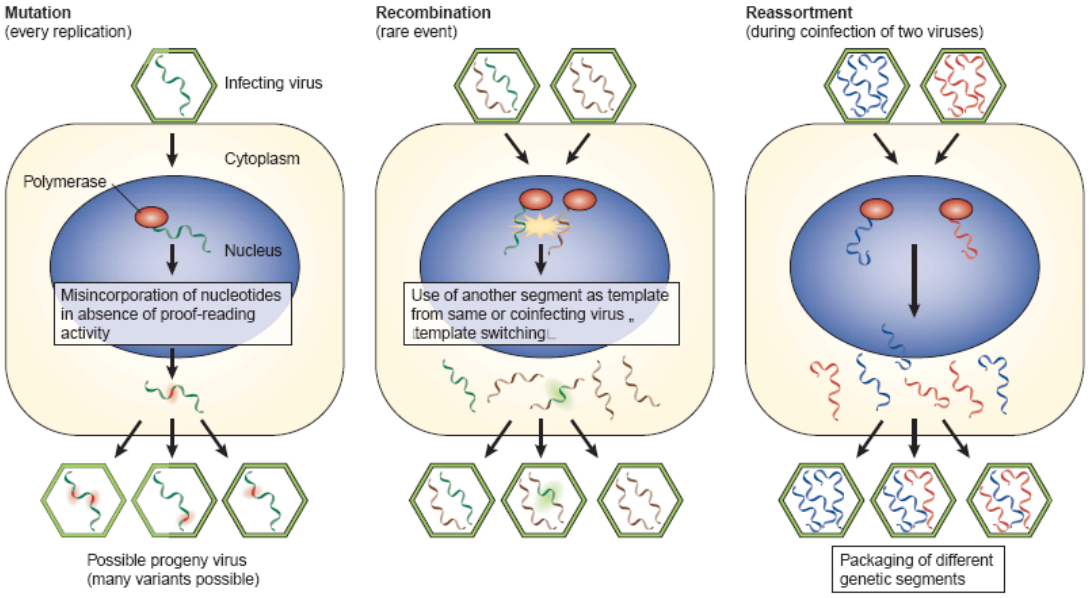


Figure 3 Molecular mechanisms for generating viral diversity (37)

3.2 History and origin of H5N1 viruses

Avian influenza viruses have now been identified as the source of novel hemagglutinin and neuraminidases (38) associated with pandemics. Two mechanisms can be responsible for the adaptation of avian viruses with novel HA and NA into viruses with the capacity to spread in human populations, genetic reassortment, or direct transmission. The segmented nature of the influenza virus genome allows for the exchange of genetic material between two subtypes of the virus coinfecting a host.

Such genetic reassortment can result in a virus with a genetic makeup different from either of the coinfecting subtypes. Sequence analysis has identified the viruses of both 1957 and 1968 pandemics as products of reassortment (39).

Because human influenza virus was first identified in 1933, it is harder to pinpoint the origin of the 1918 pandemic. However, the genetic material of the 1918 virus has been amplified and sequenced from formalin-fixed pathological specimens of the lung tissue (40). Although phylogenetic analysis of the HA from the 1918 virus places the virus firmly in the mammalian clade of the influenza virused, the sequences of the NA, NP (nucleo protein), and polymerase genes are close to the avian consensus sequences and strongly suggest that the 1918 virus was a direct introduction into human populations of an avian virus, possibly with a period of adaptation before it emerged in pandemic form (41).

Sequence analysis of H5N1 and other avian viruses has suggested that the current H5N1 virus originated by reassortment between multiple cocirculating avian influenza strains prevalent in Hong Kong in 1997. In addition, certain unique structural features like a 19 amino acid sequence deletion in the stalk of the NA and and additional glycosylation site on the head of the HA were found in the 1997 isolate of the H%N1 virus. Although these features are characteristic of highly pathogenic viruses infecting chicken, they are not common to the avain influenza viruses circulating among wild birds (42).

3.3 Avian Influenza pathogenicity

Avian influenza viruses can be sorted on the basis of virulence: highly pathogenic avian influenza viruses cause systemic lethal infection, killing birds as soon as 24 hours post-infection, and usually within one week, these viruses have been restricted to subtypes H5 and H7, although not all viruses of these subtype cause HPAI, (43) whereas low-pathogenic avian influenza viruses rarely generate outbreaks of severe disease in the field, and their associated morbidity and mortality rates are lower than those of HPAI viruses (44). Phylogenetic studies of avian influenza viruses have revealed two geographically separate sublineages (Eurasian and American),

which probably reflects the separation of viruses owing to the distinct migratory patterns of the host birds in these two regions. Both HPAI and LPAI viruses are found within these two sublineages, indicating that viral pathogenicity is not determined by geographical distribution. Without exception, all of the HPAI viruses belong to the H5 or H7 subtype, for reasons that are still unclear. There do not seem to be any associations of specific NA subtypes with HPAI viruses (26). Sometimes other infections or environment conditions may cause exacerbation of influenza infections leading to much more serious disease.

Virulence of influenza viruses is a polygenic trait and has been attributed to several viral genes; however, the HA protein is a major virulence factor in poultry (45). It has been demonstrated that the HA0 precursor of the main functional HA glycoprotein requires cleavage by host proteases before virus particles are infectious. HA0 proteins of avian influenza viruses of low virulence for poultry are limited to cleavage by host proteases such as trypsin and trypsin-like enzymes and thus restricted to replication at sites in the host where such enzymes are found, i.e. the respiratory and intestinal tracts. In contrast virulent viruses appear to be cleavable by a ubiquitous protease(s), which remains to be fully identified but appears to be one or more proprotein-processing subtilisinrelated endoproteases of which furin is the leading candidate and this enables these viruses to replicate throughout the animal, damaging vital organs and tissues which brings about disease and death in the infected bird (43, 46, 47).

Comparisons of the amino acid sequences at the HA0 cleavage site of avian influenza viruses of high and low pathogenicity revealed that while viruses of low virulence have only two basic amino acids, at positions α1 and α 4 from the cleavage site for H5 and at positions α 1 and α 3 for H7 subtype, all HPAI viruses possessed multiple basic amino acids (arginine and lysine) adjacent to the cleavage site either as a result of apparent insertion or apparent substitution. The additional basic amino acids resulting in a motif recognised and cleavable by the putative ubiquitous protease(s) (48-50). Mammals, including humans, also have furinlike proteases capable of cleaving at multiple basic amino acid motifs. However, the question as to whether or not viruses with multiple basic amino acids at the H0 cleavage site could cause systemic infections and highly pathogenic disease in humans remained unanswered up to 1997,

because the viruses known to have infected humans (H1, including the 1918 pandemic virus, H2 and H3 subtypes) all have motifs at the cleavage site indicating they would be cleaved only by trypsin like enzymes (43).

3.4 Human infection with influenza A (H5N1) viruses and Clinical Feature

In aquatic birds, the natural hosts of influenza viruses, infection is usually asymptomatic and localized to the intestinal tract. H5N1 viruses have been actively reassorting and crossing interspecies-host barriers, moving from aquatic poultry to land-based poultry and, more recently, to wild terrestrial birds and humans (43, 51, 52) Avian and human viruses preferentially bind sialic acid molecules with specific oligosaccharide side chains with α-2,3 and α-2,6 linkages, respectively. Receptor specificity was thought to be an important determinant of viral host range, as demonstrated by the fact that reports of experimental human infections with avian influenza viruses were rare (53).

Also, recent studies with differentiated human airway epithelial cells have demonstrated that cells bearing α-2,3 and α-2,6 linked sialic acids are present in human respiratory epithelium (54). Pigs are susceptible to both avian and human influenza viruses, due to the presence of sialic acid molecules with both α 2,3 and α 2,6 linkages to galactose on their tracheal epithelium, and could potentially serve as mixing vessels for the reassortment of avian and human viruses (55).

H5 viruses can also be associated with high cleavability and lethality in poultry and have caused many severe outbreaks of disease in poultry. The first known transmissions of H5N1 viruses from birds to humans took place in April–May 1997,(56) shortly after the recognition of an outbreak in sick geese on a farm in Guangdong Province of China (57). Human cases have typically followed outbreaks in poultry. Concerns about H5N1 infection becoming a global threat were raised when H5N1 infection was detected in migratory waterfowl in Qinghai Lake in western China in April 2005 (58). The virus has since spread from its original geographic focus in Southeast Asia along the flyways of migratory birds to more than 30 countries

worldwide, including Kazakhstan, Russia, parts of Europe, the Middle East, and Africa. As of August 2006, there have been a total of 238 human cases with 139 deaths (58%) with additions to the list of countries reporting confirmed human cases like Turkey, Iraq, Azerbaijan, Egypt, and Djibouti (59). In many of these countries, close human contact with poultry is common due to practices such as maintenance of backyard flocks of birds, live animal markets, and a thriving poultry industry. This combination of these factors has resulted in continued opportunities for human transmission, increasing the likelihood of further human adaptation and transmission between humans Figure 4.

To date, person-to-person transmission has been rare. Fifteen family clusters of infection involving 2 family members have been documented between January 2004 and July 2005. The largest cluster identified thus far has been in the village of Kubu Sembelang in northern Sumatra in Indonesia in May 2006 where seven confirmed cases occurred in extended family members of a 37-year-old woman who died of an acute respiratory illness and was buried prior to the establishment of a diagnosis. Sequence comparisons of viruses from human cases and those of concurrent avian cases have established that at least some of the cases in the cluster represented person-to-person transmission. In addition, there is one well-documented transmission of virus from an ill child in Thailand to her mother and a suggestive report of transmission to a health care worker (39, 52).

The available data regarding the signs and symptoms of H5N1 infection are mostly from hospitalized patients (39). The frequency of subclinical and mild forms of infection is currently unknown. Most patients present with the nonspecific complaints of a fever 38°C along with cough and shortness of breath. In many of the patients, there is a progression of symptoms leading to respiratory failure requiring ventilation and other supportive measures. Atypical symptoms like nausea, vomiting, encephalopathy, and bleeding gums and nose have been reported. Watery diarrhea may be present prior to the onset of respiratory symptoms. The majority of patients have an abnormal chest x-ray with diffuse and multifocal or patchy infiltration, but pleural effusions are rare. Laboratory abnormalities include significant lymphopenia and leucopenia, mild to moderate thrombocytopenia, and elevated transaminases. These abnormalities are poor prognostic signs (60-62). Most people have negative

sputum and blood bacteriological cultures. Pathological changes show diffuse alveolar damage in the lungs, reactive hemophagocytosis in the marrow, and lymphoid depletion with atypical lymphocytosis in the spleen and lymphoid tissues. Centrilobular hepatic necrosis and acute tubular necrosis have been reported (26, 63, 64).

3.5 Evolution of H5N1 viruses

A key concept in influenza virus evolution is that there is a marked difference in evolutionary dynamics between those viruses that infect aquatic birds and those from other host species. The evolution of H5N1 viruses in recent years has been associated with increasing virulence and an expanding host range, which beside terrestrial poultry and wild birds. Also includes mammals. While all H5N1 viruses isolated from ducks in China between 1999 and 2002 were highly pathogenic in chickens, an increasing level of pathogenicity was observed in mice with the progression of time: virus isolated in 1999 and 2000 were less pathogenic than those isolated in 2001 and 2002 (65). It has been suggested that the increasing ability to replicate in mammals has resulted from transmission between ducks and pigs as Figure 4. The expanding host range is also illustrated by successful experimental infection of domestic cats, and natural infections of tigers and leopards with recent H5N1 strains (21, 66, 67).

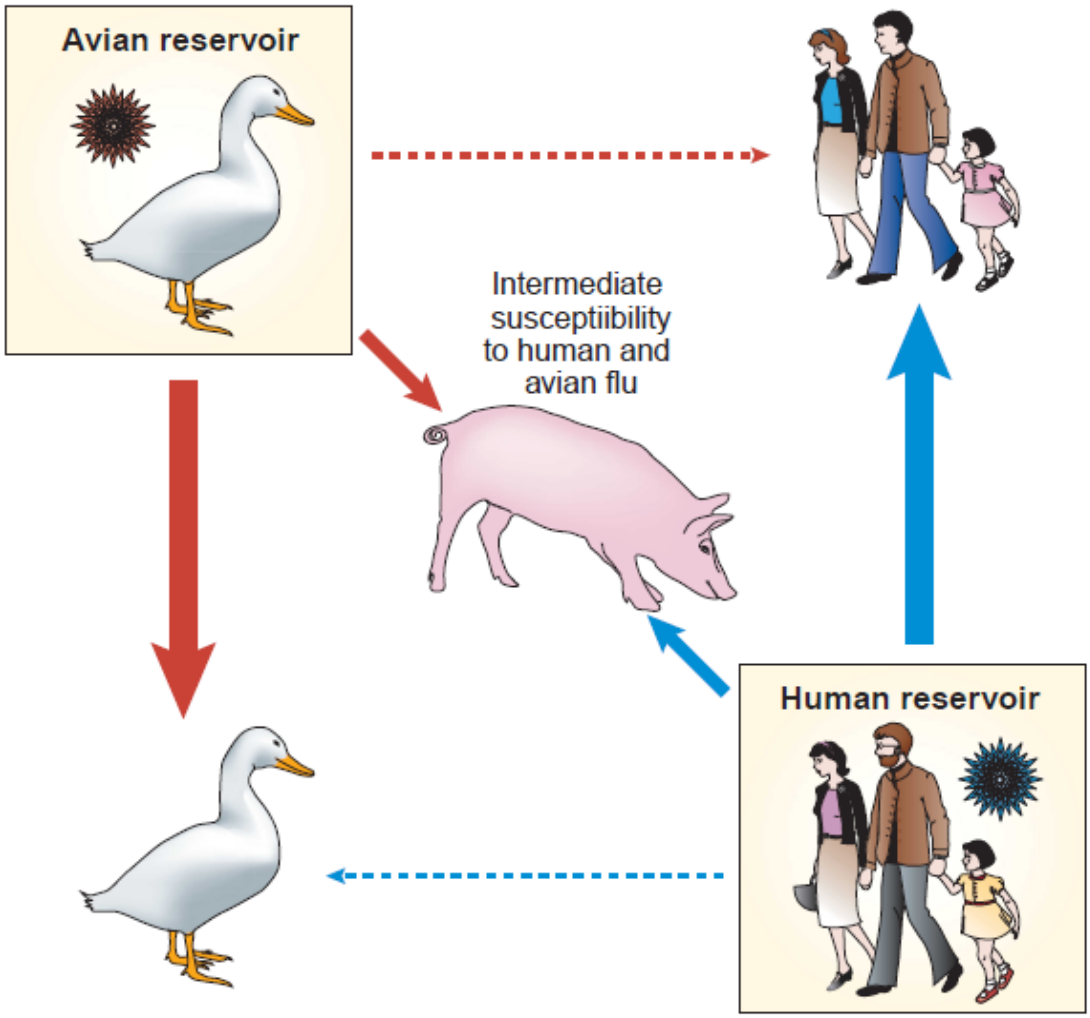


Figure 4 Possible cross-species transmission of H5N1 virus (37)

3.6 Avian influenza viruses and pandemic influenza

Evidence of bird flu spreading to humans caused enormous concern among scientists and among national and international health organizations. The reason for the concern was the fear that the virus might one day evolve into a form that spreads easily from human to human. If that were to happen, it would be difficult to avoid a pandemic causing hundreds of millions of illnesses, and millions of deaths. Pandemics of this magnitude have occurred six or seven times in the past two centuries. Avian species constitute the origin of the human H5N1 avian influenza

virus. The virus has not yet manifested effective human-to-human transmission, but the situation may change if the virus continues to mutate and assort during an epidemic (43, 68).

The viruses that case bird flu and human flu store their genetic code in RNA. These viruses frequently mutate; in other words, changes occur in their genetic code. The mutations often alter the characteristics of the viruses. The immune systems of individuals infected by earlier strains of the virus do not necessarily recognize the new strains. Thus the same individuals can become infected for a second time. Mutations in the influenza virus, including avian and human influenza, can occur either gradually or quickly. Either type of mutation can give rise to new strains of the virus. These new strains may increase or decrease the severity of the symptoms. The effects are impossible to predict (22, 69, 70).

3.7 Characteristic of H5N1 isolated in Thailand review prevalence and incidence in Thailand

The 2005 H5N1 viruses circulating in Thailand were genetically comparable with 2004 H5N1 isolates in Thailand and Vietnam, but distinct from the H5N1 virus strains circulating in Asia such as Hong Kong, China, Indonesia (1996–2005) and Europe (2005) as well as in Africa (2006). The 2005 H5N1 viruses harbor multiple insertions of basic amino acids at the HA cleavage site, which by definition is characteristic for highly pathogenic avian influenza (HPAI) (71-73). Interestingly, the HA cleavage site of 2005 H5N1 viruses in this study contained one basic amino acid “SPQREKRRKKR” differing from that of 2004 H5N1 viruses in Thailand, Indonesia, Vietnam and Eastern China “SPQRERRRKKR” (arginine (R) to lysine (K)). It is noted that HA cleavage site of the 2005 H5N1 viruses from Vietnam (mallard/347/05) and Indonesia (duck/Parepare/BBVM/05) remain unchanged “SPQRERRRKK” (57, 72). On the other hand, the HA cleavage site variability have been previously observed in some wild bird species during earlier outbreaks in Thailand in 2004 (open-bill/CU the significance of genetic characterization presenting as examples of 4 H5N1 viruses isolated from human, chickens and quail which isolated from Thailand in 2005. Sequence analysis of eight gene segments revealed that the 2005 H5N1 viruses

isolated in October 2005 were related to those recovered from chicken, tiger(s) and human(s) in January and July 2004 but not related to the viruses from the recent outbreaks in Western China, Eurasia and Africa. Interestingly, amino acid substitution especially at the HA cleavage site has been observed. However, the amino acid substitution still basic protein which can cause highly pathogenic to host. It shows that H5N1 viruses have continued to evolve since early 2004–2005 of Thailand with minor change, alternatively unchanged in pathogenicity (68). The Thailand viruses contained more avian-specific residues than the 1997 Hong Kong H5N1 viruses, suggesting that the virus may have adapted to allow a more efficient spread in avian species (72). It is still not certain how bird flu first came to Thailand. However, current evidence points towards migratory water birds such as ducks and geese. Experts believe that the virus H5N1 must have been evolving among bird populations for many years, to the point where it can infect birds without casing disease. Migratory birds are therefore a natural reservoir of disease, which they pass to other species along their migration routes (22, 42).

3.8 Laboratory diagnosis of influenza A/H5N1

Influenza A/H5N1 is HPAIV and cause disastrous epidemic disease in poultry (74, 75). Identification of a virus subtype and molecular identification of the subtype of the viral hemagglutinin and neuraminidase genes can help create a global surveillance program to monitor outbreaks. The goal of surveillance is to gather information on the influenza virus subtypes that are circulating in human and animal populations so that recommendations can be made on the content of vaccines for the next season. Surveillance systems of WHO Animal Influenza Network for detection of influenza viruses procedures used include detection of virus by direct virus direct virus isolation, by detection of specific genes or gene products by amplification procedures of polymerase chain reaction (PCR), rapid diagnostic kits or by detection of serological response (76). There are two methodologies for Influenza A/H5N1 monitoring: Subtype determination (“subtyping”) and nucleotide sequencing.

Subtype determination is essential for tracking emerging viruses and for designing appropriate influenza vaccines. The current "gold standard" for complete influenza A subtyping (determination of both HA and NA) involves virus replication in egg or tissue culture followed by a hemagglutinin inhibition assay. This method is tedious and requires several days, with the analysis time often extended to several weeks for antigenically novel viruses (76).

Influenza diagnostic method based on reverse transcription polymerase chain reaction (RT-PCR) and real-rime RT-PCR (RRT-PCR) are currently available for typing or as presumptive tests for a specific subtype (77-81). RRT-PCR-based typing of influenza uses a single primer pair that targets a short, well conserved portion of the M gene segment (7, 76, 82-84). RRT-PCR-based presumptive tests for a specific subtype use a single primer pair that targets a portion of the HA gene that is well-conserved only for a particular subtype. However, a major limitation of these singleplex PCR-based presumptive tests in terms of influenza surveillance is that they do not provide complete subtype information and are limited to a specific subtype. Clearly, the speed of this method for identifying specific A/H5 viruses could be extremely useful during and influenza pandemic. In addition, the inclusion of internal control in RT-PCR assays is highly desirable to monitor for false-negative results due to inefficient nucleic acid extraction, cDNA synthesis, or amplification (2, 11).

Nucleotide sequencing assays use further analysis of PCR-positive samples by automated sequencing. The PCR products were extracted and examined by agarose gel electrophoresis and further enzymatic reactions. The DNA sequencing reaction was performed, sing a commercially available kit and automate (85, 86). The end of this assay is use specific software that facilitates the process of sequence alignment. However, a positive amplification can be verified only by subsequent assays to elaborate sequence information.

According to global surveillance, thus rapid and specific identification of influenza A/H5N1 subtype and accurate sequence information are crucial for proper treatment. But traditional method are problematic in scenarios where new strains of virus emerge of mixture of viruses exist. Because of their sensitivity, specificity, and accuracy, DNA microarrays have become an acceptable technology for screening

samples for the presence or absence of a large variely of viruses simultaneously and identifying the genotype of and unknown specimen (10, 20, 83, 87-89).

3.9 Microarray Technology

A microarray is a series of nucleic acid targets immobilized on a solid substrate, such as a silicon wafer or glass slide, on which DNA or oligonucleotides from either host or pathogen are attached. These nucleic acids are complementary to thousands of genes of both known and unknown function. Hybridisation of fluorescently labeled probes made from nucleic acids in the test sample to these targets allows analysis of the relative concentrations of mRNA or DNA in the sample. There are two main types of microarray: spotted DNA and oligonucleotide arrays (14, 20, 89-91).

3.9.1 CombiMatrix Oligonucleotide microarray

CombiMatrix microarrays are based on a specially modified semiconductor adapted for biological applications, which contains arrays of platinum microelectrodes. The semiconductor logic circuitry directs digitally controlled simultaneous synthesis of different oligonucleotides at thousands of electrodes in response to a computer software program (20).

The ElectraSense^{TM} 12 K microarray consists of 12,544 individually addressable electrodes linked by the semiconductor circuitry. The contact pogo pads on the array side provide the electrica contact with and external device to control synthesis or ElectraSense^{TM} detection. Each working electrode is made of high purity platinum with a circular geometry (44 μm diameter), and coated with a bio-membrane, r porous reaction layer. The working electrodes are surrounded with platinum grid that is used as a counter clectrode. Voltage between the working and counter electrode is set a 0 V. A principle of action of ElectraSense^{TM} platform of DNA/RNA hybridization assays is depicted in Figure 5. An oligonucleotide probe on and electrode specifically captures a biotin labeled target molecule from the hybridization solution. This bound target is subsequently labeled with HRP using biotin-avidin

chemistry. HRP then catalyses the oxidation of tetramethylbenzidine (TMB). Due to the close proximity of HRP to the electrode surface, the oxidized TMB is easily reduced by application of a reducing potential. The generated current flux enables to quantify the HRP label and, therefore, to assay the labeled bio-components bound to the microarray (12, 17, 20).

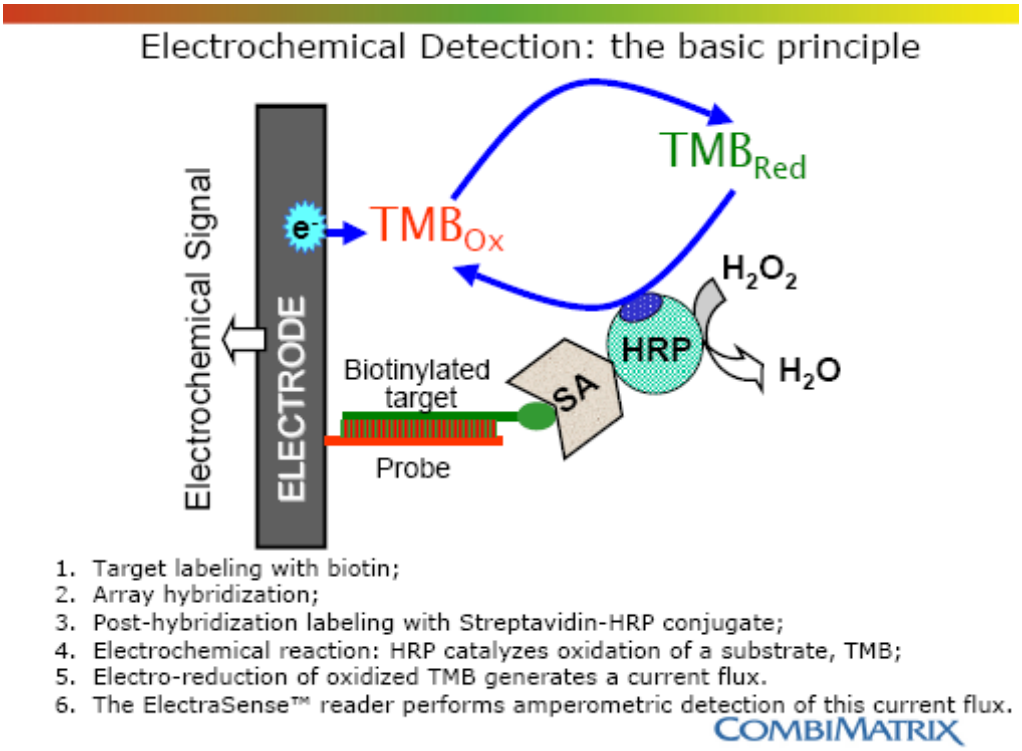


Figure 5 A general scheme of electrochemical detection for CombiMatrix ElectraSense™ microarray (20)

3.10 Bioinformatic Tool

3.10.1 Basic Local Alignment Search Tool (BLAST)

BLAST takes a query sequence and aligns it with each sequence in the database, looking for segments of high degrees of similarities. It picks out from the database those sequences that contain a segment so similar to part or the entire query that such similarity is deemed statistically significant (i.e., unlikely to occur by

chance). A BLAST search is often the most convenient method for detecting homology of a biological sequence to existing characterized sequences. BLAST looks for homology by searching for locally aligned regions of identity and/or similarity between a query sequence and sequences in a database (92).

3.10.2 Genotype nomenclature

The genotype of an influenza A viral strain is the sequential aggregate of the eight assigned gene segment lineages. A nomenclature for influenza A viral genotypes will allow researchers to unequivocally describe influenza A viral genotypes to analyze, compare and communicate the molecular epidemiology of the virus. Define a nomenclature for influenza A viral genotypes and describe a web tool developed for genotyping influenza A viruses from genome sequences. This tool facilitates identification of reassortment events between divergent lineages.

Two nomenclature conventions are used routinely in influenza research: (i) the eight segments in the influenza A genome are numbered from 1 to 8 for PB2, PB1, PA, HA, NP, NA, M and NS, respectively; (ii) There are currently 16 alleles of the HA gene termed subtypes. Likewise, there are nine alleles for NA, and two alleles for non-structural (NS) proteins. Since influenza A viruses have an unusual genomic structure, we approached the genotyping problem by first analyzing each gene segment separately. According to the above, conventions and considering that the evolutionary rate varies from segment to segment here defined a genotype as a sequential combination of the lineages for each of the eight segments in a genome. A letter was assigned to each lineage of PB2, PB1, PA, NP and M, and a number followed by a letter was assigned to each lineage of HA, NA and NS with the number representing the subtype or allele. For example, [A, D, B, 3A, A, 2A, B, 1A] is the genotype of a human seasonal subtype H3N2 virus with PB2 lineage A, PB1 lineage D, PA lineage B, HA subtype 3, lineage A and so on, following the convention for numbering of influenza genome segments. With this nomenclature, identifying genotypes and reassortment becomes an easy task accomplished by comparing the predicted genotype against all genomes that have been classified previously (92, 93).

3.10.3 FLAN (short for FLu ANnotation), the NCBI

Web server for genome annotation of influenza virus (http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/annotation.cgi) is a tool for user-provided influenza A virus or influenza B virus sequences. It can validate and predict protein sequences encoded by an input flu sequence. The input sequence is BLASTed against a database containing influenza sequences to determine the virus type (A or B), segment (1 through 8) and subtype for the hemagglutinin and neuraminidase segments of influenza A virus. For each segment/subtype of the viruses, a set of sample protein sequences is maintained. The input sequence is then aligned against the corresponding protein set with a 'Protein to nucleotide alignment tool' (ProSplign). The translated product from the best alignment to the sample protein sequence is used as the predicted protein encoded by the input sequence. The output can be a feature table that can be used for sequence submission to GenBank (by Sequin or tbl2asn), a GenBank flat file, or the predicted protein sequences in FASTA format (94).

3.10.4 The Flu Genome database

It contains three tables: Segment, Genome and Genotype. The Segment table contains information-related to sequences, including assigned lineage, strain name, segment, serotype, host, country, year, GenBank accession number, nucleotide sequence and sequence length. The Genome table contains the information for complete genomes, including assigned genotype and accession numbers of each gene segment. When more than one sequence was available for a gene segment, the longer of the two sequences was kept for the genome accession. Unique genotypes are stored in the Genotype table along with the total number of genomes that have that genotype. The Genotype table was created by querying the Genome table for distinct genotypes. Host categories were created to separate the genomes of each genotype, which include Human (Hu), Avian (Av), Swine (61), Equine (Eq), Canine (Ca) and Others (ONHM). The FluGenome database is updated automatically every night. New sequences are downloaded from the NCBI Influenza Virus Resource

(ftp://ftp.ncbi.nih.gov/genomes/INFLUENZA/) and added into the FluGenome database. The lineage information predicted for new sequences is used to update Segment, Genome and Genotype tables if necessary. For sequences already in the database, the script checks to see what information needs to be updated, and the sequences entries are flagged for further validation (93, 95).

CHAPTER IV
MATERIALS AND METHODS

4.1 Study design

The 8 H5N1-RNA samples were collected and extracted from birds and mammals samples, which were used to analyze H5N1 genotypic testing with semiconductor-base oligonucleotide microarray. The results were used to evaluate tendency for using the Bioedit.

Avian influenza A subtype H5N1 was analyzed by using the CombiMatrix Influenza A Sub-typing Primer Kit. This test kit is based on reverse transcription and PCR Technology used to target preparation for hybridization with the Influenza A research Microarray. There are 3 major processes as followed: for each sample, the two amplified biotin-labeled targets are combined and hybridized to the same Influenza A Research Microarray 12K and finally fluorescent detection of the oligonucleotide probe-bound amplified DNA by electrochemical detection.

4.2 Influenza A/H5N1 samples

H5N1 samples were originally received from the Faculty of Veterinary and Faculty of Science, Mahidol University, Nakorn Prathom, Thailand, Department of Pathology, Faculty of Veterinary Medicine, Kasetsart University, Nakorn Prathom, Thailand, and Department of Livestock Development, National Institute of Animal Health, Bangkok, Thailand.

4.3 Statistic Analysis

The quantitative detection of influenza viruses used different method of oligoneucleotide signal intensity. The tests to identify H5N1 intensity were measured by the mean and the median. The best cut off intensity for each test was assessed by statistical optimisation of sensitivity; specificity and the kappa value which available procedures from **STATA®**.

4.4 The CombiMatrix Microarray

CombiMatrix microarrays are based on silicon chip technology-the integrated circuits contain arrays of microelectrodes that are individually addressable using embedded logic circuitry on the chip. Placed in a specially designed fluidic chamber, the chip directs the synthesis of DNA in response to a digital command.

During synthesis each microelectrode is addressed to selectively generate chemical reagents by means of an electrochemical reaction. The electrochemical reaction is used to generate acid during the detritylation step of standard phosphoramidite synthesis. DNA is synthesized above a platinum electrode, withina boundary layer that coats the chip. Under the direction of a PC theCustomArraychip rapidly synthesizes 12544 different DNA molecules in parallel, each above a distinct electrode. Since a different product can be synthesized at each site, the technology eliminates the need to synthesize 12544 different oligos by conventional methods. This means arrays can be made inexpensively to any sorts of organism (e.g. flu) and bioinformatics approaches can be tested and optimized easily without having to purchase expensive oligo sets.

A generic PCR based amplification (designed to be pan-influenza strain specific) followed by specific hybridization on a microarray designed to precisely type influenza. Probes on the chip are designed to differentiate viral strain type based on the HA (types 1-16) and NA (types 1-9) serotype classifications of influenza strains. Chip designed against 1614 animal and 1937 human hemaglutinin sequences, and

against isolates 522 human and 831 animal neuraminidase sequences. Chip is laid out in a pattern to aid visual identification of HA and NA type. Chip can be read fluorescently or using a simple electrochemical detection (direct reading of the chip through an electrochemical sensing process)

4.5 Influenza A Nucleotide Sequence database

There are three groups of database for bioinfomatic analysis

1. Data of Hemagglutinin and Neuraminidase genes influenza A viruses nucleotide sequences were retrieved from the NCBI web sever for genome annotation of influenza virus (http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/annotation.cgi). This web server is a tool for user-provided influenza A virus or influenza B virus sequences. Total cases of influenza A virus are HA gene 6,463 sequences and NA gene 5,403 sequences. The database contains all of subtypes of the hemagglutinin and neuraminidase nucleotide sequences.

2. 168 Nucleotide probe sequences of H5 and 168 Nucleotide probe sequences of N1 from CombiMetrix array.

3. Sample sequences.

4.6 Nucleotide sequence analysis with Bioedit version 7.0.0

Program Bioedit version 7.0.0 was used for nucleotide sequence analysis including sequence comparison and local BLAST with the ability to compile local databases from Fasta files. BioEdit is a biological sequence editor that runs in Windows 95/98/NT/2000/XP and is intended to provide basic functions for protein and nucleic sequence editing, alignment, manipulation and analysis. They offer many quick and easy functions for sequence editing, annotation and manipulation, as well as a few links to external sequence analysis programs. The Bioedit is offered free of charge.

Table 1 Samples for this study

Sample no.	Virus name
11	Influenza A (A/Tiger/Thailand (Supanburi)/vsmu-1/2004 (H5N1)
17	Influenza A (A/Tiger/Thailand (Chonburi)/vsmu-23/2004 (H5N1)
25	Influenza A (A/ Brown-headed Gull /Thailand (Samutprakarn)/ vsmu-28/2005 (H5N1)
27	Influenza A (A/Open-bill strok/Thailand (Nakornsawan)/vsmu-30/2005(H5N1)
31	Influenza A (A/duck/Thailand(Bangkok)/2003 (H5N1)
37	Influenza A (A/duck/Thailand(Bangkok)/2003 (H5N1)
47	Influenza A/duck/Thailand (Bangkok)/2006(H5N1)
48	Influenza A/Chicken/Thailand (Bangkok)/2006(H5N1)
H1N1	Isolate from culture
H4N8	Isolate from culture
H5N1	Isolate from culture

4.7 Method

4.7.1 The Low Resolution (genotyping)

4.7.1.1 Single step preparation

For target preparation, total RNA isolated from virus-containing material is reverse-transcribed into cDNA, which is then PCR-amplified to produce biotinylated single-strand DNA. Since the influenza-specific reverse primers are grouped into two different pools (to increase sensitivity of the multiplex reactions), each sample is amplified in two tubes, one for HA subtypes 1 to 7, and another one for HA subtypes 8 to 16, and NA subtypes 1 to 9. Amplification is performed as a single-step procedure including reverse transcription and PCR. The reverse transcription conditions were 10-min at 40ºC, 20-min at 50ºC and 3-min at 94ºC. 1st PCR thermal cycling condition were 30 cycles of a 30-s 94°C denaturation step, a 45-s 50°C annealing step and a 30-s 72°C extension. Then 2nd PCR consisted of 40 cycles of a 30-s 94°C denaturation step, a 60-s 65°C annealing and extension, and finally, a 5-min extension at 72°C and 4°C after cycling was performed. The presence of PCR

products was confirmed on 4% Nusieve Seakem 3:1 agarose gels (FMC, Rocland) in 0.5XTBE (0.089 M Tris-borate, 0.089 M Boric acid and 0.002 MEDTA, pH 8), and stained with 0.5 µg/ml ethidium bromide. Gels were run at 100 V for 45 min. PCR yields were estimated by comparison with a low DNA mass ladder (Promega, USA).

Table 2 Two reaction mixtures for single step RT-PCR

Reagent	**H1-7 (50 µl)**	**H8-16+N1-9 (50 µl)**
Nuclease-free water	13.5 µl	13.5 µl
Template (total RNA)	2.5 µl	2.5 µl
2X Reaction Mix	25 µl	25 µl
0.4 mM Biotin-14-dCTP	5.0 µl	5.0 µl
Influenza Forward primer (Tube 1)	1 µl	1 µl
Influenza Reverse Primer Pool (Tubes 2 and 3)	H1-7: 2 µl	H8-16+N1-9: 2 µl
RT/ Platinum® Taq Enzyme Mix	1 µl	1 µl
Total Volume	50 µl	50 µl

4.7.1.2 Pre-hybridization

The single-stranded target was heated to 95°C for 5 min and place on ice. The chip microarray was incubated in the hybridization chamber with nuclease-free water at 65^{o}C for 10 min. Remove the microarray from the incubator and bring to room temperature. Fill the hybridization chamber with the Pre-hybridization Solution (50X Denhardt's solution, 1%SDS, 6X SSPE, 0.05% Tween-20, 20mM EDTA and nuclease free water) and loaded onto the rotisserie in the hybridization oven and incubate at 45^{o}C for 15 min with gentle rotation.

4.7.1.3 Hybridization

For each sample, mix together 10 µl aliquots of the purified target samples (H1-7 and H8-16+N1-9). Denature the HA-NA target mixture by incubating at 95°C for 5 min and place on ice. Each of samples were diluted in hybridization solution (2X Hyb Solution Stock, 1% SDS and Positive Hybridization

Control from the CombiMatrix Influenza A Sub-typing Primer Kit) to a final volume of 50 ul and load the hybridization chamber onto the rotisserie in the hybridization oven and follow by hybridization of biotinylated target in pre-block solution for 1 hour at 45^{o}C with gentle rotation.

4.7.1.4. Hybridization Washing

The array were then washed once for 5 min with 6X SSPET (6X SSPE, 0.05% Tween-20) at 45^{o}C and then 30 s each with 3X SSPET, 0.5X SSPET, PBST and PBS at room temperature.

4.7.1.5 Post-hybridization Blocking, Labeling, and Washing

The hybridized array was then blocked with 5X PBS-Casein Blocking Buffer for 15 min at room temperature and labeled for 30 min with the Fluorolink™ Cy5®-labeled streptavidin diluted 1:1000 dilution (v/v, 1 µl per 1 ml) in the 5X PBS-Casein Blocking Buffer and incubate the microarray at room temperature for 30 minutes. Protect the microarray from light by covering with aluminum foil. Remove the Dye Labeling Solution from the hybridization chamber. The arrays were washed twice with 2X PBST and twice with 2X PBS.

4.7.1.6. Imaging of Microarray 12K

The microarrays were scanned for probe and target fluorescence intensity with a scanner. After complete the scan then, save the image as a .tiff image file. The data can be extracted from the image using the CombiMatrix Microarray Imager Software. After imaging is completed, we can proceed to stripping of the microarray for subsequent rehybridization using the CombiMatrix CustomArray™ Stripping Kit (see Appendix A). Do not allow the semiconductor microarray surface to dry; keep the microarray wet, either in a tube or slide-holder containing 1X PBS, or in the Imaging Solution with the LifterSlip™ attached. Avoid prolonged storage of hybridized microarrays prior to stripping; instead, first strip the

microarray, then store wet in Imaging Solution or 1X PBS at 4^{o}C for a maximum of 2 weeks.

4.7.1.7 Data analysis

Image intensities were quantified with Microarray Imager (Combimatrix Corp.) and graphed with Microsoft Excel software. Subtype identification was accomplished by averaging HA and NA subtype DNA sequence information was generated from array intensity data with an Excel and then associate the most intense signal with the nucleotide represented by that probe. Sequence strings were then used to search the GenBank nonredundant database with the BLASTN program.

4.7.2 Microarray output combined with bioinformatics tools

Our common identification framework of influenza A/H5N1 sequences is depicted in Figure 6.

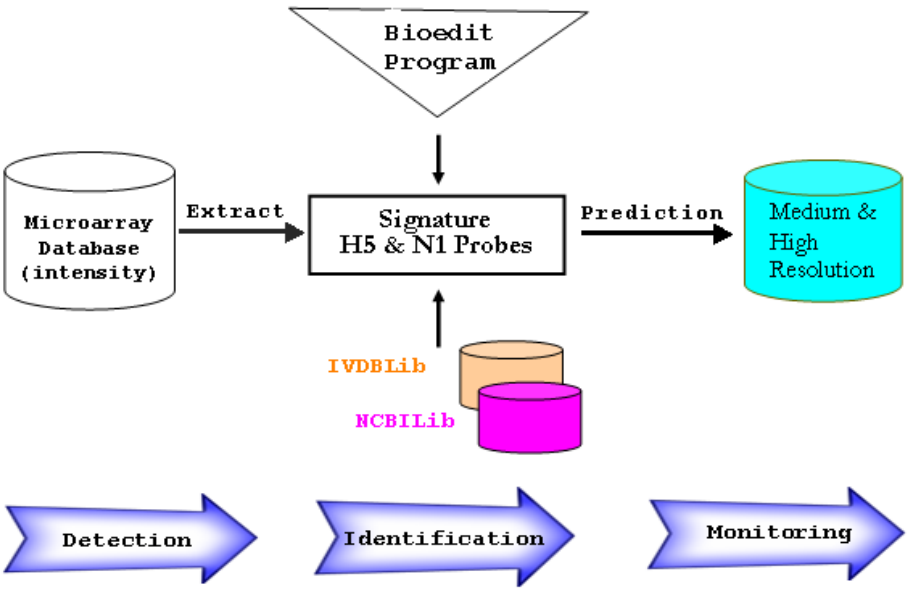


Figure 6 Identification framework of influenza A/H5N1 sequences

4.7.2.1 Statistic analysis to indicate the cut off value

Sensitivity, specificity positive and negative predictive values, and Kappa are the measures for assessing the results of diagnostic according to signal intensity. Sensitivity represents the proportion of truly diseased avian animals in a screened population who are identified as being diseased by the test. It is a measure of the probability of correctly diagnosing a condition. Specificity is the proportion of truly non-diseased avian animals that are so identified by the screening test. It is a measure of the probability of correctly identifying a non-diseased animal. Predictive value of tests in screening and diagnostic tests is the probability, that a person with a positive test is a true positive (i.e., has the disease), is referred to as the predictive value of a positive test (PPV: Positive predictive value); whereas, the predictive value of a negative (PNV) test is the probability that the animal with a negative test does not have the disease. Predictive value is related to the sensitivity and specificity of the test.

Kappa is a measure of the degree of non-random agreement between observers or measurements of the same categorical variable. Kappa is the proportion of agreement after chance agreement has been removed. If kappa=1, there is perfect agreement. If kappa = 0 the agreement is the same as would be expected by chance. The stronger the agreement, the higher the value of kappa, Negative values occur when agreement is weaker than expected by chance, but this rarely happens. Depending on the application, kappa less than 0.6 may indicate that the measurement system needs to be improved. Kappa values that greater than 0.8 are generally considered very well. The observed prevalence is the total proportion of animal oligoneucleotides found positive by both methods, and the true prevalence is the proportion of animal whose oligoneucleotides are avian flu.

Table 3 Basic contingency table and formulae

	Flu oligoneucleotide	
Cut off signal intensity test	T = +	F = -
T = +	a	b
F = -	c	d

N=a+b+c+d

Sensitivity	a/(a+c)
Specificity	d/(b+d)
Observed prevalence	(a+b)/N
True prevalence	(a+c)/N
Positive predictive value	a/(a+b)
Negative predictive value	d/(c+d)

4.7.2.2 Sequencing model

The experiment microarray data used in the present analysis were obtained using a variety of purified nucleic acid templates as oligonucleotide probes. To circumvent problems of analyzing unknown sequences (de novo sequence analysis), PCR products with a known reference sequence (resequencing analysis) were examined. But resequencing model provides base call resolution by comparing the intensities of H5 and N1 oligonucleotide probes was established. The files of scanned intensities were transferred to the ElectraSense TM for subsequent analysis. To analyze the results, we needed to determine whether a particular intensity value was indicative of hybridization. For this purpose, Base calls were made based on the intensity data of each probe. The oligonucleotide probes were represented in FASTA format.

Resequencing model is to transform the H5 and N1 oligonucleotide probes FASTA format into Bioedit program as the data representation of Influenza A/H5N1 sequences. In this study, we divided the oligonucleotide probes of each sample into ten groups by rank order of intensity (1000-10,000) and use them to generate signature “consensus” sequences for each group of intensity. This process is divided into three parts.

4.7.2.3. Overlapping probes

After genotype was determined by hybridization signal intensity produced by colourimetric assay at each strain specific probe. To test the sequencing approach in high density oligonucleotide arrays with the appropriate Bioedit program is use to evaluate oligonucleotide probe for all possible sequence H5N1 variation in Thailand during 2003-2006. Although it is possible to use oligonucleotide microarrays in mutation screen, overlapping probe must be improved as Figure 7. Initially, all sequences (the 168 oligo H5 probes and the 168 oligo N1 probe) of the oligonucleotide probes are considered as candidate sample sequence. Sets of oligo probe sequences that were collected in FASTA format. To analyze data sets from FASTA file, each oligo is first examined for the presence of complementary with H5 and N1 consensus. After that, set of H5 and N1 complementary oligo probes were rely by order of base position corresponding with H5 and N1 consensus.

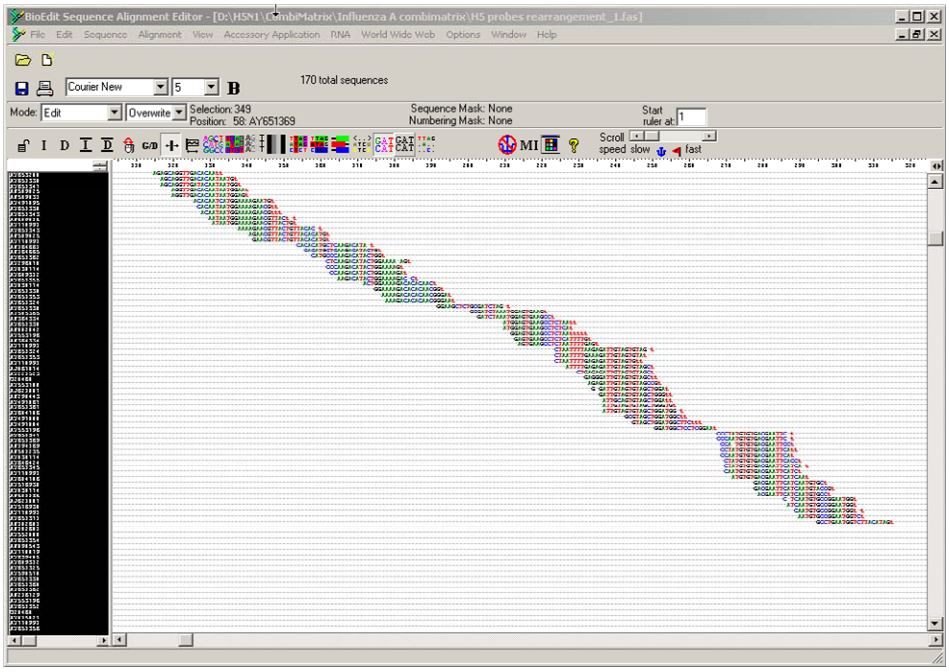


Figure 7 Overview of overlapping probe

4.7.2.4 Rank and group of probes intensity

For each sample, we divided positive signal intensity into ten groups, 1000-10,000 as Figure 8. To create consensus sequence, the oligonucleotide of each group of influenza sequences is taken multiple alignments using the Bioedit program as the pre-process of systems for finding a "consensus" sequence. The consensus sequence is nearly a signature, but it is created by taking the majority bases in the multiple alignments of a sequence collection.

Figure 8 Overview of rank and group of probes intensity

4.7.3 Database preparation

Our Sequence databases are selected from the Genbank genome database of National Center for Biotechnology Information (NCBI). In this study, we use influenza A virus is distinctly separated into Hemagglutinin (HA) and Neuraminidase (NA) segments. We picked 6,463 HA gene sequences, 5,403 NA gene sequences. For our sample, we picked 5 HA and 5 NA genes, then inserted into the

sequence database which is used for performance evaluation of signature “consensus” sequence. Figure 9 show scheme of database preparation.

4.7.4 Identification task

It is process of creating Identifiers using the signature “consensus” sequences and Bioedit program. The identification task is divided into two consecutive parts: Medium resolution and High resolution, to properly correct sequence of influenza A/H5N1 in each sample in term of % identity and E-value of BLAST tool, data for which we know the “truth’ is required. In this study we examine three data sets for which assessments can be performed where specific results are expected. Data set A provides probes for which we can assume the measurements are entirely due to non-specific binding. This permits us to study the variety intensity measures. Data set B provides the results of a spike-in experiment where gene fragments have been added at known sequence. These data can be used to assess correctly group of influenza A/H5N1 in Thailand during 2003-2006 (medium resolution). Data set C provides the results from a study in which samples were hybridize at different host. This permits us to assess similarity between consensus template and set of subsequences in each sample. To reduce the impact of repeated nucleotide of sequence in each sequence of Influenza A virus database, this resulted in database of Influenza A virus for HA and NA sequences. The robustness of each sequence was made manually editing.

4.7.4.1 Medium resolution.

Finding an oligonucleotide representation of DNA sequences is performed in a resequencing model. The oligonucleotide representation, called base call, is a substring of the DNA sequences. Members of the oligonucleotide representation were transfer to Bioedit program and then pass to multiple alignments generates the signature “consensus” sequences. Each sample generates the 8-10 consensus sequences. Sequence alignment were created using the Bioedit programs and the adjusted manually using the edit alignment according to nucleotide sequences.

Finally, All subsequence candidates passing the previous tests are examined for correctly group. The database query subtask performed a batch similarity search of a database using subsequence as the queries. The BLAST program used was the NCBI blastn version 2.12 via Bioedit with a defined set of parameters (figure 9). The masking of low complex regions was performed for the seeding phase to speed up the query; however, low complexity repeats were included in the actual scoring. The entire nucleotide database from NCBI acquired on February 7, 2006 was used as the reference database. The default gap penalty and nucleotide match score were used. The nucleotide mismatch penalty, -q, parameter was set to -1 rather than the default. The results of any BLAST query with an expect value <0.0001 were returned in tabular format from the blast all program. The information about each return (bit score, expect value, mismatches, length of match) was placed in the Return hash using the subsequence identity as the hash key for further analysis.

4.7.4.2 High resolution

The goal is to learn how RNA population differs in sequence in each sample in different host which was designed to identify unique sequence. The signature concensus is to measure the significance of each sample, called “% Identity” using variously different percent and select a set of the consensus sequence highest percent as our “Signatures”. We then pass the signatures through the pairwise process which uses the signatures for formulating a sequence identity score. If there is a query sequence, the highest identity scoring is to detect whether it is the target (influenza A/H5N1, for our experiments) or not.

When there is a query sequence, the scoring is used to measure a similarity value between the query sequence and target sequences, If the score is higher than a threshold, the query is identified as a member of the target sequences. Assume that we pass a query sequence through and identification system; subsequences that only have returns with highest identity score with pairwaise alignment in Bioedit program.

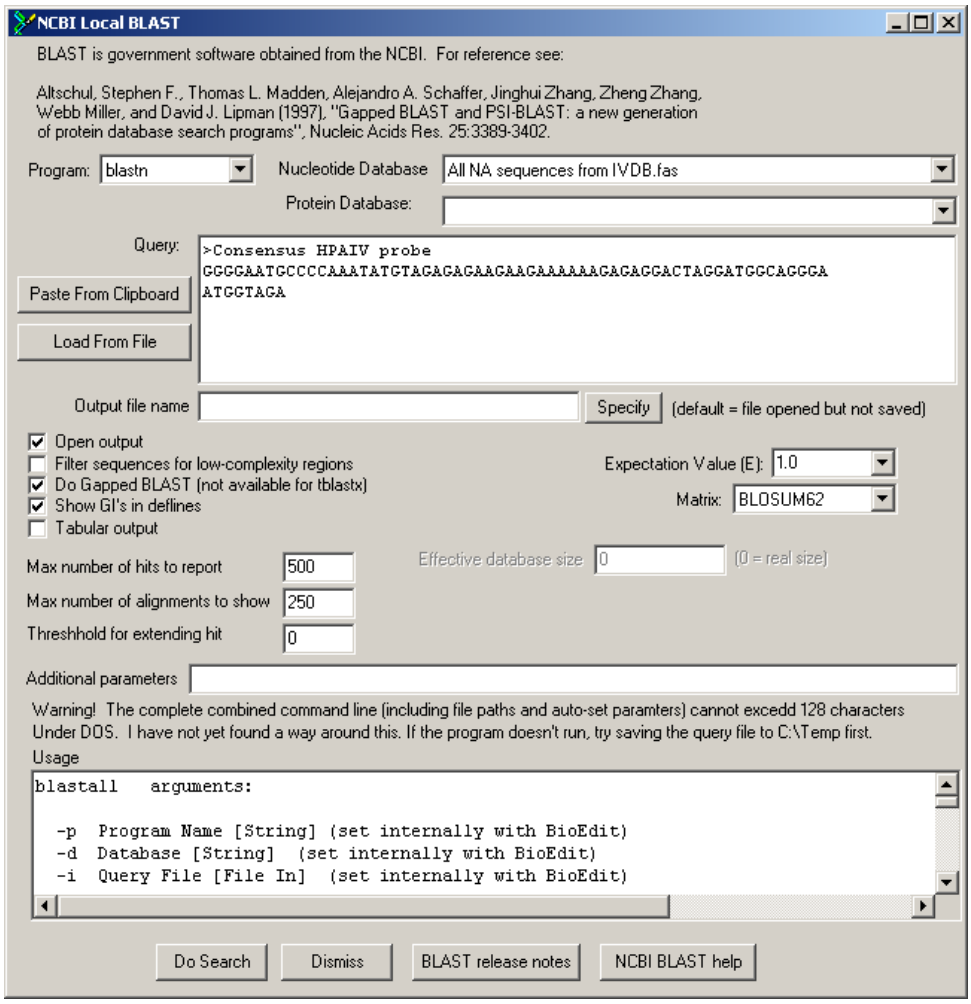


Figure 9 Screen short for Local BLAST using in Medium resolution

CHAPTER V
RESULTS

5.1 Genotyping

To assess the capacity of the semiconductor-base oligonucleotide for detection and typing of influenza A/H5N1 virus, we tested 8 influenza A/H5N1 references strains of animal origin and two non-H5N1 subtypes (H4N8 and H1N1). The avian strains represented chicken, duck, Brown-headed Gull and Open-bill stork and mammal strains represented tiger. The amplification of known influenza A virus isolates was accomplished with a pool of reverse primers at a concentration of 100 µM each. Two pools were used: one pool for the HA gene and one pool for the NA gene. All known targets were successfully amplified, as revealed by agarose gel electrophoresis (Figure 10). The reaction show one or more bands at approximately 500-600 bp. And when they were hybridized to the arrays, they were identified as subtype H5N1. The hybridization patterns of the eight known samples indicated that the HA and NA sequences from samples no. 11, 17, 25, 27, 31, 37,47 and 48 were very similar. We also tested H1N1 and H4N8 viruses isolated from culture for negative control. Avian and mammal influenza virus strains were accurately detected. All three HA (H1, H4 and H5) and two NA (N1 and N8) influenza A subtypes tested were correctly identified. Very distinct hybridization patterns were also seen for hemagglutinins H1, H4 and neuraminidase N1, N8, indicating that the arrays could potentially be used to identify subgroupings of influenza virus subtypes (Figure 12, 13). Average signal was vary from 2,000 to 10,000 for H5 and 1,000 to 10,000 for N1 (Table 4).

Visual HA and NA subtype determinations were made in two ways: (i) by marking the array images with the location of subtype-specific probes and (ii) by graphing of the subtype probe intensities. In addition, bar graphs of mean subtypes

intensity values could be used to predict the correct HA and NA subtypes. All 3 HA subtypes generated from known isolates could be correctly identified with the array, as could as 3 NA subtypes. The results of this genotyping assay are presented in Figure 11-21.

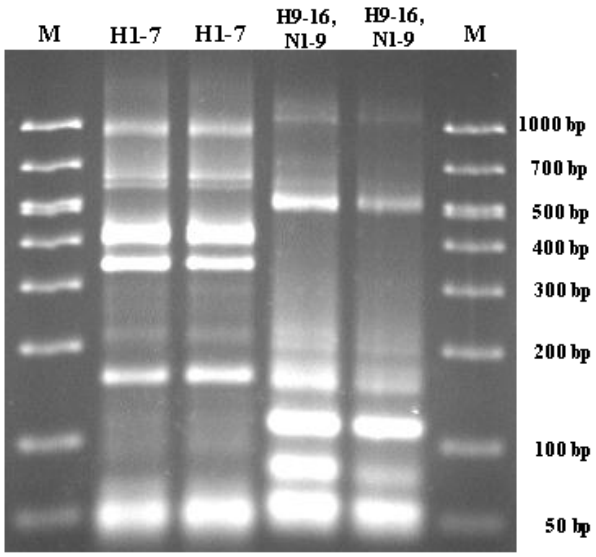


Figure 10 Gel electrophoresis of multiplex RT-PCR with H1-H16 and N1-N9 of primers products for several influenza virus samples

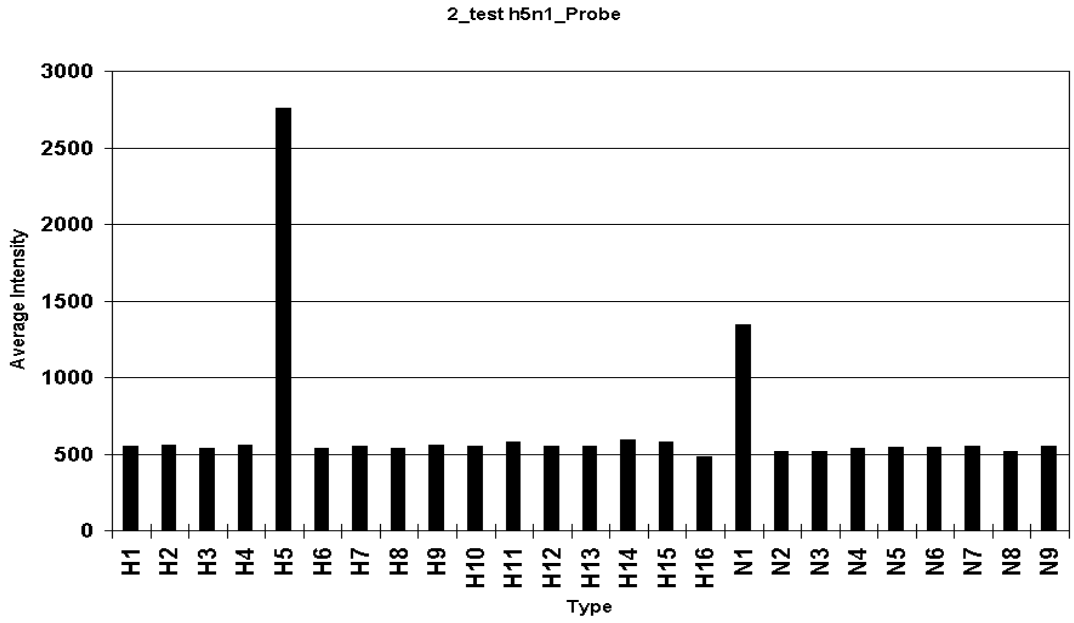


Figure 11 The average signal intensity of subtype probe group in positive control H5N1

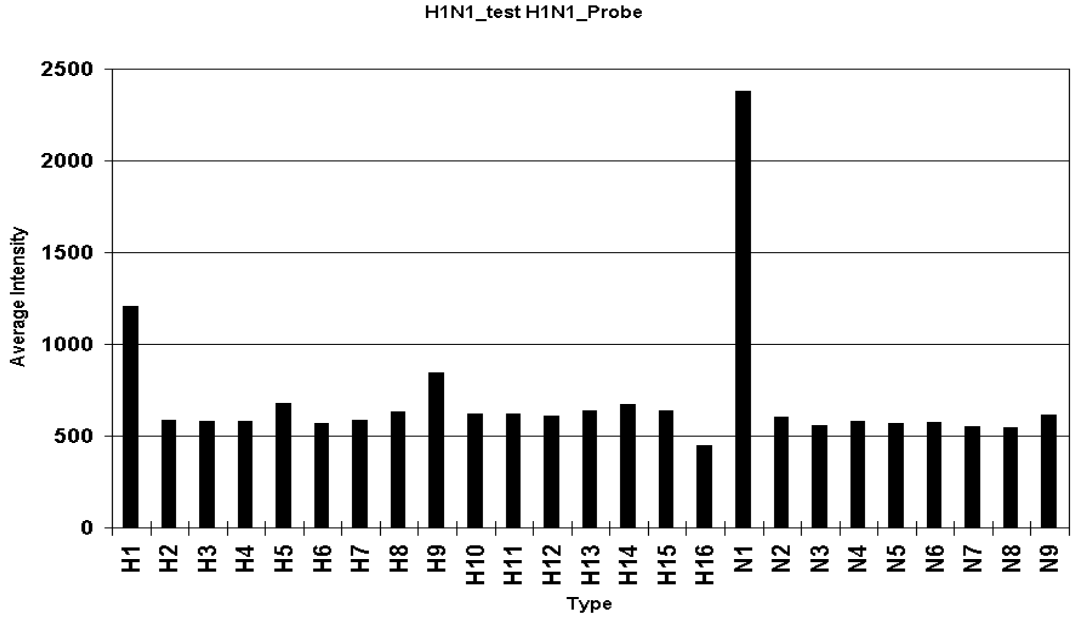


Figure 12 The average signal intensity of subtype probe group in negative control H1N1

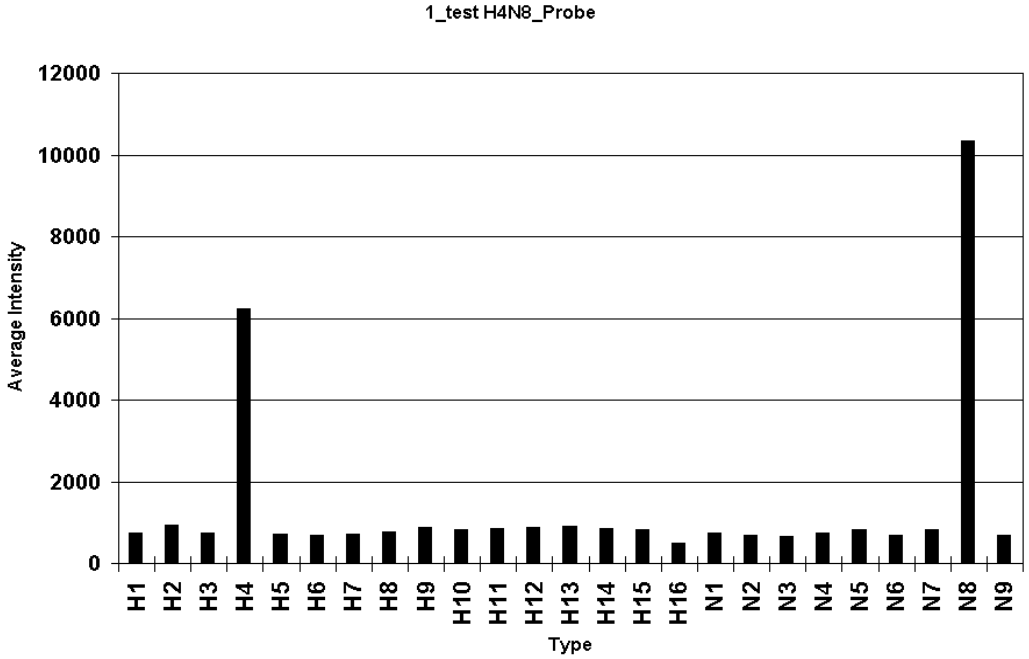


Figure 13 The average signal intensity of subtype probe group in negative control H4N8

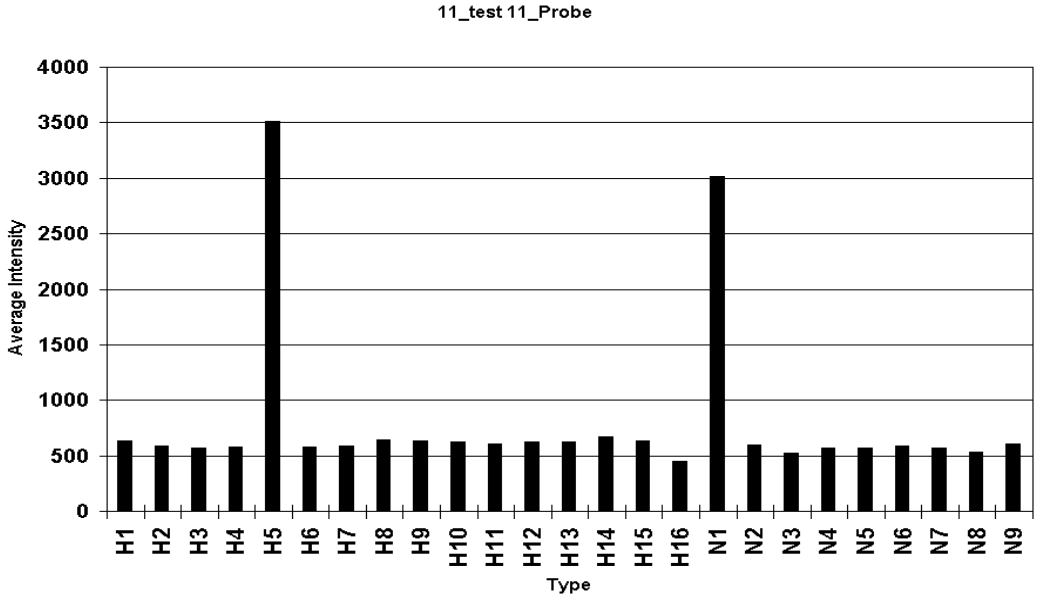


Figure 14 The average signal intensity of subtype probe group in sample no. 11

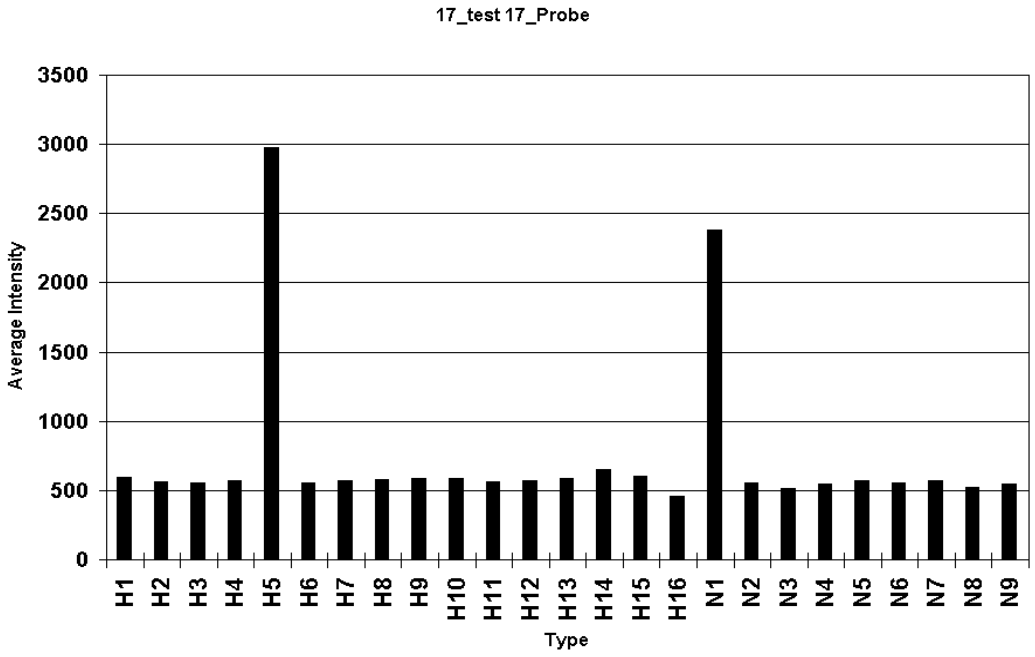


Figure 15 The average signal intensity of subtype probe group in sample no. 17

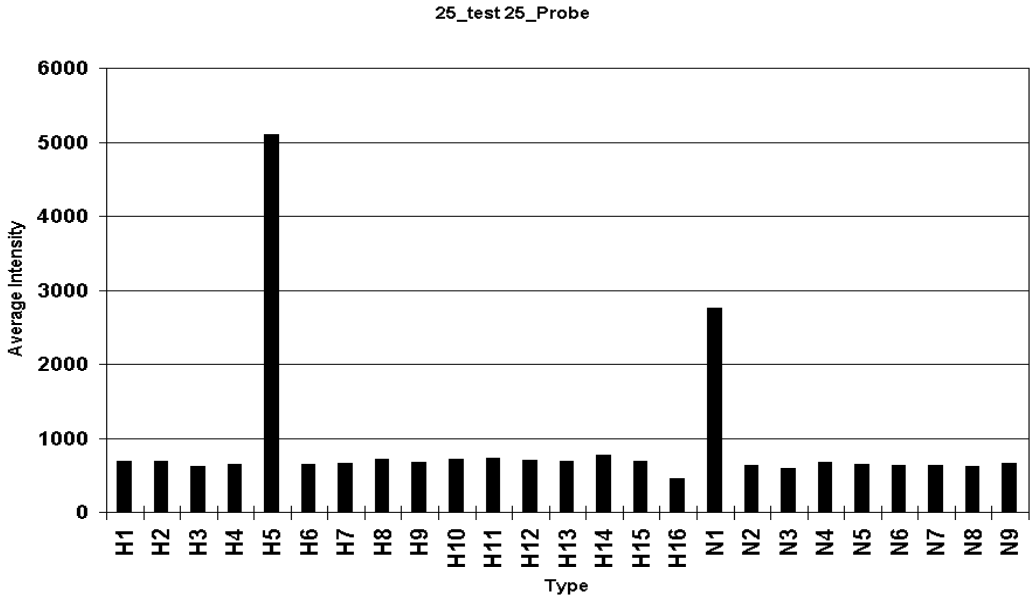


Figure 16 The average signal intensity of subtype probe group in sample no. 25

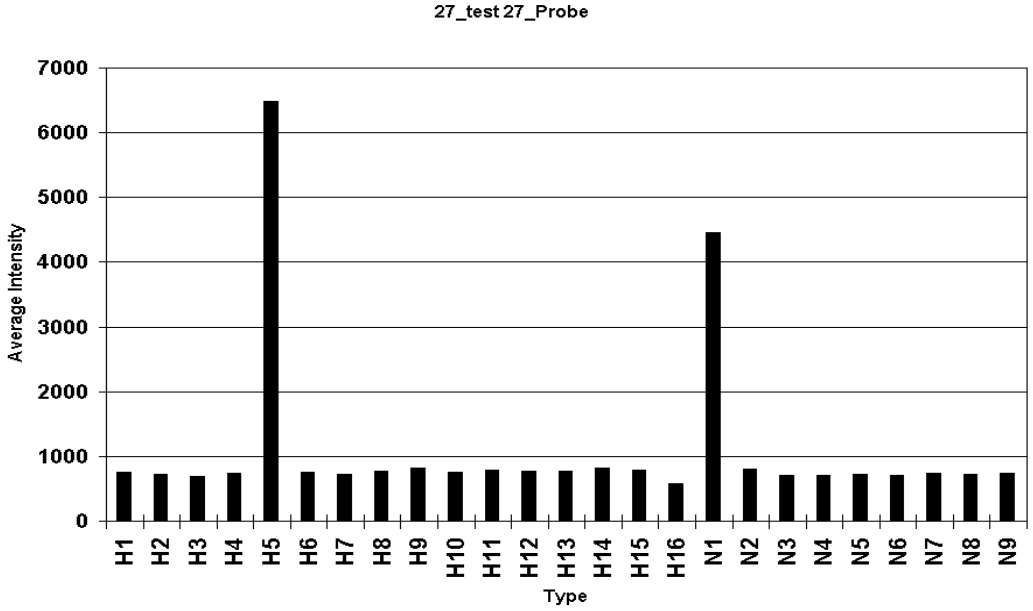


Figure 17 The average signal intensity of subtype probe group in sample no. 27

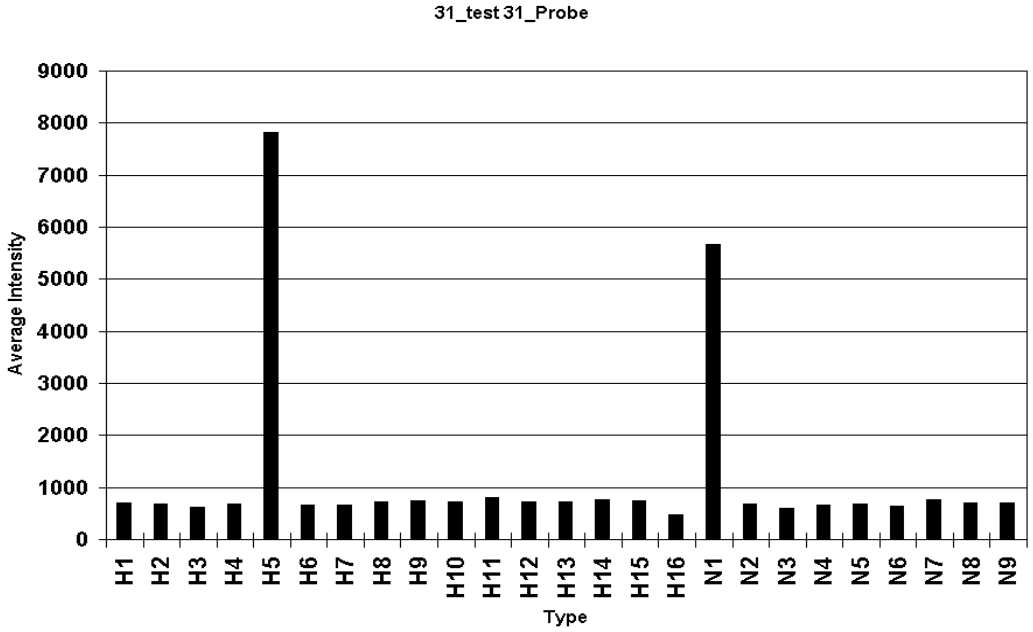


Figure 18 The average signal intensity of subtype probe group in sample no. 31

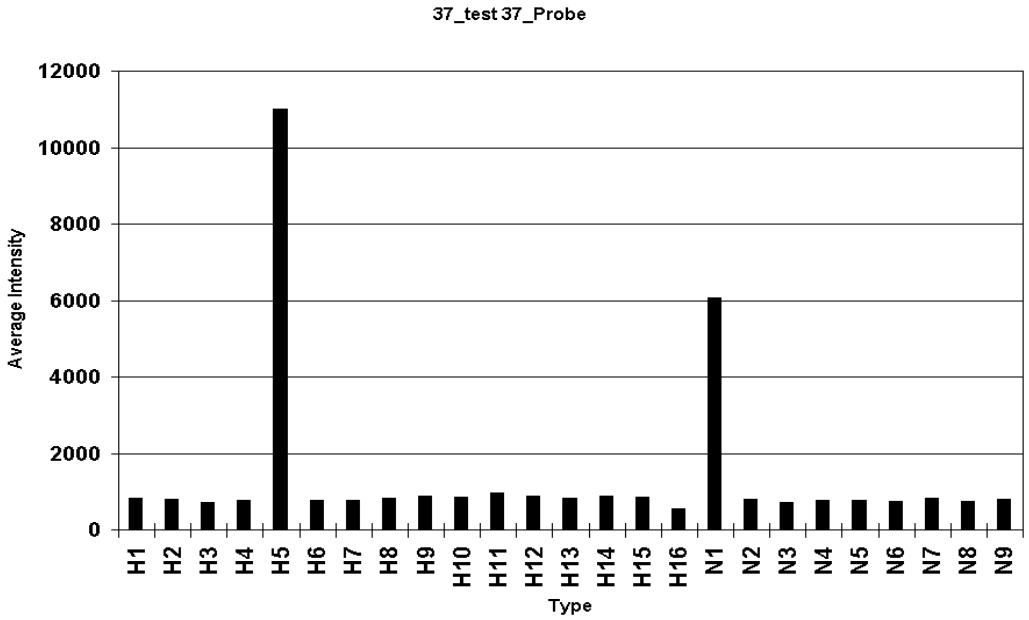


Figure 19 The average signal intensity of subtype probe group in sample no.37

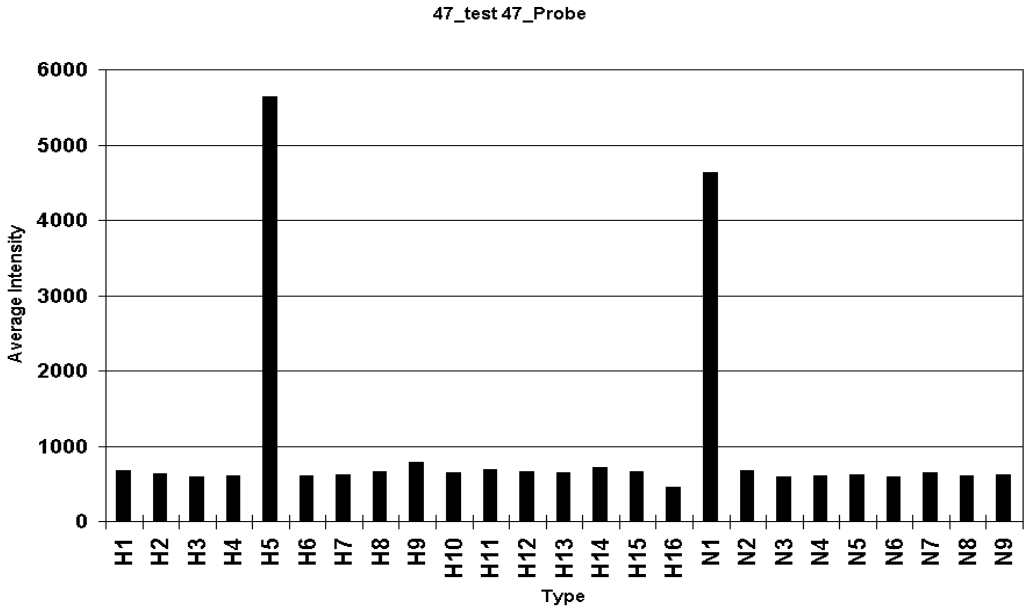


Figure 20 The average signal intensity of subtype probe group in sample no.47

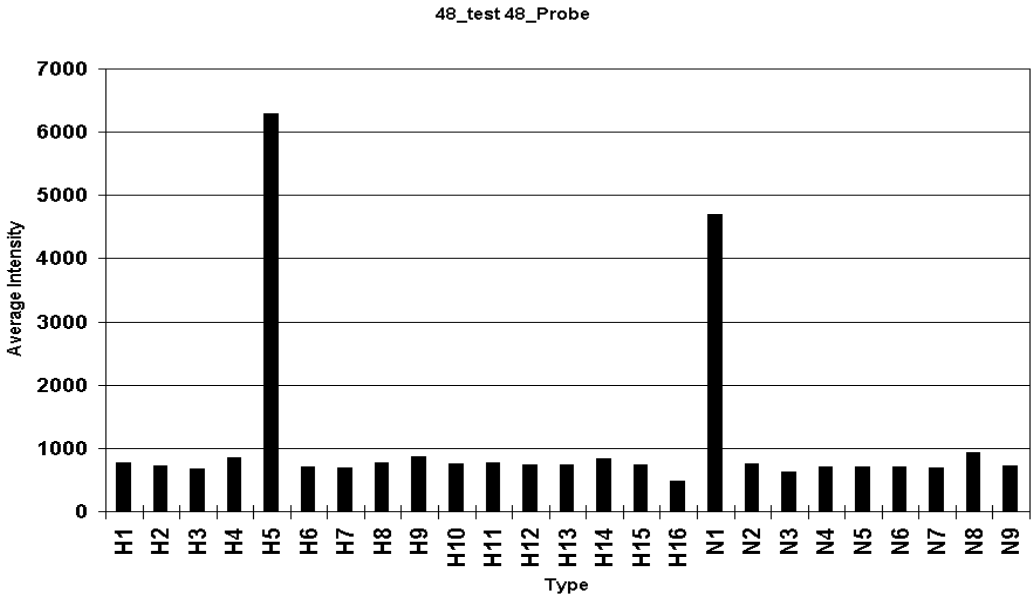


Figure 21 The average signal intensity of subtype probe group in sample no. 48

Table 4 Summary of the genotyping result using Combimatrix chip

Sample no.	Call	Assay Baseline	Type	Average Signal	Average/ Baseline	Median Signal	Median / Baseline	StDev
H5N1 test	H5N1	377	H5	2760	7.31	2245	5.95	2134.41
			N1	1342	3.56	634	1.68	1595.46
Non- H5N1 test	H4N8	424	H4	6244	14.73	4028	9.5	5556.93
			N8	10339	24.39	11039	26.05	4356.03
Non- H5N1 test	H1N1	373	H1	1206	3.23	701	1.88	1409.15
			N1	2378	6.38	651	1.74	3475.35
No.11	H5N1	386	H5	4255	11.02	4591	11.89	2252.7
			N1	3146	8.15	2249	5.82	2468.45
No.17	H5N1	436	H5	3277	7.52	2855	6.55	2229.48
			N1	2618	6.01	1491	3.42	2299.18
No.25	H5N1	413	H5	5846	14.14	6133	14.83	3051.13
			N1	3119	7.54	2171	5.25	2648.6
No.27	H5N1	451	H5	6488	14.38	6841	15.16	3260.92
			N1	4460	9.89	3481	7.72	3784.1
No.31	H5N1	373	H5	7826	21	7726	20.74	4235.37
			N1	5672	15.22	5673	15.23	4523.03
No.37	H5N1	422	H5	11001	26.09	11553	27.4	4886.61
			N1	6080	14.42	5325	12.63	4859.51
No.47	H5N1	377	H5	5643	14.96	5511	14.61	3140.53
			N1	4636	12.29	3672	9.73	3524.38
No.48	H5N1	433	H5	6286	14.52	6353	14.67	3692.81
			N1	4701	10.86	3335	7.7	3848.1

5.2 Sensitivity and Specificity

The best method to assess H5N1 was the 47_test_47. This test was associated with a high sensitivity (87.2%, 95%CI 83.6%-90.8%), and specificity (98.2%, 95%CI 97.6%-98.8%) and a kappa of 0.86 (95%CI 0.83-0.89) showing a very good agreement of the test for a cut off mean signal intensity of 1100. For the sample no. 11, 17, 25, 27, 31, 37, 48 the best Kappa were 0.83 (cut off mean signal intensity of 900), 0.76 (cut off mean signal intensity of 1100), 0.82 (cut off mean signal intensity of 1100), 0.84 (cut off mean signal intensity of 1400), 0.82 (cut off mean signal intensity of 1400), 0.83 (cut off mean signal intensity of 1700), 0.84 (cut off mean signal intensity of 1400), respectively.

Table 5 Optimised cut off and results for different tests to identify H5N1

Optimized cut off	Sensitivity			Specificity			Obs. Prev.	True prev.	PPV	NPV	Kappa		
	%	Lower 95%CI	Upper 95%CI	%	Lower 95%CI	Upper 95%CI					value	Lower 95%CI	Upper 95%CI
Sample no. 11 mean 900	81.0%	76.8%	85.2%	98.4%	97.8%	98.9%	13.5%	15.0%	89.8%	96.7%	0.83	0.79	0.86
Sample no. 17 mean 900	69.9%	65.0%	74.8%	98.7%	98.2%	99.2%	11.6%	15.0%	90.7%	94.9%	0.76	0.72	0.80
Sample no. 25 mean 1100	78.0%	73.5%	82.4%	98.7%	98.2%	99.2%	12.8%	15.0%	91.6%	96.2%	0.82	0.78	0.85
Sample no. 27 mean 1400	80.4%	76.1%	84.6%	99.1%	98.6%	99.5%	12.9%	15.0%	93.8%	96.6%	0.84	0.81	0.88
Sample no. 31 mean 1400	78.9%	74.5%	83.2%	98.5%	98.0%	99.1%	13.1%	15.0%	90.4%	96.4%	0.82	0.78	0.85
Sample no. 37 mean 1700	81.5%	77.4%	85.7%	98.3%	97.7%	98.9%	13.7%	15.0%	89.5%	96.8%	0.83	0.80	0.86
Sample no. 47 mean 1100	87.2%	83.6%	90.8%	98.2%	97.6%	98.8%	14.6%	15.0%	89.6%	97.8%	0.86	0.83	0.89
Sample no.48 mean 1400	79.8%	75.5%	84.1%	99.0%	98.6%	99.4%	12.8%	15.0%	93.4%	96.5%	0.84	0.80	0.87

5.3 Microarray output combined with bioinformatics tools

We directly recovered hybridized viral sequence in each sample from the surface of microarray by using Bioedit program. To analysis of signature subsequences for each sample, we divided the protocol into two levels: Medium resolution and High resolution. Medium resolution experiment uses a signature sequence that can be submitted as a query to BLAST. We used the subsequence to search the own influenza A virus database using BLASTn. This approach identified target sequences of a length expected to contribute to hybridization. The BLASTn result table was processed to identify, for each hit, the oligonucleotide, and the following virus characteristic: the virus type, the source host, the RNA segment and, if applicable, it's HA and NA subtypes. This model was designed to be comfortable tool to predict novo sequences. Signature oligonucleotides were defined as the oligonucleotides that recognized only one of the possible characteristics for each sample. After alignment with Local BLAST, the data from Signature sequences were selected by high intensity signal. Specially, (535-702) the 168 oligo H5 probes and (1457-1624) the 168 oligo N1 probe were chosen to enable identification of influenza A virus H5N1 that all come from Thailand and recently isolated (2003-2006). But for first Hit for BLASTn analysis were present in table. Two output of sample no. 17 and 47 were hit in BLASTn result as Table 6.

.

Table 6 Analysis of H5N1 isolates during 2003- 2006 From Thailand

Hemagglutinin

Sample	**Accession no.**	**Strain Identification by medium resolution**
Sample 11	DQ497672	Avian/4(HA)/H5N1/Viet Nam/2003
Sample 17	salaya17	Influenza A virus (A/tiger/Thailand/VSM).
Sample 25	DQ201829	Avian/4(HA)/H5N1/China/2000
Sample 47	Arunee001	49-AI-009-E-HA-TS
Sample 48	DQ497689	Avian/4(HA)/H5N1/Viet Nam/2005

Neuraminidase

Sample	**Accession no.**	**Strain Identification by medium resolution**
Sample 11	AF098551	Avian/6(NA)/H11N1/Hong Kong/1998
Sample 17	EF124300	Avian/6(NA)/H5N1/China/2005
Sample 25	DQ334770	Avian/6(NA)/H5N1/Thailand/2005
Sample 47	EF124300	Avian/6(NA)/H5N1/China/2005
Sample 48	DQ182482	Avian/6(NA)/H5N1/Belgium/2004

Table 7 Summary of signature H5 consensus Nucleotide sequence of each sample test.

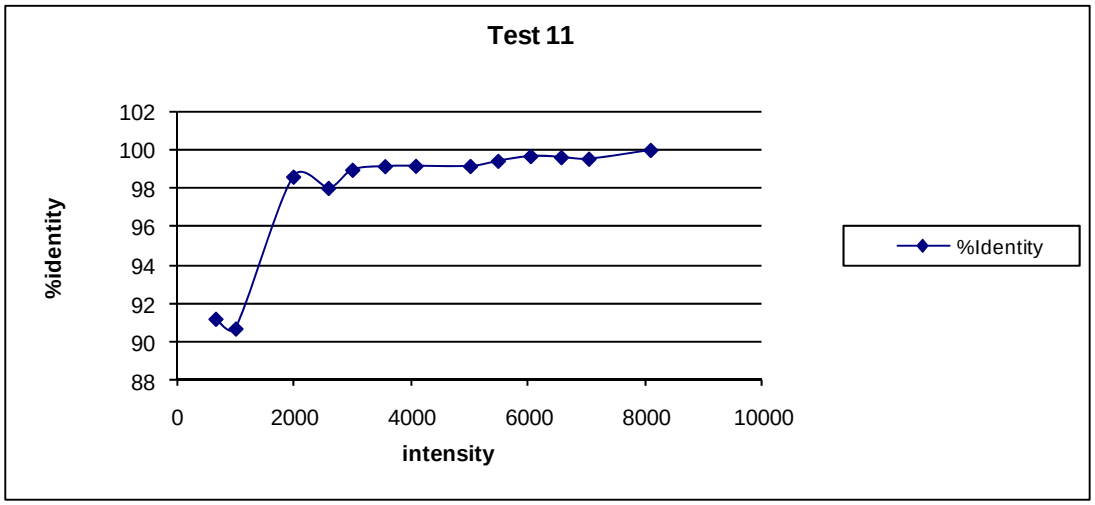
Signature consensus of H5 sample	Length (bp)	Nucleotide sequence
Sample no. 11	138	ACAATAATGGAAAAGAACGTTACTGAGTGAAGCCTCTAATTTTGAG AGATTGTAGTGTAGCCTATTGAGCAGAATAAACCATTTTGAGAGAA ATGTGGTATGGCTTATCAAACAGTACATACCCAACAATAAAGAGGA
Sample no. 17	479	AGAGCAGGTTGACACAATAATGGAAAAGAACGTTACTGTTACACAT GCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATC TAGATGGAGTGAAGCCTCTAATTTTGAGAGATTGTAGTGTAGCTGGA TGGCTTCTCGGAACCCAATGTGTGACGAATTCATCAATGTGCCGGAA TGGTCTTACATAGAGTCAATGACCTCTGTTACCCAGGAGAACTGAAA CACCTATTGAGCAGAACAAACCATTTTGAGAAAATTCAGATCATCCC CAAAAGTTCTTGGTCCAGTCATGAAGCCTCAAAGCTCAGCATGTCCA TACAAAAGTCCTCCTTTTTCAGAAATGTGGTATGGCTTATCAAAAAG AACATACCCAACAAACAATAAAGAGGACCTACAATAATACCAGGGA GGAAGAACTTCTGGTTCTCATGGAAATTCAAGAACCTTTACGACAAG GCCGACTAAGCC
Sample no. 25	308	AGCAGGTTGACACAATAATGGAAAAGAACGTTACTGTTACACATGC CCAAGACATACTGGAAAAGACACACAACGGGAAATGGAGTGAAGC CTCTAATTTTGAGAGATTGTAGTGTAGCTGGATGGCCCAATGTGTGA CGAATTCATCAATGTGCCGGAATGGTCAATTGAAACACCTATTGAGC AGAATAAACCATTTTGAGAAAATTCAGATCATCCCCAAAAGTTCTTG GTCCAGTCAAAGTCCTCCTTTTTCAGAAATGTGGTATGGCTTATCAA ACAGTACATACCCAACAATAAAGAGGAGC
Sample no. 47	305	AGAGCAGGTTGACACAATAATGGAAAAGAACGTTACTGCCCAAGAC ATACTGGAAAAGACACACAACGGATGGAGTGAAGCCTCTAATTTTG AGAGATTGTAGTGTAGCTGGATGGCCCAATGTGTGACGAATTCATCA ATGTGCCGGAATGGAATTGAAACACCTATTGAGCAGAATAAACCAT TTTGAGAAAATTCAGATCATCCCCAAAAGTTCTTGCTCCTTTTTCAG AAATGTGGTATGGCTTATCAAACAGTACATACCCAACAATAAAGAG GAGCTACAATGTCAAGAACCTTTACGA
Sample no. 48	336	AGAGCAGGTTGACACAATAATGGAAAAGAACGTTACTGTTACACAT GCCCAAGACATACTGGAAAAGACACACAACGGATGGAGTGAAGCCT CTAATTTTGAGAGATTGTAGTGTAGCTGGATGGCCCAATGTGTGACG AATTCATCAATGTGCCGGAATGGTCAATTGAAACACCTATTGAGCAG AATAAACCATTTTGAGAAAATTCAGATCATCCCCAAAAGTTCTTGGT CCAGTCAAAAGTCCTCCTTTTTCAGAAATGTGGTATGGCTTATCAAA AAGAACAGTACATACCCAACAATAAAGAGGACCTACAATGTCAAGA ACCTTTACGA

Table 8 Summary of signature N1 consensus Nucleotide sequence of each sample test.

Signature consensus of N1 sample	Length (bp)	Nucleotide sequence
Sample no. 11	503	ATAACCATTGGATCAATCTGTATGGTAATTGGAATAGTTAGCTTAATGTTACAAA TTGGGAACATTATCTCAATATGGGTTAGTCATTCAATCATTGTAACATTAGCGGG CAATTCATCTCTTTGTCCTATTAGTGGATGGGCTATATACAGTAAGGACAACRGT ATAACAAGGGGGATGTGTTTGTTATAAGAGAGCCATTCATCTCATGCTCCCACTT GGAATGCAGAACCTTCCGTCAAAGACAGAAGCCCTTATAGCTCCCCATATAACTC AAGGTTTGAGTCTGTTGCTTGGTCAGCAGTTGGTTGACAATTGGAATTTCTGGCTG GCTGTATTGAAATACAATGGCATAATAACAGACACAGTCTGAATGTGCATGTGTA AATGGCAGATCTTCAAAATGGAAAAAGGGAAAGTGGTTAAATCCTGTTATCCTGA TGCCGGTGGGTATCTTTCAATCAAAATTTGGAGTATCAAATAGGGTGGAGTTTTC GGAGACA
Sample no. 17	311	ATAACCATCGGATCAATCTGTATGGTAATTGGAATGGTTAGCTTAATGTTACAAA TTGGGAACATCTCAATATGGGTTAGTCATTCAATTCAATTAGCGGGCAATTCATCT CTTTGTCCCTATATACAGTAAGGACAACAGTATAACAAGGGGGATGTGTTTGTTA TAAGAGAGCCATTCATCTCATGCTCCCACTTGGAATGCAGAACCCCCATATAACT CAAGGTTTGAGTCTGTTGCTTGGTCAGCAGTTGGTTGACAATTGGAATTTCTGGCT GGCTGTATTGAAATACAATGGCATAATAACAGAC
Sample no. 25	254	ATCTCTTTGTCTTAGGGAGGCTTACAGAAGGACAAGTATAACAAGGGGATTGTTA AAGAGAGCATTCATCTCAGCACTTGAATGAGTTGACCTCAGACAGAAGCCCCTCC CCATATAACTCAAGGTTTGAGTCTGTTGCTTGGTCAGCAGTTGGTTGACAATTGG AATTTCTGGCTGGCTGTATTGAAATACAATGGCATAATAACAGACACTATCAAGA GTTGGAGGAATAACATACAGTCTGAATGTGCATG
Sample no. 47	312	ATAACCATCGGATCAATCTGTATGGTAATTGGAATAGTTAGCTTAATGTTACAAA TTGGGAACATCTCAATATGGGTTAGTCATTCAATTCAATTAGCGGGCAATTCATCT CTTTGTCCCATTAGTATATACAGTAAGGACAACAGTATAAAAGGGGATGTGTTTG TTATAAGAGAGCCATTCATCTCATGCTCCCACTTGGAATGCAGAACCCTCCCCAT ATAACTCAAGGTTTGAGTCTGTTGCTTGGTCAGCAGTTGGTTGACAATTGGAATTT CTGGCGTATTGAAATACAATGGCATAATAACAGAC
Sample no. 48	667	TAACCATTGGATCAATCTGTATGGTAATTGGAATAGTTAGCTTAATGTTACAAATT GGGAACATAATCTCAATATGGGTTAGTCATTCAATTCAGTAACATTAGCGGGCAA TTCATCTCTTTGTCCTATTAGTGGATGGGCTATATACAGTAAGGACAACAGTATA ACAAGGGGGATGTGTTTGTTATAAGAGAGCCATTCATCTCATGCTCCCACTTGGA ATGCAGAACCTTCGACCGTCAAAGACAGAAGCCCTTATAGCTCCCCATATAACTC AAGGTTTGAGTCTGTTGCTTGGTCAGCAGTTGGTTGACAATTGGAATTTCTGGCTG GCTGTATTGAAATACAATGGCATAATAACAGACACTATCAAGAGTTGGAGGAAT AACATACAGTCTGAATGTGCATGTGTAAATGGCTCTTGCTTTACTGTAATGACTG ACGGACCAAGTAATGGAGATCTTCAAAATGGAAAAAGGGAAAGTGGTTAAATCA GTCGAATTGGATGCTCCTAATTATCACTATGAGGAATGCTCCTGTTATCCTGATGC CGGGATAATTGGCATGGCTCGAATCGGCCATGGGTATCTTTCAATCAAAATTTGG AGTATCAAATAGGGTGGAGTTTTCGGAGACAATCCACGCCCCAATGATGGAACC AAAAGC

5.4 The high resolution

We further validated the high resolution by testing 5 HA sequences and 5 NA sequences of our Influenza A/H5N1 samples**.** After proof-of-concept of medium resolution we used the oligonucleotide sequences from probe set to search the identity for each sample. The results of Identity percent in 5 sequences of H5 gene and 5 sequences of N1 gene were shown in Figure 22-26.



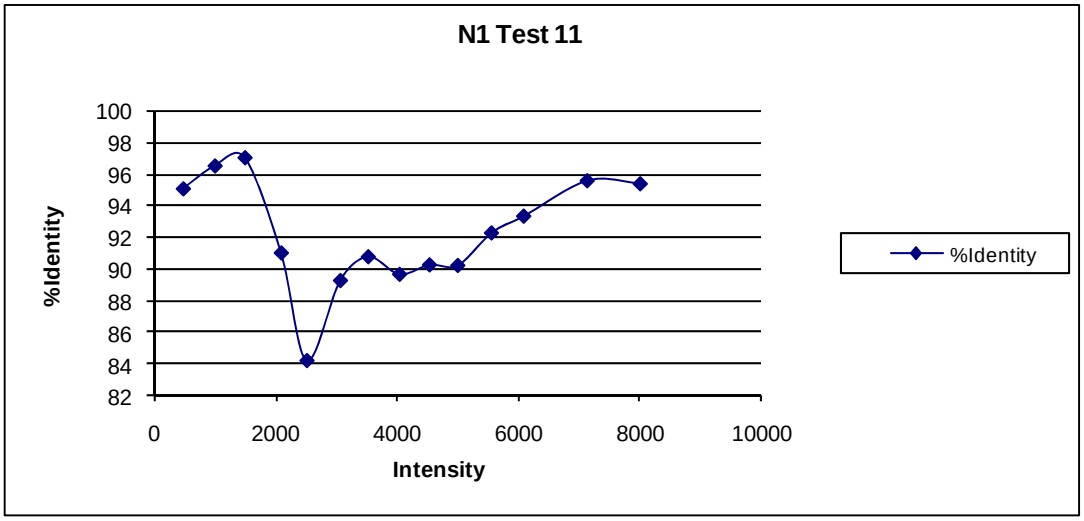
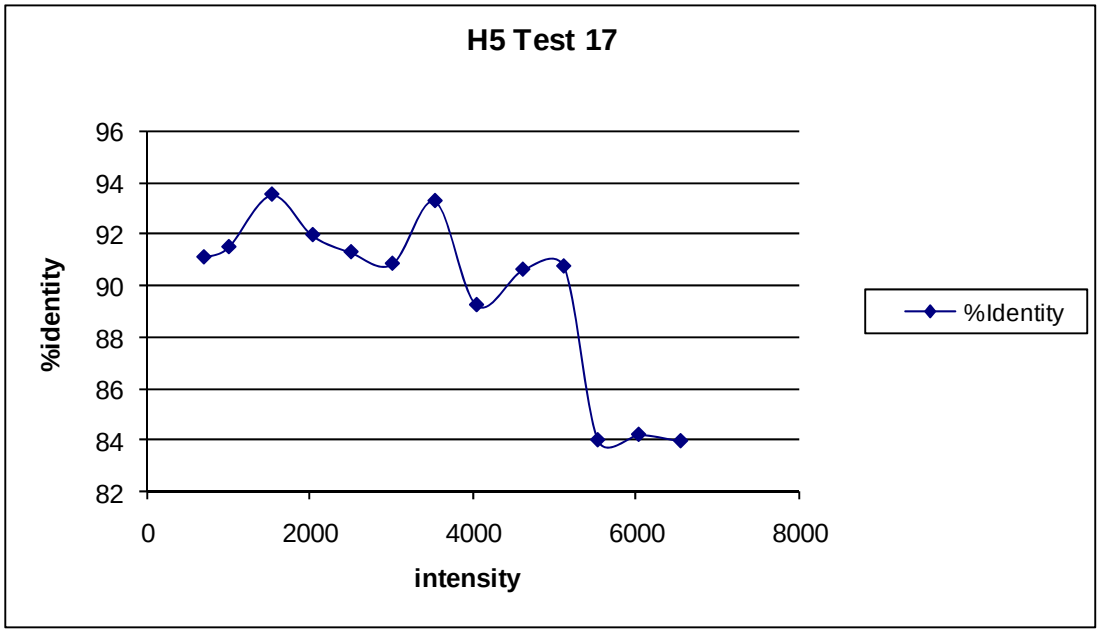


Figure 22 Identity percent sample no.11 for H5 and N1 sequences



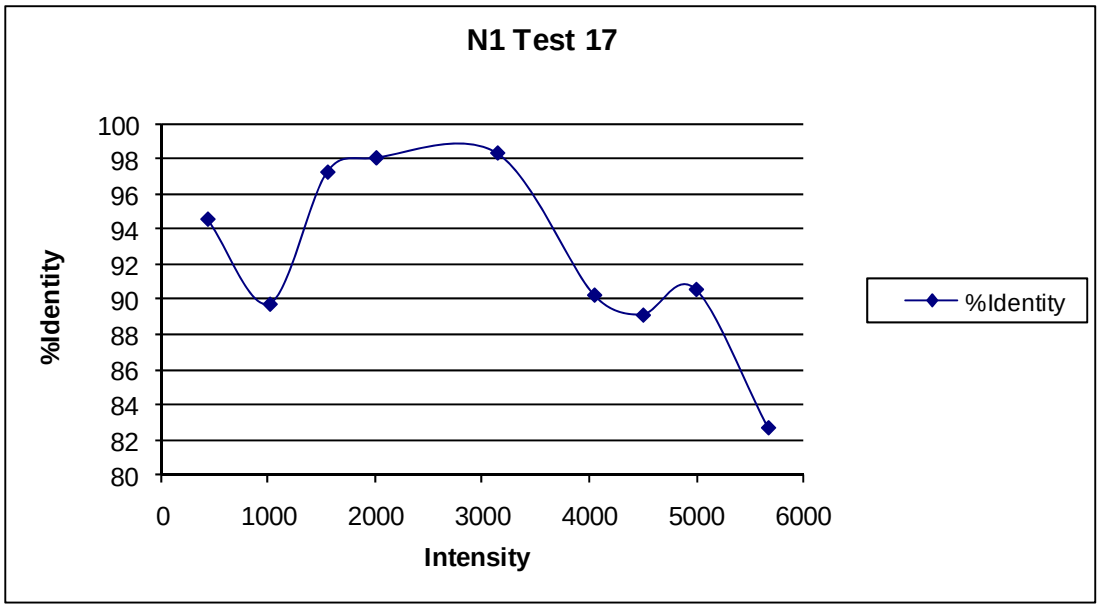
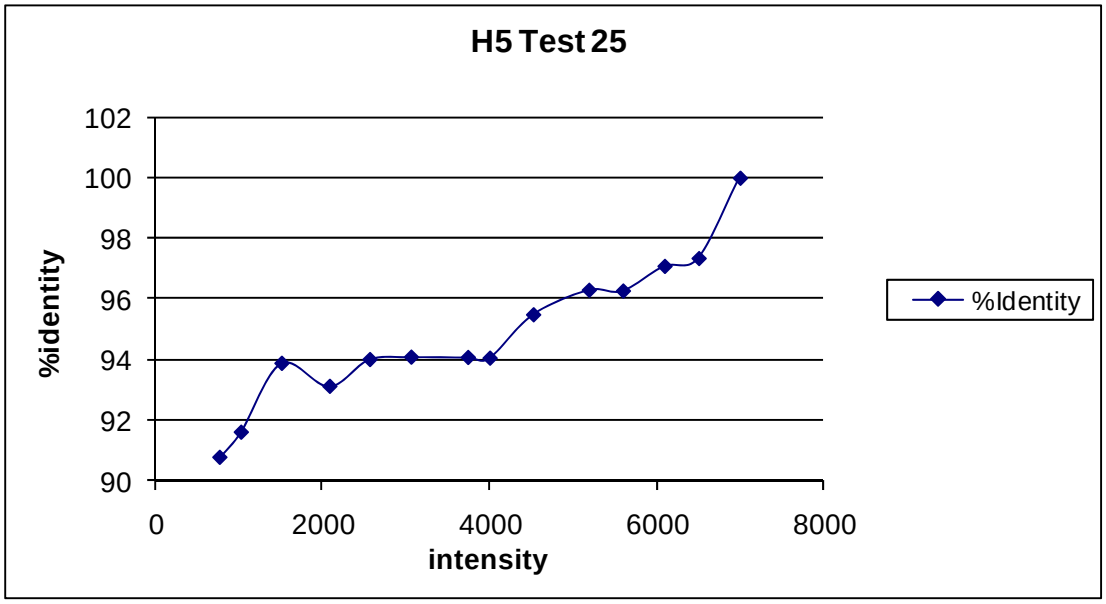


Figure 23 Identity percent sample no.17 for H5 and N1 sequences



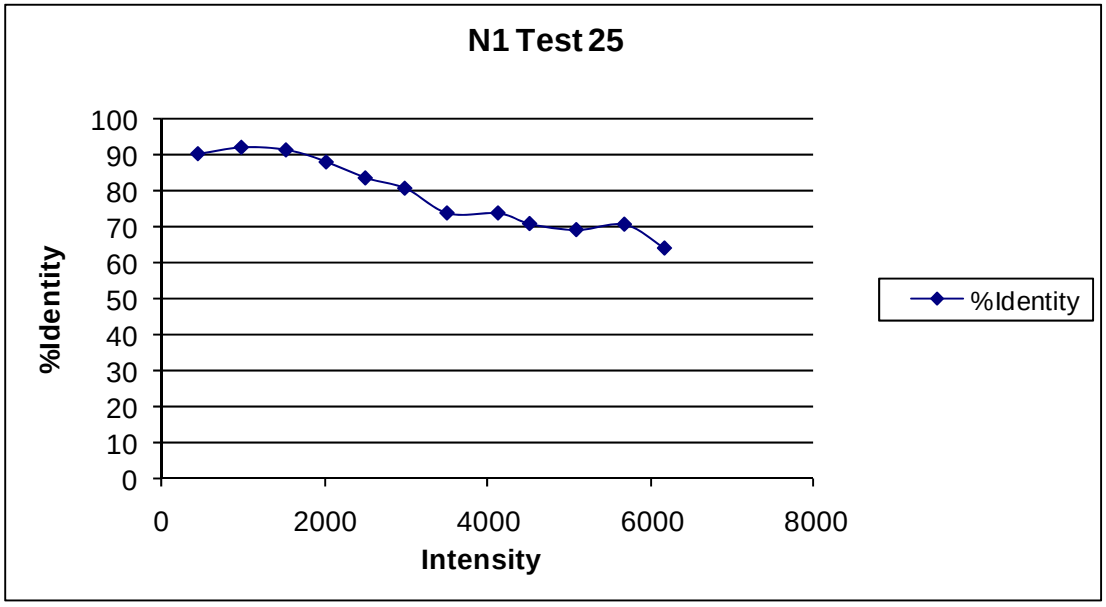
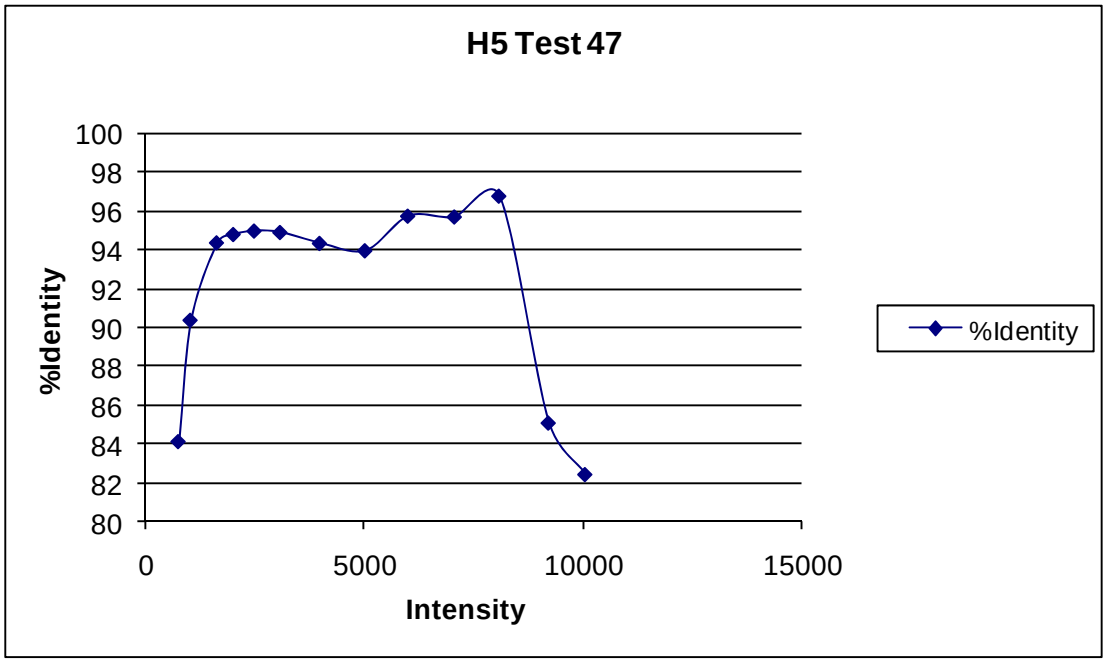


Figure 24 Identity percent sample no.125 for H5 and N1 sequences



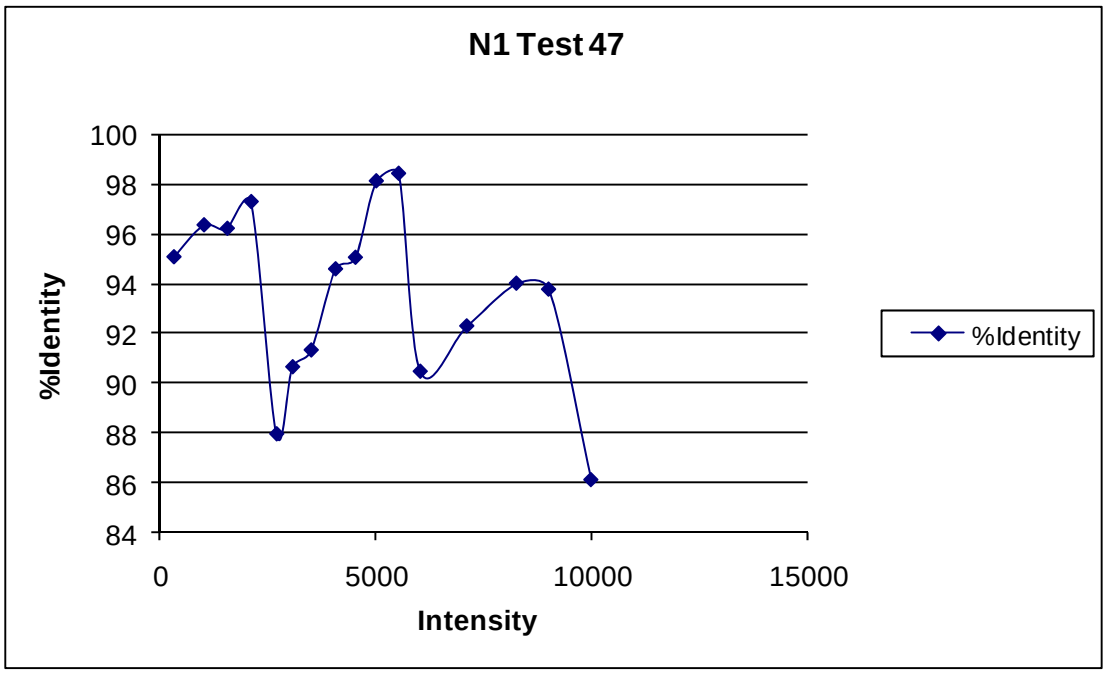
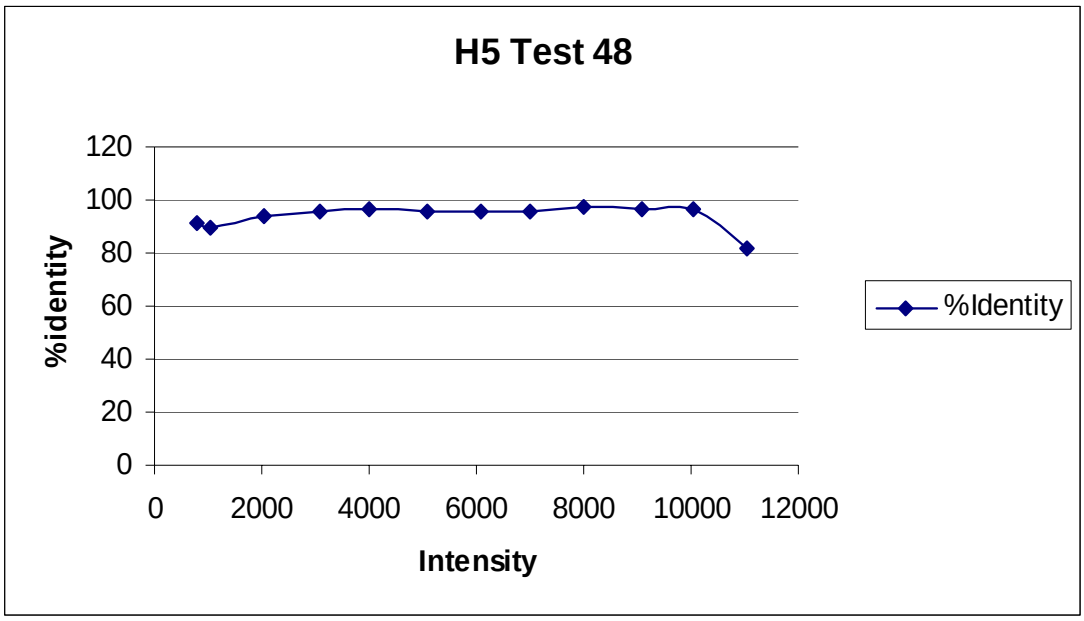


Figure 25 Identity percent sample no.47 for H5 and N1 sequences



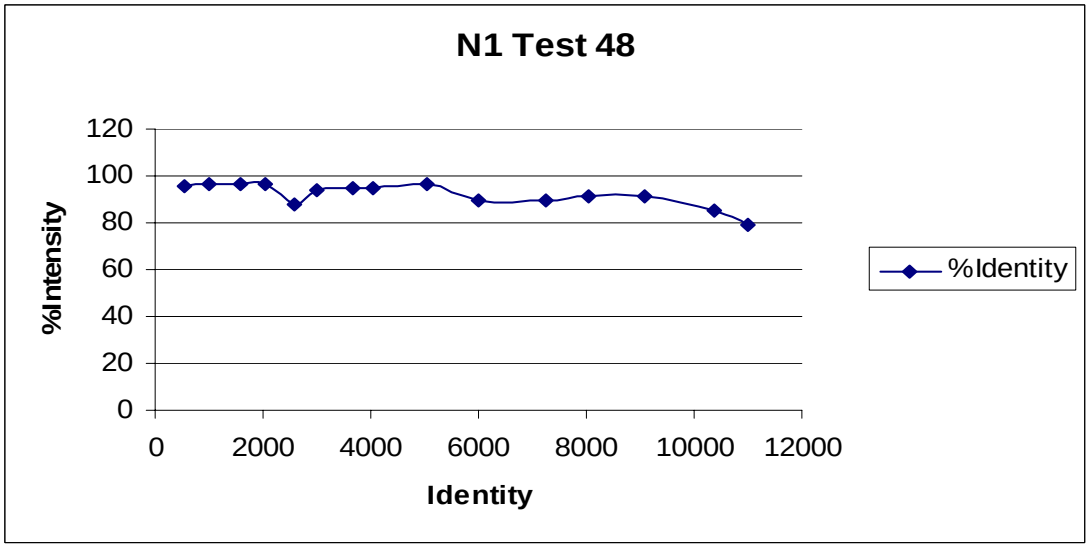


Figure 26 Identity percent sample no.48 for H5 and N1 sequences

5.5 Combine medium and high resolution result for monitoring avian influenza A H5N1

After proof-of-concept experiments, we combine medium and high resolution results to discriminate H5N1 strain. It is important not only to distinguish subtypes but also to identify the differences associated with significant shifts of the subtype from year to year by medium resolution. Not all of isolate corresponded to identification made based on the sequence obtained using the conventional DNA sequencing. For every isolate identified as influenza A virus H5 and N1 whose sequence was obtained using conventional DNA sequencing methods, the two methods identified strains that corresponded with each other except N1 in sample no.11 and no.48 in part of unEdit Sequence. After we create new database for H5 and N1 sequences, we can found all representative sequences were match against sequences in database (Table 9).

Sequence-based medium and high resolution analysis revealed the identity of the virus to be most nearly identical to each other that had been circulating during 2003-2006 in Thailand.

Table 9 Comparison of Local Blast result between uneditted and editted H5 signature consensus sequence combine medium and high resolution

Sample No.	% Identity	Local Blast Result					
		UnEdit Sequence			Edit Sequence		
		order	Bit	E-value	order	Bit	E-value
H5 No.11	100	2nd	74	7.00E-14	2nd	74	3.00E-14
H5 No.17	93.52818	1st	373	E-103	1st	373	E-103
H5 No.25	100	2nd	157	2.00-38	2nd	157	7.00E-39
H5 No.47	97.04918	1st	159	4.00-39	1st	135	3.00E-32
H5 No.48	97.02381	2nd	155	7.00E-38	2nd	155	3.00E-38
N1 No.11	97.01789	Not found	-	-	8th	177	8.00E-45
N1 No.17	98.39744	1st	145	1.00E-35	11th	125	2.00E-29
N1 No.25	92.2449	1st	137	5.00E-33	1st	125	2.00E-29
N1 No.47	98.39744	Not found	-	-	18th	117	5.00E-27
N1 No.48	96.875	1st	-	-	4th	188	3.00E-48

CHAPTER VI
DISCUSSION

Mutations in the influenza virus, including avian and human influenza, can occur either gradually or quickly. Either type of mutation can give rise to new strains of the virus. The mutations often alter the characteristics of the viruses. These new strains may increase or decrease the severity of illness. The immune systems of individuals infected by earlier strains of the virus do not necessarily recognize the new strains (22, 96). From a diagnostics standpoint, influenza is often described as a "moving target" due to the high mutation rate of HA and NA. The HA and NA genes evolve rapidly at 6.7×10^{-3} and 3.2×10^{-3} nucleotide substitutions per nucleotide per year, respectively, and are genetically diverse between subtypes (5, 67, 97). The effects are impossible to predict. For influenza virus strains, it is important not only to distinguish subtypes but also to identify the differences associated with significant shifts of the subtype from year to year. Subtype determination ("subtyping") is essential for tracking emerging viruses and for designing appropriate influenza vaccines (53, 98). The current "gold standard" for complete influenza A subtyping (determination of both HA and NA) involves virus replication in egg or tissue culture followed by a hemagglutinin inhibition assay. This method is tedious and requires several days, with the analysis time often extended to several weeks for antigenically novel viruses. RT-PCR techniques depend on specific primers, which may fail when corresponding viral sequences mutate (74, 81, 82, 99, 100). We are using a semiconductor-based oligonucleotide array technology that can be used with fluorescent labels and traditional optical scanning devices or used as a biodetector using electrochemical techniques for analysis. This platform is extremely flexible, which allows array designs to be rapidly and easily modified and synthesized, thus permitting oligonucleotides of interest to be tested empirically. In addition, the ability to use electrochemical detection with semiconductor microchips eliminates the need

for expensive optical scanning equipment. In contrast to the fluorescent detection, the ElectraSense^{TM} format has fewer limitations for further miniaturization of geometrical sizes of individual elements on the microarray. Clearly, the cost per assay and the associated costs of equipment needed to run and analyze the assay are critical for diagnostic applications such as these (12, 17, 20).

However, the influenza virus subtyping array is also compatible with traditional labeling, hybridization, and washing protocols that can be completed within 1.0 to 1.5 h. However, use of this method allows the array to be “stripped” and reused multiple times since there is no covalent coupling of the label to the array. In addition, this array can also benefit from the sectoring approach mentioned above to further bring costs to minimum. This platform is a viable alternative to RT-PCR because of the combination of assay speed, array sectoring, which would allow multiple assays on one array, the potential to strip and reuse the adaptability to inexpensive electrochemical scanning devices. The target sample preparation system used for standard RT-PCR-based method, except that it uses a very redundant consensus priming system that maximizes the chance that novel strains of influenza virus will be amplified and thus minimizes false-negative results. The chip contains multiple probes that correspond to key distinguishing elements of each HA or NA subtype. It is laid out in a visual pattern so that it can be read visually for quick identification as well as analyzed with more advanced algorithms. The system can also identify rare versus more commonly seen genetic variants based on the organization of subtype-specific probes (i.e., the probes are arranged in order from more universal to more specific for each subtype).

In many cases, the detection and identification of influenza A subtype is not sufficient for making decisions on vaccine development, patient treatment, and general surveillance, and specific sequence is needed. For example, mutations in the sequence coding the haemagglutinin receptor binding site of subtype H5N1 have been shown to be important for replication of avian viruses in humans (31, 51, 101). In this study, viral HA and NA subtype sequencing was accomplished with a bioinformatic tool. After scanning the array, intensity data were exported to an Excel worksheet. Sequence information was extracted with a routine design to associate the correct base

with the highest signal from sets of probes that were tiled by one nucleotide to cover the sequence of interest. Although microarrays are excellent for determining subtypes, they do not give detailed genetic information or information when antigenic drift occurs (102-105). The identification of HA and NA subtypes followed by sequencing with bioinformatic tools will significantly decrease the time and cost for the genetic monitoring of potential lethal virus strains (91, 102, 104, 106-109) . This study demonstrates that the detection, identification and monitoring of viral genomes in samples by using a bioinformatic tool in combination with electrochemical detection is a viable rapid approach that can complement traditional method. Two level resolutions were used to make detection and identification. Finding the signature oligonucleotide used the exclusivity of BLASTn hit to determine specificity (medium resolution). It is unambiguous but ignores the information in oligonucleotides that react with multiple characteristics, for many HA and NA subtypes there were no uniquely diagnostic probes. To compensate for the lack of signature oligonucleotides, an attempt was made to use the overrepresentation of some categories among BLASTn hits of the oligonucleotides and the underrepresentation of others to aid in the characterization.

The protocol we have developed successfully in medium resolution which provided pathogen identification to maximum level of detail possible (species or strain) depending on the quality of each signature sequence and local database. This identification capability requires minimal input on the identity of the pathogens, making non-expert use feasible. The crucial feature incorporated that allowed complete automation was the use of taxonomic databases, which classify organisms into ordered groups and provide relationships between organism entries, allowing removal of redundancies, comparison of different related signature consensus and simplification of data presentation. This allows databases such as NCBI or IVDB, which are redundant and subject to minimal curation but which constantly receive update and new sequence information, to be used with great success. For more variable or rapidly mutating H5N1, low resolution still provided accurate detailed identification, but medium and high resolution were unable to report fine scale discrimination. The comparison of the conventionally sequenced influenza A/H5N1 gene sequences illustrated that the algorithm is capable of automatically adjusting for

updates in databases. The protocol demonstrated the local database's capability to properly distinguish hybridization caused by the specified influenza A/H5N1 from those caused by genetically close (near neighbor) strains and to not make incorrect identifications, eliminating one potential cause of false positive. Improving on the level of detail reported in medium and high resolution will require more information about an individual pathogen and may have to be developed for each specific pathogen or class of pathogens. This information is also required for the protocol to identify which differences between a sample and database entries represent significant mutations.

Microarray designed to analysis of influenza virus strains have previously been study by Li et al. (16), Kessler et al (110), and Sengupta et al. (15). Wang et al. (14), incorporated influenza virus-specific probes in a microarray designed for analysis of a wide range of viruses but did not test specifically for influenza virus. On the other hand, our study, we have non-influenza virus species probes in a microarray to test specifically for influenza virus. For medium and high resolution, we need more number of samples to test and validate our algorithm. Just as we pointed to limitations of previous task, it is important to point out limitations of this study. One limitation is that, sequences were initially selected primarily for their biological interest, but then some of these candidates were excluded. This introduces a possible bias for local database into our medium resolution results. Sequences with every average intensity signal set were not excluded. These factors resulted in the selection of primarily low to high signal intensity. In medium resolution has suggested that difficulties with methods that subtract sequences arise for low intensity signal due to extreme variability. The remaining criteria used to select signature oligonucletide were the agreement of kappa analysis. We initially indicate cut off of intensity signal of H5 and N1 oligonucleotide probe. Kappa analysis result produced the best agreement between H5N1 oligonucleotide and non-H5N1 oligonucleotides. And other criteria were existing knowledge of an appropriate bioinformatics tools and update database influenza A virus library. We are unaware of any biases produced by these latter selection criteria.

However, high resolution task shows the “signature sequences” really not depend on high signal intensity. For example, Figure 22-26 (except H5 Test 11 and Test 25) show that this is exactly what happens. For set of oligonucleotides at low level of intensity, high resolutions that use identity percent have been found to extremely variable. They give high identity percent but local blast results were poor. We also found poor reliability for high resolution using set of high signal intensity oligonucleotide set. Because the signature sequence in each sample by using high resolution are very long which produced high mismatch tolerance (15, 89).

These studies are useful and important, but are not the end of the story. We need more number of influenza A strain to use for validation. A criterion often not considered in the spike-in studies is the accuracy of measurements across genes. In addition to the extremely high degree of accuracy of the method microarray is to detect RNA rather than the intermediate cDNA, absence of amplification and a more sensitivity (3). Our result suggest a more complicated scenario that each of these low-medium-high resolution perform well if combined with particular other appropriate statistic analysis (111, 112), proper bioinformatics tools (93-95, 113-117)and information of oligonucleotide probes in microarray (90, 118, 119). Future work will involve improving the use of the current taxonomic database or potentially developing a new relational database that is specific to our needs and then incorporating more specific information of target pathogens. The hierarchal design of the data analysis makes it easy to incorporate analysis that build upon the analysis already performed (120).

CHAPTER VII
CONCLUSION

In conclusion, we report exploratory analyses of signal intensity oligonucleotide array data from the semiconductor base oligonucleotide microarray with the objective of detection, identification and monitoring of the currently used measures of hemagglutinin and neuraminidase genes of H5N1 avian influenza circulating in Thailand during 2003-2006. Our analyses make use of three levels of typing (i) low resolution (detection or genotyping), experimental study consisting of single step reverse transcription, hybridization and imaging microarray, (ii) medium resolution and (iii) high resolution.

Finally, we evaluated the algorithms in term of their ability to detect known DNA sequences using the signal intensities data and bioinformatic (Bioedit program) analysis of the five hemagglutinin segments and five neuraminidase segments of five RNA H5N1 avian influenza viruses. The experiment results showed that detection, using almost all scoring algorithms achieved good results, except the use of high resolution which was produces less accurate than the others.

We have two areas of concern to date. We found that for medium and high resolution typing, it was time consuming to create the HA and NA database. Also, the specific gene sequence of each H5N1 strain is subjective, so the protocol monitoring of the new strain H5N1 is recommended for further study.

REFERENCES

1. Lee C-W, Suarez DL. Application of real-time RT-PCR for the quantitation and competitive replication study of H5 and H7 subtype avian influenza virus. Virological Methods. 2004;119:151-58.
2. Ong MDd, Hien TT. Avian influenza A (H5N1). J ClinVirol. 2006;35:2–13
3. Townsend MB, Dawson ED, Mehlmann M, Smagala JA, Dankbar DM, Moore CL, et al. Experimental evaluation of the FluChip diagnostic microarray for influenza virus surveillance. J Clin Microbiol. 2006 Aug; 44(8): 2863-71.
4. Ferguson NM, Cummings DaT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing and emerging influenza pandemic in southeast Asia. Nature. 2005;437:209-14.
5. Dawson ED, Moore CL, Smagala JA, Dankbar DM, Mehlmann M, Townsend MB, et al. MChip: a tool for influenza surveillance. Anal Chem. 2006 Nov 15; 78(22): 7610-5.
6. Kanchanachitra C, Podhisita C. Archavanitkul K, Im-em W, Institute for Popution and Social Research, Thai Health Promotion Foundation, Facing the Challenge of Bird Flu. Thailand Health Promotion Foundation 2006: 88-111.
7. Fouchier RA, Bestebroer TM, Herfst S, Van Der Kemp L, Rimmelzwaan GF, Osterhaus AD. Detection of influenza A viruses from different species by PCR amplification of conserved sequences in the matrix gene. J Clin Microbiol. 2000 Nov; 38(11): 4096-101.
8. Monto AS, McKimm-Breschkin JL, Macken C, Hampson AW, Hay A, Klimov A, et al. Detection of Influenza Viruses Resistant to Neuraminidase Inhibitors in Global Surveillance during the First 3 Years of Their Use. Antimicrob Agents Chemother. 2006 July 1, 2006;50(7):2395-402.

9. Payungporn S, Chutinimitkul S, Chaisingh A, Damrongwantanapokin S, Nuansrichay B, Pinyochon W, et al. Discrimination between Highly Pathogenic and Low Pathogenic H5 Avian Influenza A Viruses. Emerging Infectious Diseases 2006;12(4):700-01.

10. Stevens J, Blixt O, Paulson JC, Wilson IA. Glycan microarray technologies: tools to survey host specificity of influenza viruses. Microbiol Rev. 2006;4:875-64.

11. Dawson ED, Moore CL, Dankbar DM, Mehlmann M, Townsend MB, Smagala JA, et al. Identification of A/H5N1 influenza viruses using a single gene diagnostic microarray. Anal Chem. 2007 Jan 1; 79(1): 378-84.

12. Lodes MJ, Suciu D, Elliott M, Stover AG, Ross M, Caraballo M, et al. Use of Semiconductor-Based Oligonucleotide Microarrays for Influenza A Virus Subtype Identification and Sequencing. Clin Microbiol. 2006; 44(4): 1209–18.

13. Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, et al. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature. 2005 Oct 20;437(7062):1162-6.

14. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, et al. Microarray-based detection and genotyping of viral pathogens. Proc Natl Acad Sci U S A. 2002 Nov 26; 99(24): 15687-92.

15. Sengupta S, Onodera K, Lai A, Melcher U. Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization. J Clin Microbiol. 2003 Oct; 41(10): 4542-50.

16. Li J, Chen S, Evans DH. Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. J Clin Microbiol. 2001 Feb; 39(2): 696-704.

17. Liu RH, Michael J. Lodes, Tai Nguyen, Tony Siuda, Mike Slota, H. Sho Fuji, et al. Validation of A Fully Integrated Microfluidic Array Device for Influenza A Subtype Identification and Sequencing. Anal Chem. 2006; 78(12): 4184-93.

18. Majury A, Ash J, Toye B, Evans G, Mazzulli T, Winter A-L, et al. Laboratory diagnosis of human infection with avian influenza. CMAJ 2006;175(11):1371-72.

19. Kilbourne ED. WHO Manual on Animal Influenza Diagnosis and Surveillance. 2006.

20. Ghindilis AL, Smith MW, Schwarzkopf KR, Roth KM, Reyvan K, Munro DB, et al. CombiMatrix oligonucleotide arrays: Genotyping and gene expression assays employing electrochemical detection. Biosensor and Bioelectronics. 2006:1853-60.

21. Keawcharoen J, Oraveerakul K, Kuiken T, Fouchier RA, Amonsin A, Payungporn S, et al. Avian influenza H5N1 in tigers and leopards. Emerg Infect Dis. 2004 Dec; 10(12): 2189-91.

22. Facing the Challenge of Bird Flu. Thailand Health Promotion Foundation 2006:88-111.

23. Steinhauer DA, Skehel JJ. Genetics of Influenza Viruses. Annu Rev Genet 2002; 36: 305–32.

24. Horimoto T, Kawaoka Y. INFLUENZA:lessons from past pandemics, warnings from current incidents Microbiol Rev. 2005; 3: 591-600.

25. Ng EK, Cheng PK, Ng AY, Hoang TL, Lim WW. Influenza A H5N1 detection. Emerg Infect Dis. 2005 Aug; 11(8): 1303-5

26. Kamps BS, Hoffmann C, Preiser W, Werner O, Harder TC. Influenza Report 2006.

27. Gambaryan A, Tuzikov A, Pazynina G, Bovin N, Balish A, A K. Evolution of the receptor binding phenotype of influenza A (H5) viruses. J Virol. 2006; 344: 432-8.

28. Cross KJ, Burleigh LM, Steinhauer DA. Mechanisms of cell entry by influenza virus. Expert Rev Mol Med. 2001 Aug 6; 2001: 1-18.

29. Feldmann A, Schafer MK, Garten W, HD K. Targeted infection of endothelial cells by avian influenza virus A/FPV/Rostock/34 (H7N1) in chicken embryos. J Virol. 2000; 74: 8018-27.

30. Yamada S, Suzuki Y, Suzuki T, Le MQ, Nidom CA, Sakai-Tagawa Y, et al. Haemagglutinin mutations responsible for the binding of H5N1

influenza A viruses to human-type receptors. Nature. [Letters]. 2006; 444.

31. Kawaoka Y, Krauss S, Webster RG. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. J Virol. 1989; 63: 4603–8.
32. Sevill N, Domingo E, de la Torre J.C. Contribution of LCMV towards deciphering biology of quasispecies in vivo. Curr Top Microbiol Immunol. 2002; 263: 197–220.
33. Lai MM. Genetic recombination in RNA viruses. Curr Top Microbiol Immunol. 1992; 176: 21-32.
34. Nagy P.D., Simon A.E. New insights into the mechanisms of RNA recombination. J Virol. 1997; 235: 1-9.
35. Suarez DL, Senne DA, Banks J, Brown IH, Essen SC, Lee C-W, et al. Recombination resulting in virulence shift in avian influenza outbreak, Chile. Emerg Infect Dis. 2004;10:693–99.
36. Hoffmann E, Stech J, Leneva I, Krauss S, Scholtissek C, Chin P.S, et al. Characterization of the Influenza A Virus Gene Pool in Avian Species in Southern China: Was H6N1 a Derivative or a Precursor of H5N1? J Virol. 2000; 74(14): 6309–15.
37. Webby R, Hoffmann E, Webster R. Molecular constraints to interspecies transmission of viral pathogens. Nature Medicine 2004; 10(12): 77-81.
38. Wang Z, Daum LT, Vora GJ, Metzgar D, Walter EA, Canas LC, et al. Identifying influenza viruses with resequencing microarrays. Emerg Infect Dis. 2006 Apr; 12(4): 638-46.
39. Rajagopal S, Treanor J. Pandemic (Avian) Influenza. Seminars in respiratory and critical care medicine. 2007;28(2):159-68.
40. Ann H. Reid, Jeffery K. Taubenberger, Thomas G. Fanning., Initial genetic characterization of the 1918 "Spanish" influenza virus. Science. 1997; 275: 1793-6.

41. Ann H. Reid, Jeffery K. Taubenberger, Thomas G. Fanning. Evidence of and absence: the genetic origins of the 1918 pandemic influenza virus. Nat Rev Microbiol. 2004; 2: 909-14.

42. Network TWHOGIPS. Evolution of H5N1 Avian Influenza Viruses in Asia. Emerg Infect Dis 2005; 11(10): 1515-21.

43. Capua H, Alexander DJ. Avian influenza and human health. Acta Tropica. 2002; 83: 1-6.

44. Webster, R. G., Bean, W. J., Gorman, O. T., et al. Evolution and ecology of influenza A viruses. Microbiol Rev. 1992; 56: 152–79.

45. Hatta M., P. Gao, P. Halfmann , Kawaoka Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science. 2001; 293: 1840–2.

46. Rott R. The pathogenic determinant of influenza virus. Vet Microbiol. 1992; 33: 303–10.

47. Skehel, J. J., Wiley DC. Receptor binding and membrane fusion invirus entry: the influenza hemagglutinin. Annu Rev Biochem. 2000; 69: 531–69.

48. Pasick J, Handel K, Robinson J, Copps J, Ridd D, Hills K, et al. Intersegmental recombination between the haemagglutinin and matrix genes was responsible for the emergence of a highly pathogenic H7N3 avian influenza virus in British Columbia. Gen Virol 2005; 86: 727–31.

49. Wood G.W., Mccauley J.W., Bashiruddin J.B., Alexander D.J. Deduced amino acid sequences at the haemagglutinin cleavage site of avian influenza A viruses of H5 and H7 subtypes. Arch Virol. 1993;130:209–17.

50. Senne D.A., Panigraphy B., Kawaoka Y. et al. Survey of the hemagglutinin (HA) cleavage site sequence of H5 and H7 avian influenza viruses: amino acid sequence at the HA cleavage site as a marker of pathogenicity potential. Avian Dis. 1996; 40: 425–37.

51. Maines TR, Lu XH, Erb SM, Edwards L, Guarner J, Greer PW, et al. Avian influenza (H5N1) viruses isolated from humans in Asia in 2004 exhibit increased virulence in mammals. J Virol. 2005 Sep; 79(18): 11788-800.

52. Stefan Riedel M, PhD. Crossing the species barrier: the threat of an avian influenza pandemic. Proc (Bayl Univ Med Cent) 2006; 19: 16–20.

53. Gillim-Ross L, Subbarao K. Emerging Respiratory Viruses: Challenges and Vaccine Strategies. Clin Microbiol 2006; 19(4): 614–36.

54. Matrosovich, M. N., T. Y. Matrosovich, T. Gray, N. A. Roberts a, D.Klenk H, et al. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proc Natl Acad Sci USA. 2004; 101: 4620–4.

55. Ito T, Couceiro JNSS, Kelm S, Baum LG, Kruss S, Castrucci MR, et al. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. J Virol. 1998; 72: 7367–73.

56. Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. . Science. 1998; 279: 393–96.

57. Kennedy F. Shortridge, Nan Nan Zhou, Yi Guan, Peng Gao, Toshihiro Ito, Yoshihiro Kawaoka, et al. Characterization of Avian H5N1 Influenza Viruses from Poultry in Hong Kong. J Virol 1998; 252: 331–42.

58. Chen H, Smith GJ, Zhang SY. Avian flu: H5N1 virus outbreak in migratory waterfowl. Nature. 2005; 436: 191–2.

59. Cumulative Number of Confirmed Human Cases of Avian Influenza A/H5N1 Reported to WHO [database on the Internet]. (Cited. Available from: http://www.who.int/csr/disease/avian_influenza/country/cases_table_2006_08_23/en/index.html.)

60. Zambon MC. The pathogenesis of influenza in humans. Rev Med Virol. 2001 Jul-Aug;11(4):227-41.

61. Tumpey TM, Garcı´a-Sastre A, Taubenberger JK, Palese P, Swayne DE, Basler CF. Pathogenicity and immunogenicity of influenza viruses with genes from the 1918 pandemic virus. PNAS. 2004, 101(9): 3166–71.

62. Chotpitayasunondh T, Ungchusak K, Hanshaoworakul W, Chunsuthiwat S, Sawanpanyalert P, Kijphati R, et al Human Disease from Influenza A (H5N1), Thailand, 2004. Emerg Infect Dis. 2005; 11(2): 201-9.

63. Tam JS. Influenza A (H5N1) in Hong Kong: an overview. Vaccine 2002;20: S77–S81.

64. Knossow M, Skehel JJ. Variation and infectivity neutralization in influenza. Immunol. 2006 Sep; 119(1): 1-7.

65. H. Chen, G. Deng, Z. Li, G. Tian, Y. Li, P. Jiao, et al. The evolution of H5N1 influenza viruses in ducks in southern China. Proc Natl Acad Sci USA. 2004; 101: 10452-7.

66. Kuiken T, Rimmelzwaan G, Riel Dv, Amerongen Gv, Baars M, Fouchier R, et al. Avian H5N1 Influenza in Cats. Sci. 2004 8 ctober;306:241.

67. Chen R, Holmes EC. Avian Influenza Virus Exhibits Rapid Evolutionary Dynamics 2006.

68. Amonsina A, Chutinimitkul S, Pariyothorn N, Songserm T, Damrongwan-tanapokin S, Puranaveja S, et al. Genetic characterization of influenza A viruses (H5N1) isolated from 3rd wave of Thailand AI outbreaks. Virus research. 2006; 122: 194-9.

69. Dennis Normile ea. Genomic Analysis Hints at H5N1 pathogenicity. Sci. 2006; 311.

70. Webster RG, Peiris M, Chen H, Guan Y. H5N1 Outbreaks and Enzootic Influenza. Emerg Infect Dis [Overview]. 2006; 12(1).

71. Eric C J Claas, Albert D M E Osterhaus, Ruud van Beek, Jan C De Jong, Guus F Rimmelzwaan, Dennis A Senne, et al. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. THE Lancet 1998; 351: 472-7.

72. Puthavathana P, Auewarakul P, Charoenying PC, Sangsiriwut K, Pooruk P, Boonnak K, et al. Molecular characterization of the complete genome of human influenza H5N1 virus isolates from Thailand. J Gen Virol. 2005 February 1, 2005; 86(2): 423-33.

73. Uiprasertkul M, Puthavathana P, Sangsiriwut K, Pooruk P, Srisook K, Peiris M, et al. Influenza A H5N1 replication sites in humans. Emerg Infect Dis. 2005 Jul; 11(7): 1036-41.

74. Hoffmann B, Harder T, Starick E, Depner K, Werner O, Beer M. Rapid and Highly Sensitive Pathotyping of Avian Influenza A H5N1 Virus by

Using Real-Time Reverse Transcription-PCR. J Clin Microbiol. 2007 Feb; 45(2): 600-3.

75. Y. Guan, L. L. M. Poon, C. Y. Cheung, T. M. Ellis, W. Lim§, A. S. Lipatov, et al. H5N1 influenza: A protean pandemic threat. PNAS. 2004; 101(21).

76. Kilbourne ED. WHO Manual on Animal Influenza Diagnosis and Surveillance. 2006.

77. Lee C-W, Suarez DL. Application of real-time RT-PCR for the quantitation and competitive replication study of H5 and H7 subtype avian influenza virus. Virological Methods. 2004;119:151-58.

78. Ng., E.K., Cheng PK, Ng AY, Hoang TL, Lim WW. Influenza A H5N1 detection. Emerg Infect Dis. 2005; 11: 1303-5.

79. Payungporn S, Chutinimitkul S, Chaisingh A, Damrongwantanapokin S, Buranathai C, Amonsin A, et al. Single step multiplex real-time RT-PCR for H5N1 influenza A virus detection. J Virol Methods. 2006 Feb; 131(2): 143-7.

80. Choi YK, Goyal SM, Kang SW, Farnham MW, Joo HS. Detection and subtyping of swine influenza H1N1, H1N2 and H3N2 viruses in clinical samples using two multiplex RT-PCR assays. J Virol Methods. 2002 Apr; 102(1-2): 53-9.

81. Chantratita W, Sukasem C, Kaewpongsri S, Srichunrusami C, Pairoj W, Thitithanyanont A, et al. Qualitative detection of avian influenza A (H5N1) viruses: A comparative evaluation of four real-time nucleic acid amplification methods. Molecular Cell Probe. 2008 (In press).

82. Whiley DM, Sloots TP. A 5'-nuclease real-time reverse transcriptase-polymerase chain reaction assay for the detection of a broad range of influenza A subtypes, including H5N1. Diagn Microbiol Infect Dis. 2005 Dec; 53(4): 335-7.

83. Mehlmann M, Bonner AB, Williams JV, Dankbar DM, Moore CL, Kuchta RD, et al. Comparison of the MChip to Viral Culture, Reverse Transcription-PCR and the QuickVue(R) Influenza A+B Test for Rapid Diagnosis of Influenza. J Clin Microbiol. 2007 Feb 14.

84. Poddar SK. Influenza virus types and subtypes detection by single step single tube multiplex reverse transcription-polymerase chain reaction (RT-PCR) and agarose gel electrophoresis. J Virol Methods. 2002 Jan; 99(1-2): 63-70.

85. Mckimmp Breschik JL, Blick TJ, Sahasrabudhe A, Tiong T, Marshall D, Hart GJ, et al. Generation and Characterization of Variants of NWS/G70C Influenza Virus after In Vitro Passage in 4-Amino-Neu5Ac2en and 4-Guanidino-Neu5Ac2en. Antimicrobial agents and chemotherapy. 1996; 40(1): 40-46.

86. Shinya K, Hatta M, Yamada S, Takada A, Watanabe S, Halfmann P, et al. Characterization of a Human H5N1 Influenza A Virus Isolated in 2003. J Virol. 2005;79(15):9926–32.

87. Qin L-X, Beyer RP, Hudson FN, Linford NJ, Morris DE, Kerr KF. Evaluation of methods for oligonucleotide array data via quantitative real-time PCR. BMC Bioinformatics 2006; 7: 23.

88. Stevens J, Ola Blixt LG, Taubenberger JK, Peter Palese, Paulson JC, Wilson IA. Glycan Microarray Analysis of the Hemagglutinins from Modern and Pandemic Influenza Viruses Reveals Different Receptor Specificities. J Mol Biol 2006; 355: 1143–55.

89. Cutler DJ, Zwick ME, Carrasquillo MM, Yohn CT, Tobin KP, Kashuk C, et al. High-throughput variation detection and genotyping using microarrays. Genome Res. 2001 Nov; 11(11): 1913-25.

90. Stekel D. Microarray Bioinformatic. 2003.

91. Bryant PA, Venter D, Robins-Browne R, Curtis N. Chips with everything: DNA microarrays in infectious diseases. Lancet Infect Dis. 2004 Feb; 4(2): 100-11.

92. Altschul SF, Gish W, Miller W, Myers EW, Lipmanl DJ. Basic local alignment search tool. Mol Biol Evol. 1990; 215: 403-10.

93. Guoqing Lu, Thaine Rowley, Garten R, Donis RO. FluGenome: a web tool for genotyping influenza A virus. Nucleic Acids Research. 2007; 35: 275-9.

94. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Tatusova T. FLAN: a web server for influenza virus genome annotation. Nucleic Acids Research. 2007: 1–5.

95. Chang S, Zhang J, Liao X, Zhu X, Wang D, Zhu J, et al. Influenza Virus Database (IVDB): an integrated information resource and analysis platform for influenza virus research. Nucleic Acids Research. 2006,; 00(Database issue): D1–D5.

96. Wong SSY, MRCPath, Yuen K-y. Avian Influenza Virus Infections in Humans CHEST. 2006; 129(1): 156-68.

97. Catherine A. Macken, Webby RJ, Bruno WJ. Genotype turnover by reassortment of replication complex genes from avian Influenza A virus. General Virology 2006; 87: 2803–15.

98. Clercq ED. Antiviral agents active against influenza A viruses. Nature. [Reviews]. 2006; 5: 1015-25.

99. Reynisson E, Josefsen M.H, Krause M, Hoorfar J. Evaluation of probe chemistries and platforms to improve the detection limit of real-time PCR. Microbiol Method. 2005.

100. Stone B, Burrows J, Schepetiuk S, Higgins G, Hampson A, Shaw R, et al. Rapid detection and simultaneous subtype differentiation of influenza A viruses by real time PCR. J Virol Methods. 2004 May; 117(2): 103-12.

101. Voeten JT, Bestebroer TM, Nieuwkoop NJ, Fouchier RA, Osterhaus AD, Rimmelzwaan GF. Antigenic drift in the influenza A virus (H3N2) nucleoprotein and escape from recognition by cytotoxic T lymphocytes. J Virol. 2000 Aug; 74(15): 6800-7.

102. Pozhitkov AE, Tautz D. An algorithm and program for finding sequence specific oligonucleotide probes for species identification. BMC Bioinformatics. 2002; 3: 9.

103. Malanoski AP, Lin B, Wang Z, Schnur JM, Stenger DA. Automated identification of multiple micro-organisms from resequencing DNA microarrays. Nucleic Acids Res. 2006; 34(18): 5300-11.

104. Lin B, Wang Z, Vora GJ, Thornton JA, Schnur JM, Thach DC, et al. Broad-spectrum respiratory tract pathogen identification using resequencing DNA microarrays. Genome Res. 2006 Apr; 16(4): 527-35.

105. Michael D. Kane, KTimothy A. Jatkoe, Craig R. Stumpf, Jia Lu, Thomas JD, Madore SJ. Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays. Nucleic Acids Res. 2000;28(22):4552-57.

106. Chou C-C, Lee T-T, Chen C-H, Hsiao H-Y, Lin Y-L, Ho M-S, et al. Design of microarray probes for virus identification and detection of emerging viruses at the genus level. BMC Bioinformatics 2006;7:232.

107. Schattner P. Automated Querying of Genome Databases. PLoS Comput Biol. 2007; 3(1).

108. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403-10.

109. Karaman MW, Groshen S, Lee C-C, Pike BL, Hacia JG. Comparisons of substitution, insertion and deletion probes for resequencing and mutational analysis using oligonucleotide microarrays. Nucleic Acids Res. 2005;33(3).

110. Kessler N, Ferraris O, Palmer K, Marsh W, Steel A. Use of the DNA Flow-Thru Chip, a Three-Dimensional Biochip, for Typing and Subtyping of Influenza Viruses. Clin Microbiol. 2004; 42(5): 2173–85.

111. Rodin A, JR. THM, Clark AG, Sing CF, Boerwinkle E. Mining genetic epidemiology data with Bayesian networks application to APOE gene variation and plasma lipid levels. J Comput Biol. 2005; 12(1): 1-11.

112. Nordberg EK. YODA: selecting signature oligonucleotides. Bioinformatics. 2005 Apr 15; 21(8): 1365-70.

113. Smagala JA, Dawson ED, Mehlmann M, Townsend MB, Kuchta RD, Rowlen KL. ConFind: a robust tool for conserved sequence identification. Bioinformatics. 2005 Dec 15; 21(24): 4420-2.

114. Linos A, Kotsioni I, Papageorgiou E. Data mining: a new technique in medical research. Hormones (Athens). 2005 Oct-Dec; 4(4): 210-2.

115. Maurer M, Molidor R, Sturn A, Hartler J, Hackl H, Stocker G, et al. MARS: Microarray analysis, retrieval, and storage system. BMC Bioinformatics. 2005; 6(101).

116. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004 Aug 19; 5: 113.

117. Onodera K, Melcher U. VirOligo: a database of virus-specific oligonucleotides. Nucleic Acids Res. 2002 Jan 1; 30(1): 203-4.

118. Curk T, Demsar J, Xu Q, Leban G, Petrovic U, Bratko I, et al. Microarray data mining with visual programming. Bioinformatics. 2005 Feb 1; 21(3): 396-8.

119. Kothapalli R, Yoder SJ, Mane S, Loughran TP, Jr. Microarray results: how accurate are they? BMC Bioinformatics. 2002 Aug 23; 3: 22.

120. Linder R, Dew D, Sudhoff H, Theegarten D, Remberger K, Pöppl SJ, et al. The ‘subsequent artificial neural network’ (SANN) approach might bring more classificatory power to ANN-based DNA microarray analyses. Bioinformatics. 2004; 20 (18): 3544–52.

APPENDIX

APPENDIX

1. Instruments and laboratory supplies for single-step target preparation

- Top-bench microcentrifuge : JOUAN, France
- Automated pipette, P2/10/20/100/200/1000 µl : Gilson, France
- Thermocycler 9700 : Perkin Elmer Cetus, USA
- Microcentrifuge tube (0.5 ml and 1.5 ml in sizes) : Treff, Switzerland
- Pipette tips 2-10 µl : Gilson (white)
  20-200 µl : Coster (yellow)
  1000 µl : Treff (blue)
- Horizontal gel electrophoresis : Mupid, Japan
- UV- transluminator with camera computer processing : Stratagene, Eagle eye II Still Video System, USA
- Power supply : Eric line Co.Ltd.

2. Instruments and laboratory supplies for hybridization and imaging of influenza A research Microarray 12K

- Influenza A Research Microarray 12K
- Hybridization Cap (50 µl Volume)
- Clips for Hybridization Cap
- LifterSlip™ coverslip
- Imaging Solution Standard, high-resolution fluorescent microarray scanner (CombiMatrix recommends the Axon Instruments Genepix® 4000B and 4200A, and the Perkin Elmer ScanArray® 4000, 5000, Lite and Express microarray scanners).
- Adhesive tape: Scotch® Brand Magic® Transparent Tape is suitable for hybridization temperatures of 50oC or less; for extended incubations at higher temperatures use a PCR sealing tape such as Nunc

Brand PN 232702 (clear polyolefin liner) or PN 276014 (aluminum liner)

- Rotisserie oven for array hybridization
- CombiMatrix CustomArray™ holders for rotisserie oven
- 95ºC Heating Block
- Pipettors and nuclease-free tips
- Sterile plastic ware
- Gloves (powder-free)

3. Reagents and materials for single-step target preparation

- Total RNA sample (10 to 100 ng)
- CombiMatrix Influenza A Sub-typing Primer Kit for Single-Step Target Preparation, Product Number 610028
- Nuclease-free water
- 0.4 mM Biotin-14-dCTP, Invitrogen, Cat. #19518-018
- SuperScript™ One-Step RT-PCR System containing the RT/ Platinum® Taq Mix (a mixture of SuperScript™ II Reverse Transcriptase and Platinum® Taq DNA Polymerase) and the 2X Reaction Mix (a buffer system optimized for reverse transcription and PCR amplification with Mg2+, and deoxyribonucleotide triphosphates), Invitrogen Cat. #10928-034

4. Reagents for agarose gel electrophoresis

- 1X TBE buffer : 0.089 M Tris-borate, 0.89 M Boric acid, 0.02 M EDTA
- Gel loading dye solution
- Gel marker : contain linear double stranded DNA band of 1000, 700, 525, 500, 400, 300, 200, 100 and 50 base pairs (bp) : Bioactive, USA

5. **Reagent and materials for hybridization and imaging of influenza A research Microarray 12K**
    - Nuclease-free water
    - 50x Denhardt's solution
    - 0.5 M EDTA (pH 8.0)
    - 10% Tween-20
    - 20x SSPE Buffer
    - 1% SDS
    - 10x Phosphate Buffered Saline (PBS: 1.37M Sodium Chloride, 0.027M Potassium Chloride, 0.08M Sodium Phosphate dibasic, 0.02M Sodium Phosphate monobasic, pH 7.4; Ambion, Cat.# 9625)

5X PBS-Casein Blocking Buffer, BioFX Laboratories, Cat.# PBSC-0100-01 Biotin detection reagent: Fluorolink™ Cy5®-labeled streptavidin, GE Healthcare/Amersham Biosciences Cat. # PA45001

6. H5 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
535	H5|AY296080	GTACCCAACAATAAAGAGGAt
536	H5|AY728894	CCCAACAATAAAGAGGAGCt
537	H5|AY296079	CCAACAATAAAGAGGACCTACt
538	H5|AY684706	ATTGTAGTGTAGCTGGATGGt
539	H5|AY724793	GAGTTCTTCTGGACAATTTTAAt
540	H5|AY830774	CCAATGTGTGACGAATTCAt
541	H5|AJ621807	GGATTGTAGTGTAGCTGGAt
542	H5|AY651369	CCCAATGTGTGACGAATTCt
543	H5|U37179	CCATTTTGAAAAAATTCAGATAtttt
544	H5|AY296077	GCACAAACCATTTTGAGAAAt
545	H5|AJ867074	ATTTTGAGAGATTGTAGTGTAGCt
546	H5|AY590571	AGAGAGAAGAAGAAAAAAGAGAt
547	H5|AJ305306	CACAAACCATTTAGAGAAAATTCt
548	H5|AY651373	CCTTTTTCAGGAATGTGGtt
549	H5|AY651371	GAAATGTGGTATGGCTTATCAt
550	H5|AF164662	ACAATAAAGAGGACCTACAAtt
551	H5|AF509025	AGGTTGACACAATAATGGAAt
552	H5|AY770991	CTAATTTTGAGAGATTGTAGTGtt
553	H5|AY497095	ACACAATCATGGAAAAGAATGt
554	H5|AY651362	CTATTGAGCAGAATAAACCAttt
555	H5|AF164659	ATCCAACAATAAAGAGGACCt
556	H5|AY553787	AAACCATTTTGAGAAAACTCAt
557	H5|AF509031	AGGTTGACACAATAATGGAGt
558	H5|AY724795	CCAACCAAGAAGATCTTTTGt
559	H5|AY651326	CTCCTTTTTCAGAAATGTGGt
560	H5|AY724793	ACTTCGAGAGTAATGGAAAttt
561	H5|AY497070	CAAACAATAAAGAGGACCTACAt
562	H5|AY497072	GAACAATAAAGAGGACCTACAt
563	H5|U37167	AAACAATAAAGAGGACCTACAt
564	H5|AY497084	GTAGCTGGATGGCTTCttt
565	H5|M18001	CAACAATAAAGAGGACTTACAAt
566	H5|AY770991	GCAGAATAAACCATTTTGAGAt
567	H5|AY684706	GAGCAGAATAAACCATTTTGAt
568	H5|AY830774	AGAAAATTCAGATCATCCCCt
569	H5|AY296079	AGCACAAACCATTTTGAGAt
570	H5|AY609312	AAGAACTGAAACACCTATTGAt
571	H5|AY651330	CCTATTGAGCAGAATAAACCAt
572	H5|AY770991	CACCTATTGAGCAGAATAAACt
573	H5|M10243	ACAAACCTTTTTGAGAAAATTAGt
574	H5|AF164665	CACATGCTCAAGACATACTGt
575	H5|AY497067	ATTGCAGTGTAGCTGGAtt

7. H5 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
576	H5|AY684706	GAAAATTCAGATCATCCCCAt
577	H5|AY684706	ATGTGTGACGAATTCATCAAt
578	H5|AY651325	AGAACTGAAACACCTATTGAGt
579	H5|AY590570	GAACTGAAACACCTATTGAGCt
580	H5|AY075027	GAGCAGAACAAACCATTTTGt
581	H5|AY553796	AAACACCTATTGAGCAGAAtt
582	H5|AY651330	CAGAAATGTGGTATGGCttt
583	H5|AY724785	AACTTCGAGAGTAATGGAAAtt
584	H5|AY651345	CTATGTGTGACGAATTCATCAt
585	H5|AY651362	CCCTCCTTTTTCAGAAATGtt
586	H5|AF082042	CAGTACAAACCATTTTGAGAAt
587	H5|U37176	GCTCAGCATGTCCATACAt
588	H5|AY728894	ACAGTACATACCCAACAATAAt
589	H5|AY653200	AGAGCAGGTTGACACAAtt
590	H5|AF046097	AATTCAGATCATCCCCAAAAt
591	H5|AY728894	ATTCAGATCATCCCCAAAAGt
592	H5|AY639405	CAAGAACTGAAACACCTATTGt
593	H5|AY684706	ATCATCCCCAAAAGTTCTTGt
594	H5|AF164665	AACAATAAAGAAGACCTACAATAAt
595	H5|AF164661	CACACATGCTCAAGACATAt
596	H5|AY296070	CTCAAGACATACTGGAAAAAGt
597	H5|AF290443	GATTGTAGTGTAGCTGGGtt
598	H5|AY553807	CAGATCATCCCCAAAAGTTCt
599	H5|AY651362	CTGAAACACCTATTGAGCAt
600	H5|AF509025	AATAATGGAAAAGAACGTTACtt
601	H5|AF216729	AAACACCTATTGAGCAGAACt
602	H5|AY770991	ATAATGGAAAAGAACGTTACTGt
603	H5|AY770991	CAATGTGTGACGAATTCATCt
604	H5|AY221523	CTGAGAGATTGTAGTGTAGCt
605	H5|AY684706	GTCCAATCATGAAGCCTCAt
606	H5|AY651368	ACTGAAACACCTATTGAGCt
607	H5|AY651330	AGCAGGTTGACACAATAATGt
608	H5|AY651360	ATTCAGATCATCCCCAAAAAt
609	H5|AY651349	CCAATCATGAAGCCTCATCt
610	H5|AY576930	ATCAATGTGCCGGAATGGt
611	H5|AY651356	ACTTATTGAGCAGAATAAACCAt
612	H5|AY651330	CACAATAATGGAAAAGAACGtt
613	H5|AY651352	AAACATCTATTGAGCAGAATAAAt
614	H5|AY651367	ATTGTAGTGTAGCTGGGTGt
615	H5|AY651343	ACAATAATGGAAAAGAACGttt
616	H5|AY651362	ATCATGAAGCCTCATCAGGt

8. H5 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
617	H5|AY590572	CAGAAATACCCCTCAAAGAGt
618	H5|AY576930	GACGAATTCATCAATGTGCt
619	H5|AF501235	ACGAATTCATCAATGTGCCt
620	H5|AY651330	AATTGAAACACCTATTGAGCAt
621	H5|AF501235	CCTATGTGTGACGAATTCAtt
622	H5|AY770079	AATGACCTCTGTTACCCAGt
623	H5|AY590568	GATCAGATTTGCATTGGTTACt
624	H5|AF509032	CAAAAAGAACAGTGCATACCt
625	H5|AY651362	CATGCCCAAGACATACTGGt
626	H5|AY830774	CCCAAGACATACTGGAAAAGt
627	H5|AY770991	AGTGAAGCCTCTAATTTTGAGt
628	H5|U20460	GAGAAAATTCAAATCATCCCCt
629	H5|AY651367	AGAAAATTCAAATCATCCCCAt
630	H5|AY609312	CCAAGACATACTGGAAAAGAt
631	H5|AY651364	CTTATCAAAAAGAACAATGCAtt
632	H5|AY651367	GGTCCAATCATGAAGCCtt
633	H5|AY553796	GGAGTGAAGCCTCTAAttttt
634	H5|AY651354	AGCCAATGACCTCTGTTACt
635	H5|AF046088	CCAGATCATCCCCAAAAGtt
636	H5|AY770079	AAAGTGATCAGATTTGCATTGt
637	H5|AY651347	AAAAGTGATCAGATTTGCAttt
638	H5|AY651366	GCTTATCAAAAAGAACAATGCt
639	H5|AY651324	CTAATTTTAAGAGATTGTAGTGTAGt
640	H5|AY590576	AACTTCTGGTTCTCATGGAt
641	H5|AY497080	GCGTAGCTGGATGGCtt
642	H5|AF509025	AGAACGTTACTGTTACACATGt
643	H5|AY585370	AAAATTCAGATCATACCCAAAAt
644	H5|AY590576	AATGTCAAGAACCTTTACGAt
645	H5|AY651347	CCCTATGTGTGACGAATTCt
646	H5|AY651343	AAAAGAACGTTACTGTTACACt
647	H5|AY651324	AAAGACACACAACGGGAAt
648	H5|AY552000	AGTCAATGACCTCTGTTACCt
649	H5|AY651330	ATGGAGTGAAGCCTCTAAtt
650	H5|AY830774	ACTGGAAAAGACACACAACt
651	H5|AY651325	AAAGTCCTCCTTTTTTAGAAATGt
652	H5|AY553788	AGAGATTGTAGTGTAGCCGt
653	H5|AF216737	GTGATCAGATTTGCATTGGtt
654	H5|J02160	CCAGATTTGCATTGGTTACCt
655	H5|AF102681	AGCCAATGACCTCTGTTAtt
656	H5|AY651353	AAAAGACACACAACGGGAt
657	H5|AY646424	CTATGTGTGACGAATTCACCt

9. H5 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
658	H5|S68489	GAAAATTCAAATCATCCCCAGt
659	H5|AY553808	ATGTGGTATGGCTTATCCAt
660	H5|AF364334	GAGTGAAGCCTCTCATTTTGt
661	H5|AY651355	AAGACATACTGGAAAAGACCt
662	H5|AY590573	ACCCAACAATCAAGAGGAGt
663	H5|AY651371	AATGTGCCGGAATGGTCt
664	H5|AY590571	AGGGAGGATGGCAGGGt
665	H5|AY651350	CAGATCATCCCCAAAAATTCtt
666	H5|AJ621807	CTCAATGTGCCGGAATGGt
667	H5|AY770991	CAATGTGCCGGAATGGtt
668	H5|AF364334	GATCTAAATGGAGTGAAGCCt
669	H5|AF046097	AAGCTCAGCATGTCCATAt
670	H5|U20460	GAGGGATTGTAGTGTAGCtt
671	H5|U20460	AAACACCTATTGAGCAGTACt
672	H5|U79451	AGTAATGAATACCCAACAATAAAt
673	H5|AY770991	GAACGTTACTGTTACACATGt
674	H5|ISDN49024	ACCAGATTTGCATTGGTTACt
675	H5|AY590572	GACTGGGCTCAGAAATACCt
676	H5|AY651323	CCTCCTTTTTTAGAAATGTGGt
677	H5|AY590573	AACAGTACATACCCAACAATCt
678	H5|AY651330	GGAAAAGACACACAACGGt
679	H5|AY553796	AAAAGTTCTTGGTCCAGTCt
680	H5|AY585365	GCGATCTAAATGGAGTGAAGt
681	H5|AF102681	GCCTGAATGGTCTTACATAGt
682	H5|AF082042	ATGGAGTGAAGCCTCTCAt
683	H5|AY585365	ATCAGTCTTGTTAAAAGTGATCt
684	H5|AY651330	GGAAGCTCTGCGATCTAGt
685	H5|AF098543	AATGACCTCTGTTATCCAGGt
686	H5|AY553796	GGATGGCTCCTCGGAAt
687	H5|AY075027	ACACCTATTGAGCAGAACAt
688	H5|AF046097	CAGAATAAGCCATTTTGAGAAt
689	H5|AY609312	CATGAAGCCTCATCAGGGt
690	H5|AY651325	ATGAAGCCTCATCAGGGGt
691	H5|AF194169	CCATGTGTGACGAATTCCt
692	H5|AY651352	GAAAATTCAGATCTTCCCCAt
693	H5|AY651367	CTAAAAGTGATCAGATTTGCAt
694	H5|AY651352	AGGAATGTGGTATGGCttt
695	H5|AY590575	AAAGTTCTTGGTCCAGTCAt
696	H5|AF509037	ATTGAGCACAATAAACCAtttt
697	H5|AY576930	AAAGCTCTATCAAAATCCAACt
698	H5|AF509025	AAATGTGGTATGGCTCATCt
699	H5|AY830774	GACGAATTCATCAATGTACCGt
700	H5|AY651353	CTAATTTTGAAAGATTGTAGTGTAt
701	H5|AY651347	AGCAGGTTGATACAATAATGGt
702	H5|AY590576	ACAAGGTCCGACTACAGCt

10. N1 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
1457	N1|AY553818	GTATTGAAATACAATGGCATAATAt
1458	N1|AY553838	ATACAATGGCATAATAACAGACt
1459	N1|AY575892	ATTGAAATACAACGGCATAATAt
1460	N1|AY553826	GCATAATAACAGACACTATCAAt
1461	N1|AY553812	AATCAAAATTTGGAGTATCAAATAt
1462	N1|AY553826	CCTAATTATCACTATGAGGAATGt
1463	N1|AY553818	AATGGCATAATAACAGACACtt
1464	N1|AY553813	CAAAATTTGGAGTATCAAATAGGt
1465	N1|AY553826	TGGCTGTATTGAAATACAATGt
1466	N1|AY553838	CTGTATTGAAATACAATGGCAt
1467	N1|AY574188	CCACTTGGAATGCAGAACt
1468	N1|AY180836	TCACTTGGAATGCAGAACCt
1469	N1|AF398869	CTCACTTGGAATGCAGAACt
1470	N1|AY553828	TTGCTTTACTGTAATGACTGAt
1471	N1|AY553826	GACACTATCAAGAGTTGGAGt
1472	N1|AY553816	ACAGACACTATCAAGAGTTGt
1473	N1|AY553812	TGCTTTACTGTAATGACTGACt
1474	N1|AY553828	AGTCTGAATGTGCATGTGtt
1475	N1|AY576928	TTGACAATTGGAATTTCTGGCt
1476	N1|AF494254	CTTGGAATGCAGAACCtttt
1477	N1|AF509092	ATTGGATCAATCTGTATGGTAAt
1478	N1|K01031	AACCATTGGATCAATCTGTCt
1479	N1|AY553812	GCTCCTAATTATCACTATGAGGt
1480	N1|AY576928	TCAAGAGTTGGAGGAACAAt
1481	N1|AY623431	TGACAATTGGAATTTCTGGCt
1482	N1|AF028708	TTGGAATGCAGAACCttttt
1483	N1|AF509092	CAATATGGGTTAGTCATTCAAtt
1484	N1|AY553813	TATCTTTCAATCAAAATTTGGAGt
1485	N1|AF057292	TTTGGAATGCAGAACCtttt
1486	N1|AY651480	GGATCAATCTGTATGGTAATTGt
1487	N1|AY553812	GGGTATCTTTCAATCAAAATTTGt
1488	N1|Z30275	CATTCACCATTGACAAGTAAtt
1489	N1|AY590565	CGGGCAATTCATCTCTTTGt
1490	N1|AF305216	TGGCTGTATTAAAATACAACGt
1491	N1|AY590567	GGGATGTGTTTGTTATAAGAGt
1492	N1|AY653195	GGGGATGTGTTTGTTATAAGAt
1493	N1|AJ416629	TTGGAATGCAGAACCTTCtt
1494	N1|AY553826	ATGTGCATGTGTAAATGGCt
1495	N1|AY553828	TGGGTATCTTTCAATCAAAAttt
1496	N1|AY553828	CCTGTTATCCTGATGCCGt
1497	N1|AY619960	AAATTGGAAATATAATCTCAATATGGt

[a]

11. N1 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
1498	N1|AJ416628	AACATAATCTCAATATGGGTTAGt
1499	N1|ISD3BF2A798	CGGCAATTCATCTCTTTGtt
1500	N1|AF398872	CAATTCATCTCTTTGCCCtt
1501	N1|AY575883	CAATCTGTATGGTAATTGGAAtt
1502	N1|AY590565	TTTGAGTCTGTTGCTTGGtt
1503	N1|AY574188	GGTTTGAGTCTGTTGCTTGt
1504	N1|AF250481	ATTTGGAATGCAGAACCttt
1505	N1|AY575881	GGAACATAATCTCAATATGGGtt
1506	N1|AY651434	ATCAATCTGTATGGTAATTGGAt
1507	N1|AY553828	TAAATGGCTCTTGCTTTACtt
1508	N1|AJ518091	GGCAATTCATCTCTTTGTTCt
1509	N1|AY609314	ATAACCATCGGATCAATCTGt
1510	N1|M73976	GGCAATTCATCTCTTTGTCCt
1511	N1|AF509092	GAACATAATCTCAATATGGGttt
1512	N1|AJ410876	AAATTGGAAACATAATCTCAATATGt
1513	N1|AJ410561	CATTTGGAATGCAGAACCtt
1514	N1|AF509114	ATTGGATCAATCTGTATGATAAttt
1515	N1|AY553813	GTGGAGTTTTCGGAGACAt
1516	N1|AY660557	TTATAAGAGAGCCATTCATCTCt
1517	N1|AY585409	GTGGATGGGCTATATACAGtt
1518	N1|AJ410560	CAAGGTTTGAGTCTGTTGCt
1519	N1|AY651476	CCCATATAACTCAAGGTTTGAt
1520	N1|AJ416627	ATAATCTCAATATGGGTTAGCCt
1521	N1|AJ410875	TAGTGGATGGGCTATATACAGt
1522	N1|AY575883	CTCAATATGGGTTAGTCATTCAt
1523	N1|AY585403	ATCTCAATATGGGTTAGTCAttt
1524	N1|AY646176	CCCACTTGGAATGCAGAt
1525	N1|AJ412689	TGGAATGCAGAACCTTCtt
1526	N1|AY651473	CATAATCTCAATATGGGTTAGTCt
1527	N1|AY590567	ATTAGCGGGCAATTCATCtt
1528	N1|AJ410560	GATCAATCTGTATGGTAGTTGGt
1529	N1|AJ410878	TCTCCATATAACTCAAGGTTTGt
1530	N1|AY553826	AGAGTTGGAGGAATAACATACt
1531	N1|AY576928	TTGAGTCTGTTGCTTGGtt
1532	N1|AJ416625	GGAAACATAATCTCAATATGGGt
1533	N1|AY585403	ATTGGAAACATAATCTCAATATGGt
1534	N1|AY553812	CTGTTATCCTGATGCCGGt
1535	N1|AJ410558	AATAATTAGCTTAATGTTACAAATTGt
1536	N1|AF250481	GGATGGGCTATATACAGTAAAt
1537	N1|AY553828	TATCACTATGAGGAATGCTCCt
1538	N1|D31947	GTTCTCACTTGGAATGCAGt

12. N1 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
1539	N1|AY633118	CTCTCACTTGGAATGCAGGt
1540	N1|AY553814	TTAAATCAGTCGAATTGAATGCt
1541	N1|AF389120	GCAATTCATCTCTTTGTCCCt
1542	N1|AY651473	TATGGGTTAGTCATTCAATTCAt
1543	N1|AY575881	GTTAGCTTGATGTTACAAATTGt
1544	N1|AJ410875	ATTCATCTCATGCTCCCACt
1545	N1|AY651462	ATAGTTAGCTTAATGTTACAAATTGt
1546	N1|AY526746	TGGATGCTCCTAATTATCACtt
1547	N1|AY701754	GATAATTGGCATGGCTCGt
1548	N1|AY059485	TAATCTCAATATGGGTTAGCCt
1549	N1|AY576928	ATGCTCCCACTTGGAATGt
1550	N1|AY651473	TAGCGGGCAATTCATCTCt
1551	N1|AY590567	GAGTCTGTTGCTTGGTCAt
1552	N1|AY790269	TCTCACTTGGAATGCAGGt
1553	N1|AY651449	GATCAATCTGTATGGTAACTGGt
1554	N1|AY633286	GTGGATGGGCTATATACAGCt
1555	N1|AY553816	AGATCTTCAAAATGGAAAAAGGt
1556	N1|K01003	GATCAATCTGTATGGTAGTCGt
1557	N1|AJ412692	GTGTTTGTCATAAGAGAGCCt
1558	N1|K01033	AATATTATCTCAATATGGGTCAGCt
1559	N1|AY623431	TTCATCTCATGCTCCCACt
1560	N1|AY575884	CGTCAAAGACAGAAGCCCt
1561	N1|AY526746	CTGACGGACCAAGTAATGGt
1562	N1|AY779049	GTTAGCTTAATGTTACAAATTGGt
1563	N1|AY651479	GAGTCTGTTGCTTGGTCGt
1564	N1|AJ410564	GCAGGCAATTCATCTCtttt
1565	N1|AY590567	TACAGTAAGGACAACAGTATAAt
1566	N1|AF250359	CTCACTTGGAATGCAGGAt
1567	N1|AY553816	TCAAGAGTTGGAGGAATAACt
1568	N1|AY653195	CATCTCATGCTCCCACttt
1569	N1|K02252	TATTACAAATTGGAAACATAATCTCt
1570	N1|AY651434	TGCTCTCACTTGGAATGCt
1571	N1|AY553814	GAATTGAATGCTCCTAATTATCAt
1572	N1|AY535028	GATCGGGAGAACCAAAAGCt
1573	N1|AY553812	TTAAATCAGTCGAATTGGATGt
1574	N1|AY651482	GCGGGCAATTCATCTCtt
1575	N1|AY221546	ATCAATCTGTATGGTAGTTGGt
1576	N1|AY651434	GGGGGATGTGTTTGTTAttt
1577	N1|AY585410	ATAGTTAGCTTGATGTTACAAAtt
1578	N1|AY553826	TGTAAATGGCTCTTGCtttt
1579	N1|AF509097	TAGCTTAATGTTACAAATTGGGt

13. N1 oligonucleotide sequence on microarray for medium and high resolution

Probe No.	Name	Oligonucleotide sequence
1580	N1|AJ518095	AATTCATCTCTTTGTTCTATCAGt
1581	N1|AY553826	GAAAAAGGGAAAGTGGTTAAAt
1582	N1|AF057292	ATTCATCTCATGCTCCCAtt
1583	N1|AJ410557	AAGGGGATGTGTTTGTTATGt
1584	N1|AY623431	GACCGTCAAAGACAGAAGCt
1585	N1|AY553813	TTGGCATGGCTCAAATCGt
1586	N1|AF208598	TACAAATTGGAAACATAATCTCAt
1587	N1|AY526746	AAAAGGGAAAGTGGTTAAATCt
1588	N1|AY553818	TGGCTCTTGCTTTACTGtt
1589	N1|AY646425	AATAGTTAGCTTGATGTTACAAAt
1590	N1|AY575888	ATCTGTATGGTAATTGGAATAGtt
1591	N1|AY574188	CAAGGGGGATGTGTTTGtt
1592	N1|AY261521	TCATGCTCTCACTTGGAAt
1593	N1|AY590564	TCTTGTTATCCTGATGCCGt
1594	N1|AY575884	TTGATGTTACAAATTGGGAACt
1595	N1|AF208598	CACTTGGAATGCAGGACCt
1596	N1|AF250362	GATGTTACAAATTGGAAATATAATCt
1597	N1|AY619960	TAAAGACAGAAGCCCTTATAGt
1598	N1|AY590564	CATATTATCACTATGAGGAATGttt
1599	N1|AY649383	TTAATGTTACAAATTGGGAACtt
1600	N1|AY553813	AATCCACGCCCCAATGt
1601	N1|AY576928	GTTCATCTCATGCTCCCAt
1602	N1|AJ410565	GGTCAATCTGTATGGTAATTGt
1603	N1|AY122326	ATCAATCTGTATGGTAGTCGGt
1604	N1|AJ410564	TCATCTCTTTGCCCTATTAGtt
1605	N1|AJ410558	CTATTAGTGGATGGGCTATATACt
1606	N1|AY651482	GTAACATTAGCGGGCAAttt
1607	N1|AY770992	GGTTAGCTTAATGTTACAAATTGt
1608	N1|AF398422	ATAATCTCAATATGGGTTACTCAt
1609	N1|AY576928	CTCCCCATATAACTCAAGGtt
1610	N1|AY649383	ATCAATCTGTATGGTAACTGGt
1611	N1|AY553813	GGCCATGGGTATCTTTCAt
1612	N1|AF036357	GTTACAAATTGGAAACATAATATCAt
1613	N1|AF046081	TTATAAGAGAACCATTCATCTCAt
1614	N1|AY553812	TGCATGTGTAAATGGCTCt
1615	N1|AJ518095	AGTGGATGGGCTATATACACt
1616	N1|AF144304	ATAATTAGCTTGATGTTACAAAttt
1617	N1|AY180836	CAGTGGATGGGCTATATACAt
1618	N1|AY585409	TATATACAGTAAGGACAACGGt
1619	N1|AY518363	GGGGATGTGTTTGTTATTAGAt
1620	N1|AY651467	ACCGTCAAAGACAGAAGCt
1621	N1|AY651479	GATCAATCTGTATGATAATTGGAt
1622	N1|AF046081	ACATTAGCGGGCAATTCCt
1623	N1|AY651437	TTCATCTCATGCTCCCAttt
1624	N1|AY590567	CAGTTGGTTGACAATTGGAt

14. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.11

Group	Mean	N	SD	Minimum	Maximum
H1	633	168	126.2	350	1187
H10	634	50	95.23	361	820
H11	634	50	121.41	346	924
H12	623	50	137.12	337	1044
H13	625	50	109.98	424	870
H14	673	50	107.73	422	899
H15	645	50	110.9	412	892
H16	478	18	138.81	232	724
H2	602	128	111.8	352	1133
H3	581	168	124.76	216	1371
H4	598	50	108.32	358	1045
H5	***3211***	***168***	***1675.88***	***576***	***7569***
H6	599	50	117.64	280	858
H7	600	168	106.06	249	1006
H8	649	50	108.52	382	911
H9	648	168	215.04	193	1912
N1	***2677***	***168***	***2055.97***	***181***	***8348***
N2	600	168	289.92	271	2268
N3	541	50	102.02	358	734
N4	585	50	144.83	235	916
N5	597	50	118.56	303	833
N6	585	50	103.28	336	865
N7	590	168	157.98	288	1729
N8	559	50	153.92	317	1169
N9	614	50	126.63	295	962
Total	958	2240	1119.15	181	8348

15. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.17

Group	Mean	N	SD	Minimum	Maximum
H1	632	168	133.75	255	1133
H10	632	50	100.43	403	842
H11	608	50	117.44	304	906
H12	604	50	124.62	256	877
H13	618	50	118	347	898
H14	662	50	108.82	480	855
H15	666	50	134.25	323	978
H16	466	18	106.16	327	719
H2	598	128	109.77	319	823
H3	581	168	128.21	230	885
H4	627	50	117.07	385	878
H5	***2403***	***168***	***1444.01***	***425***	***7190***
H6	588	50	123.32	336	860
H7	603	168	114.79	250	968
H8	642	50	121.79	376	915
H9	618	168	110.29	332	896
N1	***1789***	***168***	***1392.01***	***203***	***6208***
N2	573	168	166.37	267	1451
N3	556	50	122.94	298	847
N4	590	50	108.87	301	824
N5	594	50	142.96	254	833
N6	586	50	122	321	844
N7	579	168	142.19	269	1225
N8	563	50	166.01	234	1215
N9	615	50	148.97	247	1033
Total	826	2240	783.55	203	7190

16. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.25

Group	Mean	N	SD	Minimum	Maximum
H1	733	168	183.02	367	1290
H10	767	50	157	443	1142
H11	784	50	215.67	424	1289
H12	743	50	225.64	321	1305
H13	729	50	158.55	446	1090
H14	828	50	176.41	497	1313
H15	743	50	167.32	281	1201
H16	530	18	217.02	187	1083
H2	735	128	198.54	248	1577
H3	677	168	176.13	207	1122
H4	688	50	127.96	404	1025
H5	***5202***	***168***	***2634.97***	***608***	***12071***
H6	695	50	157.19	342	1026
H7	717	168	176.5	301	1345
H8	754	50	139.94	410	1194
H9	737	168	214.91	268	1721
N1	***2819***	***168***	***2220.19***	***291***	***8305***
N2	718	168	335	364	2322
N3	639	50	175.41	310	1088
N4	731	50	213.25	367	1215
N5	710	50	208.79	327	1267
N6	656	50	158.19	310	1007
N7	710	168	227.66	214	2133
N8	685	50	250.44	284	1506
N9	731	50	182.25	293	1333
Total	1213	2240	1586.91	187	12071

17. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.27

Group	Mean	N	SD	Minimum	Maximum
H1	769	168	211.44	229	2064
H10	747	50	152.73	419	1155
H11	798	50	210.24	422	1336
H12	775	50	228.75	393	1422
H13	775	50	159.46	378	1211
H14	833	50	158.69	600	1291
H15	790	50	184.23	464	1322
H16	567	18	218.81	204	1005
H2	739	128	150.94	312	1240
H3	701	168	167.52	309	1406
H4	734	50	133.35	452	1011
H5	***6463***	***168***	***3234.05***	***585***	***13182***
H6	750	50	158.71	350	1059
H7	733	168	145.17	317	1151
H8	779	50	164.42	456	1248
H9	812	168	282.57	296	2450
N1	***4455***	***168***	***3754.27***	***414***	***13274***
N2	812	168	711.72	236	4921
N3	695	50	160.24	420	1140
N4	727	50	180.68	246	1143
N5	717	50	176.18	313	1113
N6	710	50	169.6	304	1173
N7	741	168	440.21	253	5869
N8	745	50	458.69	362	3705
N9	746	50	187.17	355	1259
Total	1461	2240	2210.65	204	13274

18. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.31

Group	Mean	N	SD	Minimum	Maximum
H1	716	168	266.48	324	2787
H10	714	50	172.58	313	1128
H11	821	50	441.54	338	2633
H12	730	50	225.74	318	1378
H13	702	50	191.76	285	1300
H14	759	50	208.19	344	1394
H15	743	50	185.71	319	1250
H16	486	18	211.2	160	897
H2	699	128	218.18	351	1960
H3	628	168	188.29	131	1146
H4	678	50	138.81	476	1123
H5	***7869***	***168***	***4248.99***	***451***	***16066***
H6	665	50	185.66	404	1138
H7	653	168	170.1	313	1061
H8	725	50	191.73	372	1281
H9	743	168	262.21	351	2022
N1	***5739***	***168***	***4565.9***	***226***	***15485***
N2	701	168	489.08	144	2944
N3	613	50	181.07	367	1112
N4	665	50	217.32	296	1095
N5	688	50	207.99	352	1267
N6	643	50	187.67	311	1135
N7	767	168	639.57	187	4908
N8	719	50	718.79	182	5539
N9	697	50	215.29	284	1339
Total	1616	2240	2814.06	131	16066

19. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.37

Group	Mean	N	SD	Minimum	Maximum
H1	837	168	358.17	303	4020
H10	852	50	302.56	429	2473
H11	969	50	589.95	354	3648
H12	878	50	354.54	303	2072
H13	818	50	219.88	381	1545
H14	888	50	232.25	349	1607
H15	866	50	241.95	400	1544
H16	540	18	242.74	140	985
H2	819	128	380.36	79	3526
H3	719	168	232.94	215	1405
H4	785	50	193.74	514	1501
H5	***11104***	***168***	***4941.8***	***375***	***20898***
H6	774	50	241.91	348	1496
H7	777	168	214.15	345	1555
H8	829	50	253.46	417	1763
H9	872	168	380.3	233	2984
N1	***6138***	***168***	***4898.2***	***291***	***15887***
N2	813	168	572.46	278	3571
N3	732	50	256.72	329	1401
N4	782	50	271.01	154	1445
N5	774	50	259.19	229	1479
N6	734	50	202.64	347	1233
N7	826	168	521.97	205	5238
N8	776	50	560.82	241	4309
N9	806	50	301.19	289	2016
Total	1981	2240	3526.55	79	20898

20. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.47

Group	Mean	N	SD	Minimum	Maximum
H1	670	168	220.83	288	2450
H10	654	50	130.03	366	930
H11	698	50	198	426	1259
H12	651	50	205.03	221	1271
H13	657	50	146.06	430	986
H14	713	50	161.18	446	1177
H15	670	50	175.65	308	1204
H16	480	18	159.25	239	763
H2	632	128	148.33	155	1092
H3	597	168	146.86	238	999
H4	602	50	119.89	343	988
H5	***5652***	***168***	***3116.73***	***529***	***13028***
H6	621	50	154.65	305	978
H7	615	168	129.31	273	1035
H8	668	50	150.95	429	1104
H9	787	168	917.97	296	7202
N1	***4668***	***168***	***3546.97***	***249***	***14235***
N2	684	168	629.73	286	4433
N3	593	50	169.05	288	1111
N4	612	50	177.26	276	945
N5	619	50	152.69	319	956
N6	592	50	145.93	317	989
N7	649	168	349.37	163	2996
N8	614	50	438.89	275	3470
N9	631	50	177.08	228	1196
Total	1328	2240	2101.36	155	14235

21. Distribution and signal intensity of the oligoneucleotides used to detect H5N1 sample no.48

Group	Mean	N	SD	Minimum	Maximum
H1	766	168	239.38	341	2828
H10	752	50	130.14	408	1100
H11	785	50	206.19	429	1309
H12	748	50	212.62	412	1385
H13	757	50	171.1	419	1254
H14	827	50	159.78	479	1155
H15	752	50	172.33	417	1233
H16	500	18	213.1	156	833
H2	730	128	160.67	250	1222
H3	684	168	189.8	274	1869
H4	853	50	192.11	537	1339
H5	***6297***	***168***	***3689.5***	***627***	***15047***
H6	707	50	149.59	417	1028
H7	691	168	152.06	267	1065
H8	769	50	154.43	329	1098
H9	876	168	917.06	158	7031
N1	***4726***	***168***	***3858.87***	***369***	***13729***
N2	764	168	792.83	249	5747
N3	655	50	171.44	424	1123
N4	712	50	208.72	264	1117
N5	711	50	156.95	363	1066
N6	697	50	166.2	290	1025
N7	698	168	259.81	219	2789
N8	943	50	397.82	438	2682
N9	733	50	193.14	213	1354
Total	1463	2240	2292.29	156	15047

BIOGRAPHY

NAME	Miss Supaporn Kaewpongsri
DATE OF BIRTH	11 December 1977
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Chulalongkorn University, 1999-2002 Bachelor of Science (Medical technology) Mahidol University, 2005-2008 Master of Science (Clinical pathology)
SCHOLARSHIP	Supported in part by Thesis Grant, Faculty of Graduate Studies, Mahidol University, Bangkok Thailand
HOME ADDRESS	174 Tha phutsa, Klongklung, Kamphaeng Phet Thailand 62120 E-mail: jum@shoklo-unit.com