2017

Genetic Interference and Receptor Biology of Neglected Influenza Viruses

Runxia Liu
South Dakota State University

GENETIC INTERFERENCE AND RECEPTOR BIOLOGY OF NEGLECTED

INFLUENZA VIRUSES

BY

RUNXIA LIU

A dissertation submitted in partial fulfillment of the requirement for the

Doctor of Philosophy

Major in Biological Sciences

Specialization in Microbiology

South Dakota State University

2017

GENETIC INTERFERENCE AND RECEPTOR BIOLOGY OF NEGLECTED

INFLUENZA VIRUSES

RUNXIA LIU

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy in Biological Sciences, specialization in Microbiology and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Feng Li, Ph.D.
Dissertation Advisor   Date

Volker S Brözel, Ph.D.
Head, Department of Biology   Date
and Microbiology

Dean, Graduate School   Date

ACKNOWLEDGEMENTS

CONTENTS

ABBREVIATIONS

9-O-SE: sialate-9-O-acetylesterase

A549: human lung adenocarcinoma cell

BSM: bovine submaxillary mucin

CMAH: CMP-Neu5Ac hydroxylase

CMV: cytomegalovirus

CPE: cytopathic effects

cRNA: complementary RNA

ddPCR: Droplet Digital PCR

DI: defective interfering

DMEM: Dulbecco's minimum essential medium

FBS: fetal bovine serum

HA: Hemagglutinin

HEF: Hemagglutinin-Esterase-Fusion

HEK293T: human embryonic kidney cells

HI: hemagglutination inhibition assay

hpi: hours post infection

IAV: influenza A virus

IBV: influenza B virus

ICV: influenza C virus

IDV: influenza D virus

M: matrix protein

MDCK: Madin-Darby canine kidney cell

MOI: multiplicity of infection

mRNA: messenger RNA

NA: Neuraminidase

NC: noncoding

NEP: nuclear export protein

Neu4,5Ac2:5-N-acetyl-4-O-acetyl neuraminic acid

Neu5, 9Ac2: 5-N-acetyl-9-O-acetylneuraminic acid

Neu5Ac: 5-N-acetylneuraminic acid

Neu5Gc9Ac: 9-O-acetylated N-glycolylneuraminic acid

NGS: next generation sequencing

NP: nucleoprotein

NS1: non-structural protein 1

NS2: non-structural protein 2

NTC: No template control

ORF: open reading frame

PA: polymerase acidic protein

PB1-F2: polymerase basic protein 1-F2

PB1: polymerase basic protein 1

PB2: polymerase basic protein 2

Pcc: Pearson correlation coefficients

PFU: plaque formation unit

RBCs: red blood cells

RdRp: RNA dependent RNA polymerase

RFU: relative fluorescence units

RGS: reverse genetics system

RNA: ribonucleic acid

RPKM: Reads Per Kilobase of transcript per Million mapped reads

RT-PCR: reverse transcription polymerase chain reaction

SA: sialic acid

SGM: sialylated glycan microarray

TBP: TATA-Box binding protein

TPCK: tolylsulfonyl phenylalanyl chloromethyl ketone

vRNA: viral RNA

vRNP: Ribonucleoprotein

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

GENETIC INTERFERENCE AND RECEPTOR BIOLOGY OF NEGLECTED

INFLUENZA VIRUSES

RUNXIA LIU

2017

Influenza B virus (IBV) is an important pathogen that infects humans and causes seasonal influenza epidemics. By using next-generation sequencing (NGS) approach, we analyzed total mRNAs extracted from A549 cells infected with B/Brisbane/60/2008, and identified four defective genomes in IBV with two from the polymerase basic subunit 1 (PB1) segment and the other two from the matrix (M) segment. Each of them can potently inhibit the replication of IBV. One derived from PB1 segment was able to interfere modestly with influenza A virus (IAV) replication. The productions of the four defective RNAs are not dependent on the cell types. The important initial insights into IBV defective genomes can be further explored toward better understanding of the replication, pathogenesis, and evolution of IBV.

The second study demonstrated that influenza D virus (IDV) is more efficient in recognizing both human Neu5,9Ac$_2$ and non-human Neu5Gc9Ac receptors than influenza C virus (ICV). ICV prefers human Neu5,9Ac$_2$ over non-human Neu5Gc9Ac. The results reveal that IDV and ICV diverge in communicating with both O-acetyl group at the C9 position and acetyl/glycolyl groups at the C5 position in terminal 9-carbon SAs. Our findings provide evidence that IDV has acquired the unique ability to infect and transmit

xvi

among agricultural animals that are enriched in Neu5Gc9Ac, in addition to pose a zoonotic risk to humans only expressing Neu5,9Ac$_2$.

Characterization of a contemporary human ICV is needed. C/Victoria/2/2012 (C/Vic) isolated in 2012 was used in this study. Phylogenetic studies demonstrated that C/Vic is a reassortant virus composed of segments derived from multiple ICV lineages or strains, which evolved independently. We identified two mutations in the 170-loop of the HEF protein around the receptor binding pocket as a possible antigenic determinant responsible for the discrepant hemagglutinin inhibition results. C/Vic replicates more efficiently at the cool temperature, which should be further investigated toward elucidating the molecular determinants of temperature-dependent growth. The study on this contemporary ICV shall aid in the further investigation of biology, evolution, and pathogenesis of ICV.

Chapter 1. Introduction and Literature Review

1.1 Classification of Influenza viruses

Influenza viruses are members of Orthomyxoviridae family. Influenza viruses contain a segmented genome of single-stranded negative-sense RNA. They are a group of respiratory pathogens that has diversified through evolution into three types: A, B, and C. This classification is based on their distinct antigenic properties residing in two viral structural proteins: nucleoprotein (NP) and matrix (M) protein (1980; W. R. Dowdle et al., 1974). Recently, a new type of influenza virus with bovine as a primary reservoir, provisionally designated influenza type D, has been described (B. M. Hause et al., 2014; B. M. Hause et al., 2013). The major difference between influenza A/B viruses (IAV/IBV) and influenza C virus (ICV) is that IAV/IBV possess eight segments while ICV contains seven segments in viral particles. IAV and IBV contain two spike proteins: Hemagglutinin (HA) and Neuraminidase (NA), while ICV possesses only one glycoprotein, named Hemagglutinin-Esterase-Fusion (HEF) protein, which combines both the functions of HA (receptor binding, membrane fusion) and NA (receptor destroying).

	IAV	IBV	ICV	IDV
Segment	8	8	7	7
Subtype	18	1	1	1
Disease	Severe	Severe	Mild	Mild
Host	Multiple species	Human	Human	Multiple species
Reservoir	Waterfowl /Bird	Human	Human	Bovine
Evolution	Fast	Slow	Slow	Slow

Table 1.1. Overview of four types of influenza viruses.

Among these influenza viruses, IAV is of most veterinary importance. Various strains of subtypes of IAV can infect and cause severe diseases in several agricultural animal species, as well as human populations (J. K. Taubenberger and J. C. Kash, 2010). IAVs are divided into subtypes based on two surface proteins: the hemagglutinin (HA) and the neuraminidase (NA). There are 18 different hemagglutinin subtypes and 11 different neuraminidase subtypes (N. M. Bouvier and P. Palese, 2008; S. Tong et al., 2013; Y. Wu et al., 2014).

IBV and ICV are thought to be human pathogens, despite IBV or ICV can be isolated from swine or seals (A. D. Osterhaus et al., 2000; Z. Ran et al., 2015). IBV infection of humans can result in clinical diseases similar to IAV, ranging from mild to severe respiratory illness (A. S. Monto, 2008; W. Paul Glezen et al., 2013). IBV can cause high mortality in seasonal influenza epidemics particularly in children (A. Gutierrez-Pizarraya

et al., 2012; P. Wu et al., 2012). Two antigenic and genetic lineages (Victoria and Yamagata lineages) co-circulated since the 1980s (I. G. Barr et al., 2010; J. A. McCullers et al., 2004; P. A. Rota et al., 1990).

ICV usually causes mild upper respiratory diseases in humans. It has the ability in causing severe lower respiratory illness in infants (S. Katagiri et al., 1983; Y. Matsuzaki et al., 2006; H. Moriuchi et al., 1991). ICV were divided into six discrete lineages by antigenic and phylogenetic analysis, represented by C/ Taylor/1233/47, C/Yamagata/26/81, C/Kanagawa/1/76, C/Aichi/1/81, C/Sao Paulo/378/82 and C/Mississippi/80 (Y. Matsuzaki et al., 2003; Y. Muraki et al., 1996; K. Sugawara et al., 1993). ICV is distributed worldwide and and multiple genetic lineages co-circulate globally (D. A. Buonagurio et al., 1986; A. C. Dykes et al., 1980; Y. Furuse et al., 2016; Y. Matsuzaki et al., 2000; Y. Matsuzaki et al., 2016; H. Nishimura et al., 1987).

1.2 Structure of influenza viruses

Influenza virus is enveloped, spherical. The out layer is a lipid membrane that is taken from the host plasma membrane when influenza virus budding from infected cells. Influenza A viruses possess eight segments that encodes for eleven proteins. Two "Spike" proteins hemagglutinin (HA) and neuraminidase (NA) are embedded into the lipid membrane of influenza A and B viruses. HA and NA determine the subtype of influenza A virus. There are 18 different hemagglutinin subtypes and 11 different neuraminidase subtypes (N. M. Bouvier and P. Palese, 2008; S. Tong et al., 2013; Y. Wu et al., 2014).

HA and NA are important proteins that induce host immune system to produce internalization antibodies against the glycoprotein to prevent influenza virus infection. In infected cells, HA gene express a single polypeptide chain HA0 with a length of approximately 560 amino acids. HA0 is cleaved into two subunits, HA1 and HA2, which linked with each other by disulphide bonds. Cleavage of HA0 into HA1 and HA2 is essential for viral envelope and endosome membrane fusion during virus infection. In influenza A and B virus the receptor-destroying activity is performed by the NA protein, which hydrolyzes the glycosidic bond between sialic acid and galactosyl residues, help newly made viral particle release from the plasma membrane.
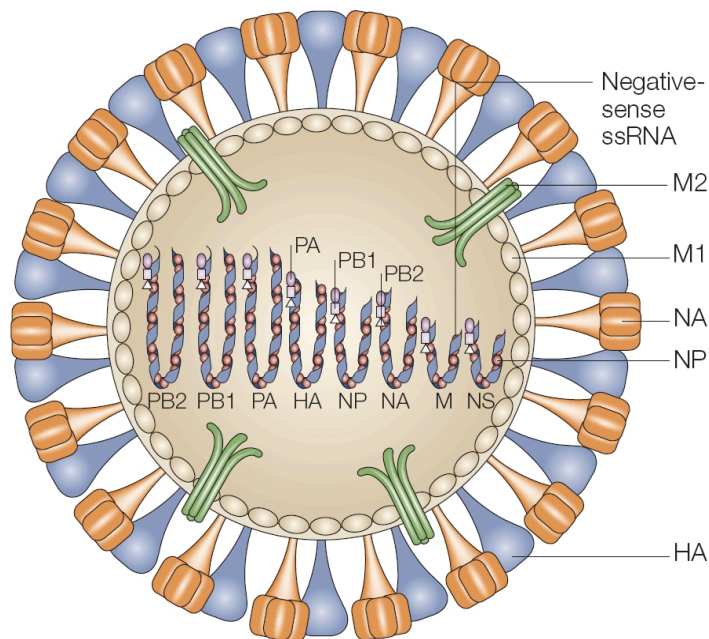
Figure 1.1. Diagrammatic illustration of the influenza A virion (T. Horimoto and Y. Kawaoka, 2005)

Ribonucleoproteins (vRNPs) complex is the core of influenza viruses. vRNPs are made up of viral genome RNAs wrapped up around multiple copies of nucleoprotein (NP). At the terminal ends of each viral segment is the three polymerase proteins (PB1, PB2 and PA) complex. A very small amount of nuclear export protein (NEP) are made up of vRNPs. M1 protein is underneath the viral lipid membrane, forms a matrix holding the viral vRNPs (D. P. Nayak et al., 2009; D. P. Nayak et al., 2004).

M2 is another protein that is inserted in the lipid membrane of influenza virions. M2 is a type III transmembrane protein that forms tetramers in the membrane (D. P. Nayak et al., 2009; C. Schroeder et al., 2005). M2 acts as an ion channel, which will be open during viral and endosomal membranes fusion. Opening the M2 ion channel acidifies the viral core and help the virion release vRNP into cytosol. In influenza A virus, alternative splicing of the M mRNA produces AM2 which forms a proton channel in the lipid membrane. The expression of IBV M2 protein is produced from the primary RNA transcript of M segment through a coupled termination and reinitiation mechanism via a UAAUG pentanucleotide motif near M1 stop codon (C. M. Horvath et al., 1990a). ICV M2 protein is a cleaved translation product of the full-length, colinear mRNA derived from influenza C virus RNA segment (A. Pekosz and R. A. Lamb, 1998).

The polymerase basic protein 1-F2 (PB1-F2) of influenza A virus is expressed from a second open reading frame (+1) of the PB1 gene (W. Chen et al., 2001; R. A. Lamb and M. Takeda, 2001). PB1-F2 localizes at inner and outer mitochondrial membranes. PB1-

F2 has been shown to play a role in mitochondrial morphology, cell apoptosis and up-regulate polymerase activity (J. S. Gibbs et al., 2003; H. Yamada et al., 2004; D. Zamarin et al., 2005).

There are some differences between type C influenza virus and type A and B viruses. Genome of ICV has 7 segments, whereas genome of IAV and IBV has 8 segments. In ICV, a hemagglutinin-esterase-fusion (HEF) protein has receptor binding, receptor destroying (acetylesterase), and membrane fusion activities, whereas in IAV and IBV, separate HA and NA proteins perform these functions in a cooperative fashion. Another difference is that ICV uses Neu5,9Ac2 (N-acetyl-9-O-acetylneuraminic acid), while IAV and IBV utilizes Neu5AC (N-acetylneuraminic acid) for viral entry (G. Herrler et al., 1988). Both neuraminic acid derivatives are the terminal sugars in carbohydrate chains attached to glycolipids or glycoproteins located at the cellular surface. Subtypes of influenza A virus HA discriminate between an α2-6 and α2-3 linkage to the second galactosyl residue, a property that (partially) explains species specificity. For example, human IAV prefer bind to α2-6 linkage, avian IAV recognize α2-3 linkage, and swine IAV can binds to both α2-6 and α2-3 linkages. This explains swine is a mixing vessel of human and avian influenza virus. ICV's HEF recognizes Neu5,9Ac2 independent of its linkage to the next sugar. HEF in ICV has also esterase activity that cleaves acetyl from the C9 position (G. Herrler et al., 1985). In influenza A and B virus the receptor-destroying activity is performed by the NA protein, which hydrolyzes the glycosidic bond between sialic acid and galactosyl residues.

1.3 Replication cycle of influenza viruses

The life cycle of influenza virus can be divided into five stages: receptor-mediated endocytosis; release viral ribonucleoprotein (vRNP) complexes; vRNA replication and transcription, protein synthesis; transport of viral proteins and vRNA to plasma membrane; virus assembly and release.
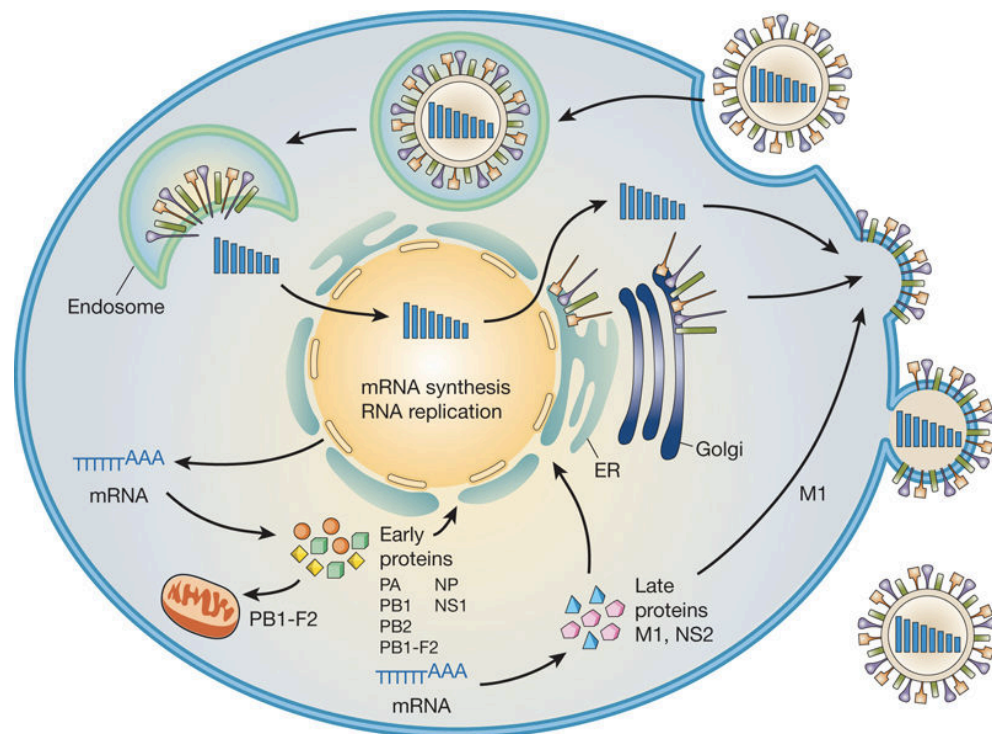
Figure 1.2. A schematic diagram of the influenza virus life cycle (G. Neumann et al., 2009).

The entry of influenza virus into cells begins with HA (IAV or IBV) or HEF (ICV) protein recognize its receptor sialic acid. HA1 and HEF1 contain receptor binding domain, while HA2 and HEF2 contain fusion peptide. After HA or HEF bindings to its

receptor, virus enter the host cells by receptor-mediated endocytosis and travel to late endosome. The PH is around 5.5 at late endosome. The low PH triggers HA2 or HEF2 go through conformational change and expose their fusion peptide. The fusion peptide insert itself into endosome membrane and mediate viral and endosome membrane fusion. Fusion is followed by release of viral core into cytosol where the M1 proteins are disassemble and vRNPs are released. The low PH also opens M2 proton channel. Opening M2 ion channel acidifies the viral core and weakens the interaction of M1 layer with vRNPs.

The vRNPs then import into nucleus where replication and transcription take place. vRNPs contains PB2, PB1, PA (or P3) and NP proteins. The nuclear localization signals (NLS) of these proteins bind to the cellular import proteins then import vRNPs into nucleus. The replication and transcription of vRNA are by RNA dependent RNA polymerase (RdRp). RdRp is made up of three polymerase proteins: PB2, PB1 and PA (or P3). During replication, the negative sensed vRNA are transcribed into complementaty RNA (cRNA) which serves as a template to synthesize more vRNAs. Meanwhile, vRNA are transcribed into messenger RNA (mRNA). mRNA is exported into cytoplasm for translation. Newly synthesized polymerase, NP and NS1 are imported back into nucleus to help with vRNA replication and transcription.

Late in replication cycle, the M1 and NS2 proteins bind to newly synthesized vRNPs and help vRNP export from nucleus. The attachment of vRNPs to M1 proteins traiggers viral budding from host cells. Budding occurs at apical side of polarized epithelial cells. HA,

NA and M2 proteins are export on the membrane of infected cells. At the last step, NA cleaves sialic acid residues from glycoproteins and glycolipids to help progeny viral particle release from infected cells.

1.4 Defective interfering (DI) RNAs of influenza A virus

Defective interfering (DI) genomes are produced during the replication of almost all viruses. DI genomes are truncated at highly variable positions, with either central or a terminal deletion depending on the virus species (N. J. Dimmock and A. J. Easton, 2014, 2015). Early studies showed DI viruses were produced by high multiplicity-of-infection (MOI) passages and contained a mixture of heterologous DI genome sequences (S. D. Duhaut and N. J. Dimmock, 1998; P. A. Jennings et al., 1983). The underlying mechanism for DI RNA formation has not been fully understood yet. However, there is evidence indicating that a viral RNA polymerase slippage-based faulty replication process is a likely mechanism (N. J. Dimmock and A. J. Easton, 2014). DI viruses contain the truncated DI genome with the capacity to interfere with infectious virus production. DI genomes has been detected in patients infected with different viruses, indicating DIs may play a role in transmission and pathogenesis (D. Li et al., 2011; T. T. Yuan et al., 1998).
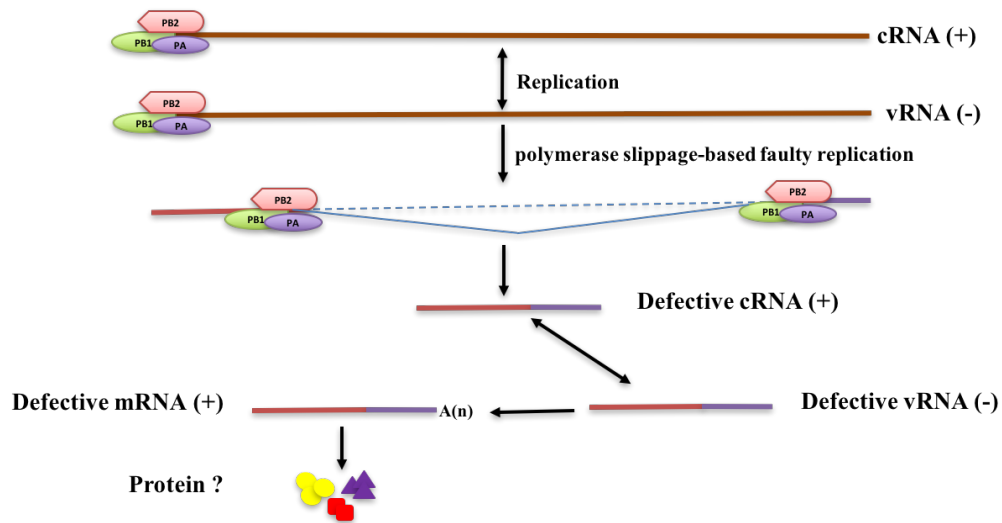
Figure 1.3. The production mechanism of defective interfering RNA in influenza A virus.

The existence of DI genomes in IBV has been reported in 1954 (P. Von Magnus, 1954) and DI genomes are well known in IAV. So far, there are over 50 different species of DI RNAs have been reported during IAV replication (N. J. Dimmock and A. J. Easton, 2015; S. Noble and N. J. Dimmock, 1995). These DI RNAs are small in size and viral RNA (vRNA) sensed. They possess similar or identical terminal sequences to their parental segments but lack a major internal region. DI genomes inhibit with IAV replication by competition with the parental full-length segments in replication, transcription, and genome packaging (T. Odagiri and M. Tashiro, 1997). There is evidence showed that the panhandle structure and the packaging signals are related in the segment-specific inhibition in viral replication (A. Baum et al., 2010; T. Frensing et al., 2014; J. M. Ngunjiri et al., 2013; M. Perez-Cidoncha et al., 2014). The DI genomes can be produced by any segment, but almost all DI RNAs originated from the three largest polymerase segments (N. J. Dimmock and A. J. Easton, 2014, 2015). DI RNAs compete with their parental full-length segments at two steps. First, when DI genomes are produced in the

nucleus, they engage the polymerase complexes so less full-length RNAs would be synthesized. Second, DI genomes can compete their parental segments and be packaged into budding virions to produce DI particles. DI particles are non-infectious. Their replication needs co-infection with wildtype helper viruses. In addition, IAV-derived DI genomes can potentially induce innate immune response (A. Baum et al., 2010; T. Frensing et al., 2014; J. M. Ngunjiri et al., 2013; M. Perez-Cidoncha et al., 2014) and protect cells or animals from infections of IAV, IBV, and other respiratory viruses (N. J. Dimmock et al., 2008; A. J. Easton et al., 2011; P. D. Scott et al., 2011).
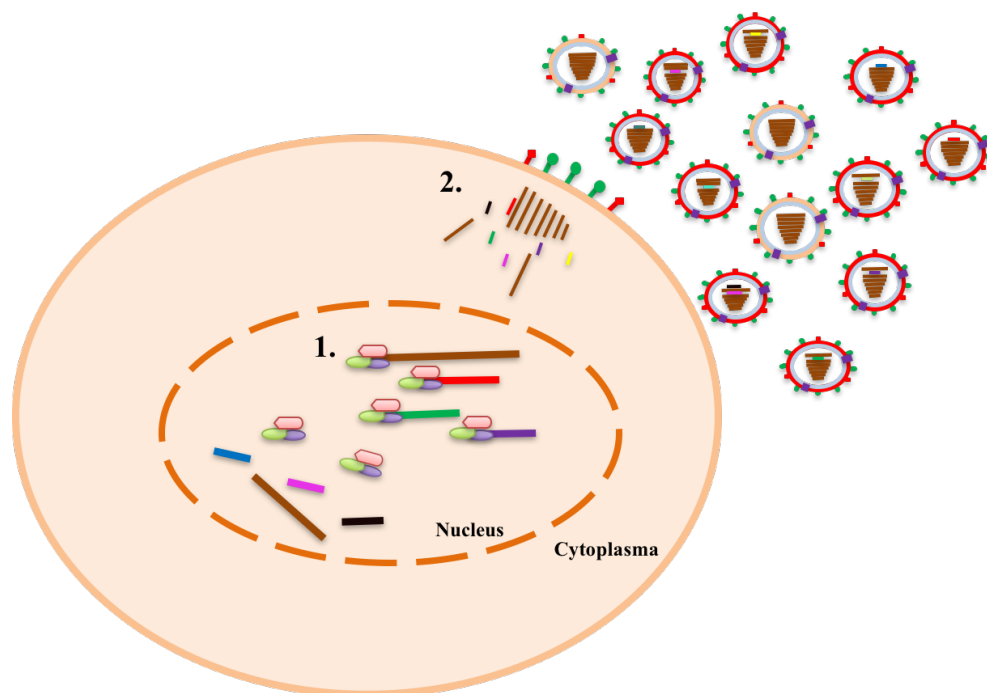
Figure 1.4. Defective interfering RNAs interfere with wild-type influenza virus replication. The competition of DI genomes with the parental full-length segments in replication and transcription by engaging polymerase in nucleus; in packaging by out competing their corresponding full-length segment.

1.5 Current advances on influenza D virus (IDV)

IDV is a new genus of the family Orthomyxoviridae, initially isolated from a 15-week old swine with influenza-like symptoms from Oklahoma in 2011 (B. M. Hause et al., 2013). This virus possesses seven RNA segments and shows 50% protein sequence identity to ICV, similar to the level of divergence between IAV and IBV. Evolutional analysis showed IDV and ICV were diverged about 300 to 1200 years ago (Z. Sheng et al., 2014). ICV is the genetically closest to IDV, but had no cross reactivity to ICV antibodies and failed to undergo reassortment with ICV. Next generation sequencing (NGS) analysis revealed a novel M1 splicing strategy differed from that of ICV. Crystal structure of its HEF reveals an open receptor-binding cavity of HEF compared to that of ICV, which explains the observed broad cell tropism of IDV (H. Song et al., 2016). Subsequently, IDV was detected and isolated in cattle from United States, Europe (France, Italy) and Asia (China, Japan) (C. Chiapponi et al., 2016; E. A. Collin et al., 2015; M. F. Ducatez et al., 2015; E. Foni et al., 2017; B. M. Hause et al., 2014; W. M. Jiang et al., 2014; S. Murakami et al., 2016). Besides swine and cattle, IDV was also detected in small ruminants (sheep and goats) and horses from recent study (M. Quast et al., 2015). Cattle have been proposed as the reservoirs for IDV (B. M. Hause et al., 2014) and two distinct genetic and antigenic IDV lineages were co-circulating in cattle in United States. IAV and IBV are more closely related to each other than to ICV and IDV. ICV is genetically closest to IDV, with only 50% amino acids homology between each other (B. M. Hause et al., 2013).

1.5.1 Epidemiology and pathogenesis of influenza D virus

Influenza viruses can cause epidemics and pandemics by causing severe respiratory disease spreading rapidly among all age groups. IAV can infect a wide variety of animals, including human, horses, pigs, ferrets and birds (J. K. Taubenberger and J. C. Kash, 2010). Both IAV and IBV can cause substantial morbidity and mortality during seasonal epidemics. IAV have caused several pandemics so far, such as Spanish influenza (H1N1), outbreaks of avian H5N1 viruses in Hong Kong and the most recent H1N1 influenza virus pandemic in 2009 (G. Neumann and Y. Kawaoka, 2015). IBV doesn't have the potential to cause serious pandemics due to the lack of an animal reservoir. ICV co-circulating with IAV and IBV cause local epidemics (A. Anton et al., 2011; Y. Matsuzaki et al., 2007). ICV have been isolated from pigs, dogs and human (Y. J. Guo et al., 1983; J. C. Manuguerra and C. Hannoun, 1992). Genetic reassorting among ICV occurs frequently in nature and most of the circulating ICV are newly emerged reassortants (Y. Matsuzaki et al., 2003). Co-circulation of ICV belonging to different genetic and antigenic lineages in a small community has been documented (Y. Matsuzaki et al., 1994). ICV usually cause upper tract infection with mild clinical symptoms, such as rhinorrhea, cough and fever (E. Greenbaum et al., 1998). Severe lower respiratory infections to infants and young children caused by ICV have also been reported (S. Gouarin et al., 2008). ICV co-circulate with IAV and IBV and it is hard to distinguish the clinical differentiation among different types of influenza virus infection (Y. Matsuzaki et al., 2007). ICV can cause natural infection in pigs and transmit from pigs to pigs by direct

contact. Pigs inoculated with ICV had mild respiratory clinical signs, consisting of breathing difficulty and nasal secretion increase (Y. J. Guo et al., 1983).

IDV has been widely geographic distributed since the first IDV was reported and isolated from pig exhibiting severe influenza-like illness in Oklahoma in 2011(B. M. Hause et al., 2013). Detection and isolation of IDVs have been reported in diseased calves in United Stated, China, Japan, France and Italy (C. Chiapponi et al., 2016; E. A. Collin et al., 2015; M. F. Ducatez et al., 2015; E. Foni et al., 2017; B. M. Hause et al., 2014; W. M. Jiang et al., 2014; T. F. Ng et al., 2015). Bovine has been proposed as the natural reservoir for IDV. Phylogenetic analysis of IDV genome sequence revealed two distinct genetic and antigenic lineages represented by D/swine/Oklahoma/1334/2011 (D/OK) and D/bovine/Oklahoma/660/2013 (D/660), which co-circulating and frequently reassort with each another (E. A. Collin et al., 2015).

Serological investigation showed that 9.5% of surveyed swine possessed antibodies specific to IDV indicating that this previously unidentified virus circulates in U.S. swine (B. M. Hause et al., 2013). A recent study showed 0.6% swine sera samples collected in 2009 in Italy were positive to IDV while this number increased to 11.7% in 2015 (E. Foni et al., 2017). It indicated IDV are circulating in Italy swine herds. Interestingly, 36.8% nasal swabs from pigs in Guangzhou, China were tested positive to IDV (S. L. Zhai et al., 2017). It suggested IDV could transmit in pigs in South China. Seroprevalence of IDV was investigated in small ruminants, horses and poultry. 5.2% of sheep sera and 8.8% of goat sera were tested positive to IDV, demonstrated sheep and goat are likely susceptible

to IDV infection (M. Quast et al., 2015). Horse serum has also been detected with positive antibodies against IDV (unpublished data). No IDV antibodies have been detected in poultry including chicken and turkey so far (M. Quast et al., 2015). The prevalence of IDV in cattle especially in calves is higher. Cattle sera collected in 2003-2004 from 40 selected farms in Nebraska, United States were tested positive for IDV (J. Luo et al., 2017). In 2013, 18% of bovine sera from six states in U.S. were tested positive to IDV using real time reverse transcription PCR and HI assay (B. M. Hause et al., 2014). Over 90% of newborn calf sera collected in 2014 had high levels of maternal antibodies against IDV in Nebraska cattle population (J. Luo et al., 2017). It seems calves under six-month old are more permissive for IDV infection and transmission. This study showed IDV has been existing as early as 2003 and has been circulating in Nebraska cattle herds. A nationwide distribution of IDV in cattle herds in Japan has been reported. An average 30% sera collected from 2010 to 2015 were positive in cattle, but the positivity rates varied among regions (T. Horimoto et al., 2016). More animal species (bufflo, dairy cattle, yellow cattle) has been reported positive to IDV in South China (S. L. Zhai et al., 2017). In human, serological studies found 1.3% of human sera were positive to IDV and titers are low (20-40, range) (B. M. Hause et al., 2013). Similar results were obtained from recent serological survey in Eastern South Dakota. 1.1 % of human sera were cross react to IDV (unpublished data). Interestingly, the positive samples were from individuals exposed and non-exposed to livestock. This indicated IDV transmission may occur without exposed to virus and virus infected animals. In contrast with those observations, another study reported 91% seropositive among 35 cattle-exposed workers and 18% seropositive for non-cattle-exposed individuals (S. K. White et al., 2016). This indicates

that IDV has the potential transmission from cattle, the natural viral reservoir to human, causing zoonotic risk among human population. However, ICV antibody cross-activity with IDV could be a part reason that cause the high sera-positivity rate in this study.
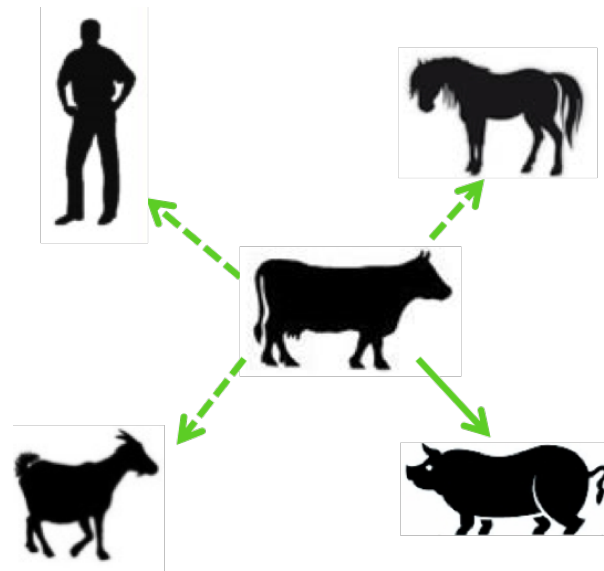
Figure 1.5. Transmission model of influenza D virus. IDV utilizes bovine as natural reservoir and is transmitted from bovine to swine as indicated by the solid line. Antibodies specific for IDV were detected from small ruminants (Sheep and goat), horses and human. Since no influenza D virus has been isolated from these species, the potential transmission of IDV from bovine to small ruminants, horses and human are indicated by dotted line.

Animals infected with IDV usually exhibit no clinical signs or mild clinical respiratory symptoms. In some cases, IDV could cause severe pneumonia in calves which may be

involved with secondary bacterial infection (E. A. Collin et al., 2015). From experimental IDV infection in cattle, virus was detected in both upper and lower respiratory tracts. IDV was efficiently transmitted in contacted cattle but not from cattle to ferret. IDV infected cattle exhibited mild respiratory symptoms and tracheal inflammation (L. Ferguson et al., 2016).

IDV can readily infect and transmit by direct contact with both ferrets and pigs (B. M. Hause et al., 2013). Infection was limited to the upper respiratory tract (nasal swabs, in nasal turbinates) and no virus was detected in trachea and lung tissue. Since ferrets are the best surrogate animal model for human with influenza infection (J. A. Belser et al., 2009), IDV infection could pose a potential threat to human health. In addition, IDV can infect guinea pigs and transmit by direct contact (C. Sreenivasan et al., 2015). Virus was detected in both upper and lower (lung) respiratory tracts but no clinical symptoms were observed. Interestingly, IDV infection can alter the structural integrity of the respiratory epithelium and as a result trigger a significant increase in neutrophils in the trachea of infected animals. Virus replication dynamics and transmission in in guinea pigs are in a good agreement with those findings obtained in native animals.

Ng *et al* detected IDV in calves with acute respiratory disease frequently but not in clinically healthy animals (T. F. Ng et al., 2015). Additionally, the common pathogens, bovine viral diarrhea virus (BVDV), bovine coronavirus (BoCV), bovine herpesvirus 1 (BHV-1) and bovine respiratory syncytial virus (BRSV), associated with bovine respirotary disease were not detected in IDV infected sick calves. This pathological effect

seems to suggest an etiological role of IDV in bovine respiratory disease complex (RDC). Frequent isolations or detections of IDVs from diseased cattle in Northern America (US and Mexico), Asia, and Europe with RDC apparently support this assumption. It is worth mentioning that in addition to swine and cattle, small ruminants (sheep and goats) and equine are probably additional hosts for IDV.

Taken together, IDV has the potential to infect humans, which is currently supported by experimental infections of both human surrogate models (ferret and guinea pig) as well as by serologic evidence of exposure to IDV among persons with occupational contact with cattle. The widespread distribution of IDV in agricultural animals (in goat, sheep, horses) and its interspecies movement pose a potential threat to human and agricultural health across the globe. Currently, there is no vaccine to protect animals from IDV infection. IDVs have antigenically evolved into multiple lineages. So the IDV vaccine research effort in the future needs to take such antigenic drift phenomena into consideration toward a universal vaccine targeting multiple animal species. Routine diagnosis of IDV infection in agricultural animals can be achieved by the standard reverse transcription polymerase chain reaction (RT-PCR) and hemagglutination inhibition assay (HI).

1.5.2   Biology of influenza D virus

IAV and IBV contain eight RNA segments with two spike proteins Hemagglutinin (HA) and neuraminidase (NA) on the surface utilizing 5-N-acetylneuraminic acid (Neu5Ac) as their primary receptor for virus entry. ICV and IDV possess seven RNA segments lacking

NA. The only glycoprotein Hemagglutinin-esterase-fusion (HEF) protein combines the functions of HA (receptor binding and membrane fusion activity) and NA (receptor destroying). However, the esterase cleaves the O-deacetylation of the N-acetyl-9-O-acetylneuraminic acid but not α-ketosidic linkage between the sialic (N-acetylneuraminic) acid and an adjacent sugar residue (J. N. Varghese and P. M. Colman, 1991). Both ICV and IDV utilize 5-N-acetyl-9-O-acetylneuraminic acid (Neu5, 9Ac2) as their receptor (H. Song et al., 2016).

In cell culture, IDV demonstrates a much broader cell tropism than ICV. In vitro cellular tropism study shows IDV can infect several cell lines such as swine testicle (ST), Madin-Darby canine kidney (MDCK), Green African monkey kidney (Marc-145), human rectal tumor (HRT-18G) and human lung adenocarcinoma (A549) cells, while ICV replicates only in ST and HRT cells (B. M. Hause et al., 2013). IDV infected swine testicle (ST) cells shows typical influenza virus cytopathic effects (CPE) (B. M. Hause et al., 2014; B. M. Hause et al., 2013). Structural modeling of HEF proteins of IDV showed a variable receptor binding site compared to ICV (B. M. Hause et al., 2013). It possesses an identical binding site for 9-O-acetyl group but a smaller 5-N-acetyl group binding pocket. This indicated ICD utilize the same receptor 9-O-acetylated sialic acids as ICV. Two variable amino acids in the 9-O-acetyl group binding site could influence binding efficiency and affinity to its receptor which may related to the observed broader celluar tropism of IDV. Structurally, despite the fact that the hemagglutinin-esterase-fusion glycoprotein (HEF) of both IDV and ICV possesses an extremely similar structural fold in communicating with the cellular receptor, IDV's HEF has an open receptor-binding

cavity that may accommodate diverse extended glycan moieties constituting the cellular receptor (H. Song et al., 2016). This structural flexibility may explain why the IDV has a broad cell tropism.
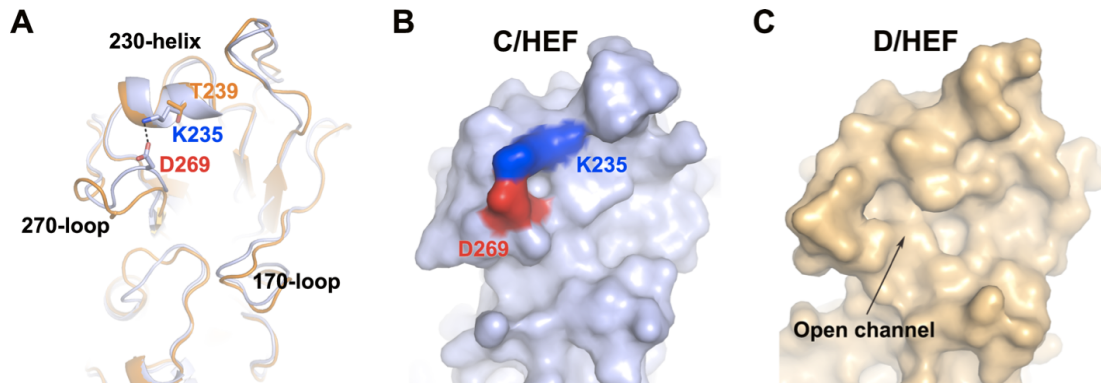
Figure 1.6. An open channel in IDV HEF receptor-binding domain. In ICV, the positively charged Lysine (K235) and negatively charged Aspartate (D269) forms a salt bridge, pulling 230-helix and 270-loop together (A and B). While in IDV, the amino acids at equivalent position can't form salt bridge, leaving an open cavity (C).

IDV has a wide range of temperature for its replication. IDV also differs from ICV in that its replication is permissive at 37 °C. ICV replicated well at optimal temperature 33°C while IDV replicate efficiently at both 33 °C and 37°C. Further analysis showed differences in the fusion domain of HEF proteins between ICV and IDV (Z. Sheng et al., 2014). The variable residues located at N- and C- terminal regions of fusion domain 1 and fusion domain 3. Since temperature affect fusion activity of HEF protein (E. Takashita et al., 2012), the variable fusion domains of the two viruses could contribute to the different replication temperature.

A recent study showed IDV is most stable in high temperature and low pH (J. Yu et al., 2017). After being treated at 53°C for 120 minutes, IDV still retained infectivity while other influenza viruses have lost infectivity completely. After being treated at pH 3.0 for 30min, IDV lost approximately 20% of its infectivity while other influenza viruses completely lost their infectivity. Chimeric IAV with HA and NA proteins replaced by HEF protein of IDV can tolerant higher temperature and low pH. HEF glycoprotein plays an important role in IDV thermal and acid stability.

1.5.3 Genomics and Evolution of Influenza D virus

IDV, like ICV, contains seven segments with negative polarity. Its overall amino acid sequence shares approximately 50% identity with that of ICV (B. M. Hause et al., 2013). It is most closely related to ICV with seven RNA segments, rather than to IAV and IBV containing eight segments. However, the distance between IDV and ICV is similar to the differences between IAV and IBV for the most of genomic segments (Z. Sheng et al., 2014).

Sequence identity was compared between D/OK and human ICV, ranging from 69-72% for Polymerase basic protein 1(PB1) to 29-33% for (nonstructural protein) NS segment, with about 50% overall identity (B. M. Hause et al., 2013). PB1 is the most highly conserved polymerase protein among influenza virus genus. Identity of PB1 amino acids sequence is 69%-72% between the two viruses and 39%-41% between D/OK and

IAV/IBV. PB2 and P3 protein sequences homology between D/OK and human ICV are 53% and 50%, respectively, which are lower than PB1 alignment. The Pka of polymerase acidic (PA) proteins of IAV/IBV is approximately 5.2. Polymerase protein 3 (P3) of ICV has a Pka of 7.2. The predicted Pka of D/OK P3 is 6.2, which is between those of IAV/IBV and ICV.

Hemagglutinin (HA) is the major glycoprotein of IAV/IBV, and hemagglutinin-esterase-fusion (HEF) protein of ICV is homologous to HA. The amino acids pairwise identity is 49% among IAV HA subtypes (S. Tong et al., 2012) and approximately 25-30% between IAV and IBV (B. M. Hause et al., 2013). The amino acid sequence of the D/OK HEF has 53% identity to ICV, similar to the observed homology among IAV HA subtypes. ICV and D/OK possess seven and six predicted N-linked glycosylation sites in their HEF proteins, respectively (Z. Sheng et al., 2014). D/OK lack one predicted N-linked glycosylation at receptor binding domain, implying different strategies for IDV to escape antibody-mediated neutralization.

NP and M proteins are highly conserved among the members of each genus of influenza virus, which share 85% homology (B. M. Hause et al., 2013). The homologies of NP and M proteins are only 20-30% between different types of influenza viruses (M. Yamashita et al., 1988). The homology of D/OK and ICV NP proteins is 38-41%. The amino acids sequence of D/OK precursor protein P42 has 38% identity with its counterpart of ICV. The non-structural proteins NS1 of D/OK and human ICV have only 29-33% identity, similar to 22% homology between IAV and IBV NS1 proteins.

A major difference among influenza viruses occurs in the matrix (M) segment with regard to the expression of M1 and M2 proteins (P. Palese and M. Shaw, 2007). IDV utilizes a novel coding strategy for expression of M1 and M2 proteins, differs from all other influenza virus types. ICV and IDV utilize similar splicing strategies to produce M1 protein. The unspliced mRNA translate a precursor protein P42 which is proteolytically cleaved into the M1'and CM2 and proteins. M1 proteins of both ICV and IDV are translated from splicing events. In ICV, the splicing introduces a termination codon to M1 transcript. While in IDV, it adds a second exon encoding four residues to the primary M1 transcript (B. M. Hause et al., 2014).

The 3' and 5'-terminal ends of noncoding (NC) regions of each segment were highly conserved in D/OK. The conserved terminal sequences show partial and inverted complementarity which enables single-strand RNAs to form "panhandle" structures (U. Desselberger et al., 1980). "Panhandle" structures are required for mRNA synthesis and vRNA packaging (M. T. Hsu et al., 1987). The 3' and 5'-terminal sequences are similar to those of ICV with only one nucleotide change at position 5 ("A" in D/OK and "U" in ICV) from 3'terminal.

Bovine IDVs and D/OK, the two representative strains of IDV lineages, had greater than 96% identity in all segments (B. M. Hause et al., 2014). PB1 (98.9 to 99.1% identity) and NS (98.8 to 99.2% identity) segments have highest homology, while HEF (96.7 to 99.0%

identity) and P42 (96.9 to 99.2% identity) are most diverged. The 5' and 3' noncoding regions for each segment are quite conserved between of bovine and swine IDVs. Phylogenetic analysis showed bovine and swine IDVs clustered together and most closely related to ICV. This indicates IDV diverged from ICV after they diverged from IAV and IBV (B. M. Hause et al., 2014). Evolution evidence showed the divergence time of IDV and ICV from the most recent common ancestor ranges from 334 years for PB1 to 1299 years for HE (Z. Sheng et al., 2014).

1.5.4   Hemagglutinin-esterase-fusion protein of influenza D virus

Hemagglutinin-esterase-fusion(HEF) protein is the only glycoprotein on the surface of ICV and IDV. It contains the receptor binding domain (R), esterase domain(E) and fusion(F) domain. Crystal structure of IDV HEF revealed E1, E' and E2 subdomains of esterase domain and F1, F2 and F3 subdomains in fusion domain. The overall HEF structures and the individual subdomains are quite similar between IDV and ICV. R domains show 46.3% sequence identity. Both structural modeling and crystal structure of HEFs of IDV and ICV reveal a broader receptor-binding site (B. M. Hause et al., 2013; H. Song et al., 2016). Crystal structure of IDV HEF further reveals an open cavity formed by 230-helix and 270-loop at the receptor-binding site. While in ICV, a salt bridge forming by K235 and D269 between the two secondary structures causes a closed cavity. In IDV HEF, the two amino acids are replaced by T239 and A273, which cannot form salt bridge interaction. Other residues that form the bottom of the receptor-binding site are similar between the two species.

Glycan array showed that the HEF protein of IDV binds to both 5-N-Ac and 5-N-Gc sialic acids carrying 9-O-acetyl group. This result indicates that the open cavity of receptor-binding site residing in the HEF protein can tolerate different modifications at the C5 position (H. Song et al., 2016). It is consistent with the observed broader cellular tropism of IDV (H. Song et al., 2016). Crystal structure study recently revealed that the major interaction between HEF and the receptor analogue is derived from 9-O-Ac sialic acid. A hydrogen bond is formed between the interaction of acetyl carbonyl oxygen of 9-O-Ac sialic acid and the hydroxyl group of Y231 which is Y227 in ICV. Two hydrogen bonds are absent in IDV due to the conserved residue Y127 in IAV and IBV changed to F127, which is same with IBV (H. Song et al., 2016; Q. Wang et al., 2007). However, two hydrogen bonds formed between carbonyl group of C4 and T171 ensures that the 9-O-acetyl group has similar orientation with binding to ICV. Additional hydrogen bonds formed between receptor analogue and IDV HEF are one from 5-N-acetyl group with A172 and two from the carboxyl group of C1 with S173.

Esterase domains are most conserved. The E1, E' and E2 subdomains share 66.7%, 68.8% and 56.6% sequence identity between ICV and IDV (H. Song et al., 2016). Both HEF proteins form conserved sialate-9-O-acetylesterase catalytic triad with S71/H369/D365 for ICV and S73/H375/D372 for IDV. The conserved receptor-destroying pocket indicated IDV utilizes 9-O-acetyl sialic acid as the cellular receptor for virus entry (B. M. Hause et al., 2013). It is reported that esterase domains of HEF proteins are conserved in some Nidovirus, such as bovine coronavirus (BCoV), porcine

torovirus (PToV) and bovine torovirus (BToV) as well (M. A. Langereis et al., 2009; H. Song et al., 2016).

1.6 Summary

Influenza viruses pose a constant threat to human and animal health. Among four types of influenza viruses, IAV and IBV infect humans and can cause severe respiratory diseases in annual influenza epidemics, which can result up to 20,000 deaths in U.S. alone. ICV is a ubiquitous childhood pathogen typically causing mild respiratory symptoms, in rare cases causing severe respiratory disease in infants (S. Katagiri et al., 1983; Y. Matsuzaki et al., 2006). IDV with bovines as a primary reservoir and amplification host can periodically spillover to humans and other agriculture animals such as swine. Better understanding of the replication mechanism, pathogenesis, and antigenic evolution of influenza viruses will help develop effective vaccines and control strategies against this important group of respiratory pathogens of humans and animals.

BIBLIOGRAPHY

(1980). A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bulletin of the World Health Organization *58*, 585-591.

Anton, A., Marcos, M. A., Codoner, F. M., de Molina, P., Martinez, A., Cardenosa, N., Godoy, P., Torner, N., Martinez, M. J., Ramon, S., Tudo, G., Isanta, R., Gonzalo, V., de Anta, M. T., and Pumarola, T. (2011). Influenza C virus surveillance during the first influenza A (H1N1) 2009 pandemic wave in Catalonia, Spain. Diagnostic microbiology and infectious disease *69*, 419-427.

Barr, I. G., McCauley, J., Cox, N., Daniels, R., Engelhardt, O. G., Fukuda, K., Grohmann, G., Hay, A., Kelso, A., Klimov, A., Odagiri, T., Smith, D., Russell, C., Tashiro, M., Webby, R., Wood, J., Ye, Z., Zhang, W., and Writing Committee of the World Health Organization Consultation on Northern Hemisphere Influenza Vaccine Composition, for (2010). Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009-2010 northern hemisphere season. Vaccine *28*, 1156-1167.

Baum, A., Sachidanandam, R., and Garcia-Sastre, A. (2010). Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. Proceedings of the National Academy of Sciences of the United States of America *107*, 16303-16308.

Bean, W. J., Kawaoka, Y., Wood, J. M., Pearson, J. E., and Webster, R. G. (1985). Characterization of virulent and avirulent A/chicken/Pennsylvania/83 influenza A viruses: potential role of defective interfering RNAs in nature. J Virol *54*, 151-160.

Belser, J. A., Szretter, K. J., Katz, J. M., and Tumpey, T. M. (2009). Use of animal models to understand the pandemic potential of highly pathogenic avian influenza viruses. Advances in virus research *73*, 55-97.

Bouvier, N. M., and Palese, P. (2008). The biology of influenza viruses. Vaccine *26 Suppl 4*, D49-53.

Buonagurio, D. A., Nakada, S., Fitch, W. M., and Palese, P. (1986). Epidemiology of influenza C virus in man: multiple evolutionary lineages and low rate of change. Virology *153*, 12-21.

Chambers, T. M., and Webster, R. G. (1987). Defective interfering virus associated with A/Chicken/Pennsylvania/83 influenza virus. J Virol *61*, 1517-1523.

Chen, W., Calvo, P. A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J. R., and Yewdell, J. W. (2001). A novel influenza A virus mitochondrial protein that induces cell death. Nat Med *7*, 1306-1312.

Chiapponi, C., Faccini, S., De Mattia, A., Baioni, L., Barbieri, I., Rosignoli, C., Nigrelli, A., and Foni, E. (2016). Detection of Influenza D Virus among Swine and Cattle, Italy. Emerging infectious diseases *22*, 352-354.

Collin, E. A., Sheng, Z., Lang, Y., Ma, W., Hause, B. M., and Li, F. (2015). Cocirculation of two distinct genetic and antigenic lineages of proposed influenza D virus in cattle. Journal of virology *89*, 1036-1042.

Desselberger, U., Racaniello, V. R., Zazra, J. J., and Palese, P. (1980). The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. Gene *8*, 315-328.

Dimmock, N. J., and Easton, A. J. (2014). Defective interfering influenza virus RNAs: time to reevaluate their clinical potential as broad-spectrum antivirals? Journal of virology *88*, 5217-5227.

Dimmock, N. J., and Easton, A. J. (2015). Cloned Defective Interfering Influenza RNA and a Possible Pan-Specific Treatment of Respiratory Virus Diseases. Viruses *7*, 3768-3788.

Dimmock, N. J., Rainsford, E. W., Scott, P. D., and Marriott, A. C. (2008). Influenza virus protecting RNA: an effective prophylactic and therapeutic antiviral. J Virol *82*, 8570-8578.

Dowdle, W. R., Galphin, J. C., Coleman, M. T., and Schild, G. C. (1974). A simple double immunodiffusion test for typing influenza viruses. Bulletin of the World Health Organization *51*, 213-215.

Dubois, Julia, Terrier, Olivier, and Rosa-Calatrava, Manuel (2014). Influenza Viruses and mRNA Splicing: Doing More with Less. mBio *5*.

Ducatez, M. F., Pelletier, C., and Meyer, G. (2015). Influenza D virus in cattle, France, 2011-2014. Emerging infectious diseases *21*, 368-371.

Duhaut, S. D., and Dimmock, N. J. (1998). Heterologous protection of mice from a lethal human H1N1 influenza A virus infection by H3N8 equine defective interfering virus: comparison of defective RNA sequences isolated from the DI inoculum and mouse lung. Virology *248*, 241-253.

Dykes, A. C., Cherry, J. D., and Nolan, C. E. (1980). A clinical, epidemiologic, serologic, and virologic study of influenza C virus infection. Archives of internal medicine *140*, 1295-1298.

Easton, A. J., Scott, P. D., Edworthy, N. L., Meng, B., Marriott, A. C., and Dimmock, N. J. (2011). A novel broad-spectrum treatment for respiratory virus infections: influenza-based defective interfering virus provides protection against pneumovirus infection in vivo. Vaccine *29*, 2777-2784.

Ferguson, L., Olivier, A. K., Genova, S., Epperson, W. B., Smith, D. R., Schneider, L., Barton, K., McCuan, K., Webby, R. J., and Wan, X. F. (2016). Pathogenesis of Influenza D Virus in Cattle. Journal of virology *90*, 5636-5642.

Foni, E., Chiapponi, C., Baioni, L., Zanni, I., Merenda, M., Rosignoli, C., Kyriakis, C. S., Luini, M. V., Mandola, M. L., Bolzoni, L., Nigrelli, A. D., and Faccini, S. (2017). Influenza D in Italy: towards a better understanding of an emerging viral infection in swine. Sci Rep *7*, 11660.

Frensing, T., Pflugmacher, A., Bachmann, M., Peschel, B., and Reichl, U. (2014). Impact of defective interfering particles on virus replication and antiviral host response in cell culture-based influenza vaccine production. Applied microbiology and biotechnology *98*, 8999-9008.

Furuse, Y., Matsuzaki, Y., Nishimura, H., and Oshitani, H. (2016). Analyses of Evolutionary Characteristics of the Hemagglutinin-Esterase Gene of Influenza C Virus during a Period of 68 Years Reveals Evolutionary Patterns Different from Influenza A and B Viruses. Viruses *8*.

Gibbs, J. S., Malide, D., Hornung, F., Bennink, J. R., and Yewdell, J. W. (2003). The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function. Journal of virology *77*, 7214-7224.

Gouarin, S., Vabret, A., Dina, J., Petitjean, J., Brouard, J., Cuvillon-Nimal, D., and Freymuth, F. (2008). Study of influenza C virus infection in France. Journal of medical virology *80*, 1441-1446.

Greenbaum, E., Morag, A., and Zakay-Rones, Z. (1998). Isolation of influenza C virus during an outbreak of influenza A and B viruses. Journal of clinical microbiology *36*, 1441-1442.

Guo, Y. J., Jin, F. G., Wang, P., Wang, M., and Zhu, J. M. (1983). Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus. The Journal of general virology *64 (Pt 1)*, 177-182.

Gutierrez-Pizarraya, A., Perez-Romero, P., Alvarez, R., Aydillo, T. A., Osorio-Gomez, G., Milara-Ibanez, C., Sanchez, M., Pachon, J., and Cordero, E. (2012). Unexpected severity of cases of influenza B infection in patients that required hospitalization during the first postpandemic wave. J Infect *65*, 423-430.

Hause, B. M., Collin, E. A., Liu, R., Huang, B., Sheng, Z., Lu, W., Wang, D., Nelson, E. A., and Li, F. (2014). Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. mBio *5*, e00031-00014.

Hause, B. M., Ducatez, M., Collin, E. A., Ran, Z., Liu, R., Sheng, Z., Armien, A., Kaplan, B., Chakravarty, S., Hoppe, A. D., Webby, R. J., Simonson, R. R., and Li, F. (2013). Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens *9*, e1003176.

Herrler, G., Durkop, I., Becht, H., and Klenk, H. D. (1988). The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. The Journal of general virology *69 (Pt 4)*, 839-846.

Herrler, G., Rott, R., Klenk, H. D., Muller, H. P., Shukla, A. K., and Schauer, R. (1985). The receptor-destroying enzyme of influenza C virus is neuraminate-O-acetylesterase. The EMBO journal *4*, 1503-1506.

Horimoto, T., Hiono, T., Mekata, H., Odagiri, T., Lei, Z., Kobayashi, T., Norimine, J., Inoshima, Y., Hikono, H., Murakami, K., Sato, R., Murakami, H., Sakaguchi, M., Ishii, K., Ando, T., Otomaru, K., Ozawa, M., Sakoda, Y., and Murakami, S. (2016). Nationwide Distribution of Bovine Influenza D Virus Infection in Japan. PLoS One *11*, e0163828.

Horimoto, T., and Kawaoka, Y. (2005). Influenza: lessons from past pandemics, warnings from current incidents. Nat Rev Microbiol *3*, 591-600.

Horvath, C. M., Williams, M. A., and Lamb, R. A. (1990a). Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. The EMBO journal *9*, 2639-2647.

Horvath, C. M., Williams, M. A., and Lamb, R. A. (1990b). Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. The EMBO Journal *9*, 2639-2647.

Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M., and Palese, P. (1987). Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. Proceedings of the National Academy of Sciences of the United States of America *84*, 8140-8144.

Jennings, P. A., Finch, J. T., Winter, G., and Robertson, J. S. (1983). Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? Cell *34*, 619-627.

Jiang, W. M., Wang, S. C., Peng, C., Yu, J. M., Zhuang, Q. Y., Hou, G. Y., Liu, S., Li, J. P., and Chen, J. M. (2014). Identification of a potential novel type of influenza virus in Bovine in China. Virus genes *49*, 493-496.

Katagiri, S., Ohizumi, A., and Homma, M. (1983). An outbreak of type C influenza in a children's home. The Journal of infectious diseases *148*, 51-56.

Lamb, R A, Lai, C J, and Choppin, P W (1981). Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proceedings of the National Academy of Sciences *78*, 4170-4174.

Lamb, R. A., and Takeda, M. (2001). Death by influenza virus protein. Nat Med *7*, 1286-1288.

Langereis, M. A., Zeng, Q., Gerwig, G. J., Frey, B., von Itzstein, M., Kamerling, J. P., de Groot, R. J., and Huizinga, E. G. (2009). Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases. Proceedings of the National Academy of Sciences of the United States of America *106*, 15897-15902.

Li, D., Lott, W. B., Lowry, K., Jones, A., Thu, H. M., and Aaskov, J. (2011). Defective interfering viral particles in acute dengue infections. PLoS One *6*, e19447.

Luo, J., Ferguson, L., Smith, D. R., Woolums, A. R., Epperson, W. B., and Wan, X. F. (2017). Serological evidence for high prevalence of Influenza D Viruses in Cattle, Nebraska, United States, 2003-2004. Virology *501*, 88-91.

Manuguerra, J. C., and Hannoun, C. (1992). Natural infection of dogs by influenza C virus. Research in virology *143*, 199-204.

Matsuzaki, Y., Abiko, C., Mizuta, K., Sugawara, K., Takashita, E., Muraki, Y., Suzuki, H., Mikawa, M., Shimada, S., Sato, K., Kuzuya, M., Takao, S., Wakatsuki, K., Itagaki,

T., Hongo, S., and Nishimura, H. (2007). A nationwide epidemic of influenza C virus infection in Japan in 2004. Journal of clinical microbiology *45*, 783-788.

Matsuzaki, Y., Katsushima, N., Nagai, Y., Shoji, M., Itagaki, T., Sakamoto, M., Kitaoka, S., Mizuta, K., and Nishimura, H. (2006). Clinical features of influenza C virus infection in children. The Journal of infectious diseases *193*, 1229-1235.

Matsuzaki, Y., Mizuta, K., Kimura, H., Sugawara, K., Tsuchiya, E., Suzuki, H., Hongo, S., and Nakamura, K. (2000). Characterization of antigenically unique influenza C virus strains isolated in Yamagata and Sendai cities, Japan, during 1992-1993. The Journal of general virology *81*, 1447-1452.

Matsuzaki, Y., Mizuta, K., Sugawara, K., Tsuchiya, E., Muraki, Y., Hongo, S., Suzuki, H., and Nishimura, H. (2003). Frequent reassortment among influenza C viruses. Journal of virology *77*, 871-881.

Matsuzaki, Y., Muraki, Y., Sugawara, K., Hongo, S., Nishimura, H., Kitame, F., Katsushima, N., Numazaki, Y., and Nakamura, K. (1994). Cocirculation of two distinct groups of influenza C virus in Yamagata City, Japan. Virology *202*, 796-802.

Matsuzaki, Y., Sugawara, K., Furuse, Y., Shimotai, Y., Hongo, S., Oshitani, H., Mizuta, K., and Nishimura, H. (2016). Genetic Lineage and Reassortment of Influenza C Viruses Circulating between 1947 and 2014. Journal of virology *90*, 8251-8265.

McCullers, J. A., Saito, T., and Iverson, A. R. (2004). Multiple genotypes of influenza B virus circulated between 1979 and 2003. Journal of virology *78*, 12817-12828.

Monto, A. S. (2008). Epidemiology of influenza. Vaccine *26 Suppl 4*, D45-48.

Moriuchi, H., Katsushima, N., Nishimura, H., Nakamura, K., and Numazaki, Y. (1991). Community-acquired influenza C virus infection in children. The Journal of pediatrics *118*, 235-238.

Murakami, S., Endoh, M., Kobayashi, T., Takenaka-Uema, A., Chambers, J. K., Uchida, K., Nishihara, M., Hause, B., and Horimoto, T. (2016). Influenza D Virus Infection in Herd of Cattle, Japan. Emerging infectious diseases *22*, 1517-1519.

Muraki, Y., Hongo, S., Sugawara, K., Kitame, F., and Nakamura, K. (1996). Evolution of the haemagglutinin-esterase gene of influenza C virus. The Journal of general virology *77 (Pt 4)*, 673-679.

Nayak, D. P., Balogun, R. A., Yamada, H., Zhou, Z. H., and Barman, S. (2009). Influenza virus morphogenesis and budding. Virus research *143*, 147-161.

Nayak, D. P., Hui, E. K., and Barman, S. (2004). Assembly and budding of influenza virus. Virus research *106*, 147-165.

Neumann, G., and Kawaoka, Y. (2015). Transmission of influenza A viruses. Virology *479-480*, 234-246.

Neumann, G., Noda, T., and Kawaoka, Y. (2009). Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature *459*, 931-939.

Ng, T. F., Kondov, N. O., Deng, X., Van Eenennaam, A., Neibergs, H. L., and Delwart, E. (2015). A metagenomics and case-control study to identify viruses associated with bovine respiratory disease. Journal of virology *89*, 5340-5349.

Ngunjiri, J. M., Buchek, G. M., Mohni, K. N., Sekellick, M. J., and Marcus, P. I. (2013). Influenza virus subpopulations: exchange of lethal H5N1 virus NS for H1N1 virus NS triggers de novo generation of defective-interfering particles and enhances interferon-

inducing particle efficiency. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research *33*, 99-107.

Nishimura, H., Sugawara, K., Kitame, F., Nakamura, K., and Sasaki, H. (1987). Prevalence of the antibody to influenza C virus in a northern Luzon Highland Village, Philippines. Microbiology and immunology *31*, 1137-1143.

Noble, S., and Dimmock, N. J. (1995). Characterization of Putative Defective Interfering (DI) A/WSN RNAs Isolated from the Lungs of Mice Protected from an Otherwise Lethal Respiratory Infection with Influenza Virus A/WSN (H1N1): A Subset of the Inoculum DI RNAs. Virology *210*, 9-19.

Odagiri, T., and Tashiro, M. (1997). Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step. J Virol *71*, 2138-2145.

Osterhaus, A. D., Rimmelzwaan, G. F., Martina, B. E., Bestebroer, T. M., and Fouchier, R. A. (2000). Influenza B virus in seals. Science *288*, 1051-1053.

Palese, P., and Shaw, M. (2007). Orthomyxoviridae: The viruses and their replication. In Field's Virology, B.N. Fields, D.M. Knipe, and P.M. Howley, eds. (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins), pp. 1647-1689.

Paul Glezen, W., Schmier, J. K., Kuehn, C. M., Ryan, K. J., and Oxford, J. (2013). The burden of influenza B: a structured literature review. Am J Public Health *103*, e43-51.

Pekosz, A., and Lamb, R. A. (1998). Influenza C virus CM2 integral membrane glycoprotein is produced from a polypeptide precursor by cleavage of an internal signal

sequence. Proceedings of the National Academy of Sciences of the United States of America *95*, 13233-13238.

Perez-Cidoncha, M., Killip, M. J., Oliveros, J. C., Asensio, V. J., Fernandez, Y., Bengoechea, J. A., Randall, R. E., and Ortin, J. (2014). An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. J Virol *88*, 4632-4646.

Quast, M., Sreenivasan, C., Sexton, G., Nedland, H., Singrey, A., Fawcett, L., Miller, G., Lauer, D., Voss, S., Pollock, S., Cunha, C. W., Christopher-Hennings, J., Nelson, E., and Li, F. (2015). Serological evidence for the presence of influenza D virus in small ruminants. Veterinary microbiology *180*, 281-285.

Ran, Z., Shen, H., Lang, Y., Kolb, E. A., Turan, N., Zhu, L., Ma, J., Bawa, B., Liu, Q., Liu, H., Quast, M., Sexton, G., Krammer, F., Hause, B. M., Christopher-Hennings, J., Nelson, E. A., Richt, J., Li, F., and Ma, W. (2015). Domestic pigs are susceptible to infection with influenza B viruses. Journal of virology *89*, 4818-4826.

Robb, N. C., and Fodor, E. (2012). The accumulation of influenza A virus segment 7 spliced mRNAs is regulated by the NS1 protein. The Journal of general virology *93*, 113-118.

Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C. H., Meier-Ewert, H., Skehel, J. J., and Wiley, D. C. (1998). Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature *396*, 92-96.

Rota, P. A., Wallis, T. R., Harmon, M. W., Rota, J. S., Kendal, A. P., and Nerome, K. (1990). Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. Virology *175*, 59-68.

Saira, K., Lin, X., DePasse, J. V., Halpin, R., Twaddle, A., Stockwell, T., Angus, B., Cozzi-Lepri, A., Delfino, M., Dugan, V., Dwyer, D. E., Freiberg, M., Horban, A., Losso, M., Lynfield, R., Wentworth, D. N., Holmes, E. C., Davey, R., Wentworth, D. E., and Ghedin, E. (2013). Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. J Virol *87*, 8064-8074.

Schroeder, C., Heider, H., Moncke-Buchner, E., and Lin, T. I. (2005). The influenza virus ion channel and maturation cofactor M2 is a cholesterol-binding protein. Eur Biophys J *34*, 52-66.

Scott, P. D., Meng, B., Marriott, A. C., Easton, A. J., and Dimmock, N. J. (2011). Defective interfering influenza A virus protects in vivo against disease caused by a heterologous influenza B virus. The Journal of general virology *92*, 2122-2132.

Sheng, Z., Ran, Z., Wang, D., Hoppe, A. D., Simonson, R., Chakravarty, S., Hause, B. M., and Li, F. (2014). Genomic and evolutionary characterization of a novel influenza-C-like virus from swine. Archives of virology *159*, 249-255.

Song, H., Qi, J., Khedri, Z., Diaz, S., Yu, H., Chen, X., Varki, A., Shi, Y., and Gao, G. F. (2016). An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. PLoS pathogens *12*, e1005411.

Sreenivasan, C., Thomas, M., Sheng, Z., Hause, B. M., Collin, E. A., Knudsen, D. E., Pillatzki, A., Nelson, E., Wang, D., Kaushik, R. S., and Li, F. (2015). Replication and Transmission of the Novel Bovine Influenza D Virus in a Guinea Pig Model. Journal of virology *89*, 11990-12001.

Sugawara, K., Nishimura, H., Hongo, S., Kitame, F., and Nakamura, K. (1991). Antigenic characterization of the nucleoprotein and matrix protein of influenza C virus with monoclonal antibodies. The Journal of general virology *72 (Pt 1)*, 103-109.

Sugawara, K., Nishimura, H., Hongo, S., Muraki, Y., Kitame, F., and Nakamura, K. (1993). Construction of an antigenic map of the haemagglutinin-esterase protein of influenza C virus. The Journal of general virology *74 (Pt 8)*, 1661-1666.

Takashita, E., Muraki, Y., Sugawara, K., Asao, H., Nishimura, H., Suzuki, K., Tsuji, T., Hongo, S., Ohara, Y., Kawaoka, Y., Ozawa, M., and Matsuzaki, Y. (2012). Intrinsic temperature sensitivity of influenza C virus hemagglutinin-esterase-fusion protein. Journal of virology *86*, 13108-13111.

Taubenberger, J. K., and Kash, J. C. (2010). Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe *7*, 440-451.

Tong, S., Li, Y., Rivailler, P., Conrardy, C., Castillo, D. A., Chen, L. M., Recuenco, S., Ellison, J. A., Davis, C. T., York, I. A., Turmelle, A. S., Moran, D., Rogers, S., Shi, M., Tao, Y., Weil, M. R., Tang, K., Rowe, L. A., Sammons, S., Xu, X., Frace, M., Lindblade, K. A., Cox, N. J., Anderson, L. J., Rupprecht, C. E., and Donis, R. O. (2012). A distinct lineage of influenza A virus from bats. Proceedings of the National Academy of Sciences of the United States of America *109*, 4269-4274.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L. M., Johnson, A., Tao, Y., Dreyfus, C., Yu, W., McBride, R., Carney, P. J., Gilbert, A. T., Chang, J., Guo, Z., Davis, C. T., Paulson, J. C., Stevens, J., Rupprecht, C. E., Holmes, E. C., Wilson, I. A., and Donis, R. O. (2013). New world bats harbor diverse influenza A viruses. PLoS pathogens *9*, e1003657.

Varghese, J. N., and Colman, P. M. (1991). Three-dimensional structure of the neuraminidase of influenza virus A/Tokyo/3/67 at 2.2 A resolution. Journal of molecular biology *221*, 473-486.

Von Magnus, P. (1954). Incomplete forms of influenza virus. Advances in virus research *2*, 59-79.

Wang, Q., Tian, X., Chen, X., and Ma, J. (2007). Structural basis for receptor specificity of influenza B virus hemagglutinin. Proceedings of the National Academy of Sciences of the United States of America *104*, 16874-16879.

White, S. K., Ma, W., McDaniel, C. J., Gray, G. C., and Lednicky, J. A. (2016). Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. J Clin Virol *81*, 31-33.

Wu, P., Goldstein, E., Ho, L. M., Yang, L., Nishiura, H., Wu, J. T., Ip, D. K., Chuang, S. K., Tsang, T., and Cowling, B. J. (2012). Excess mortality associated with influenza A and B virus in Hong Kong, 1998-2009. The Journal of infectious diseases *206*, 1862-1871.

Wu, Y., Wu, Y., Tefsen, B., Shi, Y., and Gao, G. F. (2014). Bat-derived influenza-like viruses H17N10 and H18N11. Trends in microbiology *22*, 183-191.

Yamada, H., Chounan, R., Higashi, Y., Kurihara, N., and Kido, H. (2004). Mitochondrial targeting sequence of the influenza A virus PB1-F2 protein and its function in mitochondria. FEBS Lett *578*, 331-336.

Yamashita, M., Krystal, M., and Palese, P. (1988). Evidence that the matrix protein of influenza C virus is coded for by a spliced mRNA. Journal of virology *62*, 3348-3355.

Yu, J., Hika, B., Liu, R., Sheng, Z., Hause, B. M., Li, F., and Wang, D. (2017). The Hemagglutinin-Esterase Fusion Glycoprotein Is a Primary Determinant of the Exceptional Thermal and Acid Stability of Influenza D Virus. mSphere *2*.

Yuan, T. T., Lin, M. H., Chen, D. S., and Shih, C. (1998). A defective interference-like phenomenon of human hepatitis B virus in chronic carriers. Journal of virology *72*, 578-584.

Zamarin, D., Garcia-Sastre, A., Xiao, X., Wang, R., and Palese, P. (2005). Influenza virus PB1-F2 protein induces cell death through mitochondrial ANT3 and VDAC1. PLoS pathogens *1*, e4.

Zhai, S. L., Zhang, H., Chen, S. N., Zhou, X., Lin, T., Liu, R., Lv, D. H., Wen, X. H., Wei, W. K., Wang, D., and Li, F. (2017). Influenza D Virus in Animal Species in Guangdong Province, Southern China. Emerging infectious diseases *23*, 1392-1396.

Chapter 2. Identification and Characterization of Viral Defective RNA Genomes in Influenza B Virus

Abstract

Influenza B virus (IBV) is an important pathogen that infects humans and causes seasonal influenza epidemics. To date, little is known about defective genomes of IBV and their roles in viral replication. In this study, by using next-generation sequencing (NGS) approach, we analyzed total mRNAs extracted from A549 cells infected with B/Brisbane/60/2008 virus (Victoria lineage), and identified four defective genomes in IBV with two (PB1ΔA and PB1ΔB) from the polymerase basic subunit 1 (PB1) segment and the other two (MΔA and MΔB) from the matrix (M) segment. These defective genomes retained the terminal sequences of the relevant segments but contained significant deletions in the central regions with each having the potential for encoding a novel polypeptide. Significantly, each of them can potently inhibit the replication of B/Yamanashi/166/98 (Yamagata lineage). Furthermore, PB1ΔA was able to interfere modestly with influenza A virus (IAV) replication. Finally, we presented evidence that the productions of the four defective RNAs are not dependent on the cell types. In summary, our study provides important initial insights into IBV defective genomes, which can be further explored toward better understanding of the replication, pathogenesis, and evolution of IBV.

2.1 Introduction

Influenza viruses are classified as types A, B, and C according to their distinct antigenic properties residing in two major structural proteins (matrix 1 and nucleoprotein) (P. Palese and M. Shaw, 2007). Recently, a new type of influenza virus with bovine as a primary reservoir, designated influenza type D, has been described (B. M. Hause et al., 2014; B. M. Hause et al., 2013). Influenza B virus (IBV) is a clinically important pathogen. It has been well established that IBVs infect humans, and cause seasonal influenza epidemics along with IAV H3N2 and H1N1 strains (N. J. Cox and K. Subbarao, 2000; J. A. McCullers et al., 1999a; J. A. McCullers et al., 2004; A. S. Monto, 2008). Similar to IAV, IBV can result in a spectrum of clinical diseases in humans, ranging from mild to severe respiratory illness requiring hospitalization and medical treatment (A. S. Monto, 2008; W. Paul Glezen et al., 2013). IBV-associated mortality has been frequently observed in seasonal influenza epidemics particularly in children and adolescents (A. Gutierrez-Pizarraya et al., 2012; P. Wu et al., 2012). The burden of respiratory diseases caused by IBV, coupled with the increasing strong links between IBV and fatal infections in humans (J. A. McCullers and F. G. Hayden, 2012; C. D. Paddock et al., 2012), calls for an urgent need for more basic and clinical investigations of IBV.

Humans are thought to be the primary host and reservoir of IBV, although sporadic infection episodes of IBVs have been described in seals and pigs (A. D. Osterhaus et al., 2000; Z. Ran et al., 2015). In contrast, IAV circulates in various mammals including humans, swine, and migratory or domestic waterfowl (P. Palese and M. Shaw, 2007). IBV evolves at an estimated rate of $2x10^{-3}$ nucleotide substitutions per site per year for both hemagglutinin (HA) and neuraminidase (NA) segments (N.

Nakagawa et al., 2009; R. Nerome et al., 1998; J. Shen et al., 2009). This evolutionary

rate is approximately 2-3 times slower than those observed in IAVs. Despite the

relatively slow evolutionary rate, new antigenic variants of IBV emerge, largely driven

by genetic reassortments of co-circulating IBV strains (J. H. Lin et al., 2007; J. A.

McCullers et al., 2004; J. A. McCullers et al., 1999b). IBV has evolved into two

antigenically and genetically distinct lineages, Yamagata and Victoria (W. P. Glezen,

2014; D. Vijaykrishna et al., 2015). The antigenic diversity of IBVs has posed a

significant challenge for the selection of an appropriate strain for annual vaccine

production.

Similar to IAV, IBV contains eight negative-sense, single-stranded RNA

segments. Despite similarities in the coding strategy and viral protein expression, there

are some notable differences between these two viruses. The most striking difference lies

in the neuraminidase (NA) segment (P. Palese and M. Shaw, 2007). This segment in

IAV is monocistronic, only encoding the NA protein, while the NA segment of IBV is

bicistronic, expressing two integral membrane proteins, NA and NB. The open reading

frame of NB protein starts four nucleotides before the initiating AUG codon that directs

the NA protein synthesis. Another difference is in the polymerase basic 1 (PB1) segment.

In IBV, the PB1 segment only encodes the PB1 protein, whereas in some strains of IAV,

it codes for PB1 and a small accessory protein PB1-F2 (R. Hai et al., 2010). Recent

studies have also provided evidence that IAV's PB1 has the capacity to synthesize PB1-

N40 by using alternative translation initiation strategy (H. M. Wise et al., 2009). Another

major difference between IAV and IBV occurs in the matrix (M) segment with regard to

the expression of M1 and M2 proteins (P. Palese and M. Shaw, 2007). The M2 protein of IAV is translated from a spliced mRNA originating from the primary M segment-derived RNA transcript (R A Lamb et al., 1981), while the expression of IBV M2 protein is produced from the primary RNA transcript of M segment through a coupled termination and reinitiation mechanism, also called translational stop-start, around the sequence encoding M1 stop codon via a UAAUG pentanucleotide motif (C. M. Horvath et al., 1990). Moreover, in a marked contrast to the nonspliced IBV M segment, the IAV M segment can undergo three different splicing events (R A Lamb et al., 1981) and more than 90% of M segment-derived mRNAs are spliced transcripts (N. C. Robb and E. Fodor, 2012). These spliced transcripts in the M and NS segments of IAV were recently speculated for important roles in viral replication, host range, and pathogenesis (Julia Dubois et al., 2014).

In addition to spliced transcripts, the replication and transcription of IAV segments give rise to more than 50 different defective interfering (DI) RNAs (A. R. Davis et al., 1980; A. R. Davis and D. P. Nayak, 1979; J. Michael Janda et al., 1979; D. P. Nayak et al., 1982; S. Noble and N. J. Dimmock, 1995). A DI RNA is a smaller (ranging from 200-700 nucleotides in size) viral RNA (vRNA) with similar or identical terminal sequences to its parental segment and a large internal deletion. The competition of DI genomes with the parental full-length segments in replication, transcription, and genome packaging results in a segment-specific inhibition of IAV replication (T. Odagiri and M. Tashiro, 1997). The underlying mechanism for DI RNA formation has not been fully understood yet. However, there is evidence indicating that a viral RNA polymerase

slippage-based faulty replication process is a likely mechanism (N. J. Dimmock and A. J. Easton, 2014). Despite the observations that DI genomes can originate from any segment, the general consensus is that the three polymerase segments of IAV are the primary donors of DI RNAs (N. J. Dimmock and A. J. Easton, 2014, 2015). DI vRNAs can be packaged into budding virions from the surface of infected cells in a way to out compete their parental segments, leading to the formation of DI particles. DI particles containing incomplete genomes are unable to replicate, unless the missing viral protein(s) is supplied by coinfection with replication-competent viruses. In addition to directly interfering with wild-type IAV replication, some DI genomes are potent inducers of the innate immune response (A. Baum et al., 2010; T. Frensing et al., 2014; J. M. Ngunjiri et al., 2013; M. Perez-Cidoncha et al., 2014). Recent emerging data consistently indicate that the IAV-derived DI genomes can protect MDCK cells or mice from infections of IAV, IBV, and other respiratory viral pathogens such as respiratory syncytial virus (N. J. Dimmock et al., 2008; A. J. Easton et al., 2011; P. D. Scott et al., 2011). Furthermore, IAV DI particles were identified in the respiratory tract of infected chickens and humans (W. J. Bean et al., 1985; T. M. Chambers and R. G. Webster, 1987; K. Saira et al., 2013), suggesting a potential role in viral transmission and pathogenesis.

Despite significant progress on the characterization of the DI genomes in IAV and harnessing them for broad-spectrum antivirals, and an early description of DI genome phenomena in IBV (P. Von Magnus, 1954), little is known about the molecular nature of IBV DI genomes and their roles in viral replication. The investigation of DI genomes of IBV will improve our understanding of IBV pathogenesis and host immune response,

which may inform alternative strategies for vaccine design. A high-resolution analysis of dynamic changes in viral gene transcription following IBV infection has not been reported. In this study, we investigated the viral gene expression dynamics of IBV in IBV-infected human lung epithelial A549 cells using next-generation sequencing (NGS) approach. Specifically, we identified four defective RNAs of IBV. Among them, two defective RNAs (PB1ΔA and PB1ΔB) were derived from the polymerase segment 1 (PB1). The other two defective RNAs (MΔA and MΔB) were derived from the M segment. Interestingly, the majority of the defective transcripts were from the M segment with MΔA being the most abundant, which is substantially different from IAVs in which most DI transcripts are generated from the three polymerase segments. Furthermore, we showed that these four defective vRNAs strongly inhibited IBV replication and, to a lesser degree, IAV replication. Production of these four defective vRNAs appeared to be independent of cell types used for cultivation of IBV. Our findings provide new insights into IBV-specific defective genomes and further characterization of these small vRNAs may help to better understand the biology and evolution of human IBV.

2.2 Materials and Methods

2.2.1 Cells and virus.

Human lung alveolar carcinoma epithelial cell line A549 (ATCC® CCL-185) cells, Madin-Darby canine kidney (MDCK) cells (ATCC® CCL-34), and human Calu-3 (ATCC ® HTB-55™) maintained in Dulbecco's minimum essential medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS, PAA Laboratories Inc., Dartmouth, MA, USA) and 100 U/ml penicillin-streptomycin (Life Technologies,

Carlsbad, CA, USA). Influenza B/Brisbane/60/2008 virus was provided by Drs. Ruben Donis and Xiyan Xu (Center of Disease Control and Prevention, USA). The virus was propagated using MDCK cells at an MOI of 0.1. The cells in a T75 tissue culture flask were allowed to reach only 60% to 70% confluency at the time of infection. The virus inoculum was suspended in 2 ml DMEM and incubated at 37°C in 5% CO_2 for 1 h. Following infection, fresh DMEM with 0.5 μg/ml tolylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin (Sigma, Saint Louis, MO, USA) was added for further incubation at 37°C in 5% CO_2 for 3 days. After 3 days, the supernatant was harvested and spun at 500 × g for 10 min at 4°C to remove the cellular debris followed by $TCID_{50}$ determination in MDCK cells. DMEM supplemented with 200 U/ml penicillin-streptomycin (Life Technologies, Carlsbad, CA, USA) and 1 μg/ml TPCK-treated trypsin (Sigma, Saint Louis, MO, USA) was used as the virus growth medium.

2.2.2 Viral growth kinetics in A549 cells

A549 cells were inoculated with influenza B/Brisbane/60/2008 virus at an MOI of 1.0. After 1 h-incubation at 37°C in 5% CO_2, the viral inoculum was removed from the cells. Cells were washed 3 times with PBS and then fresh virus growth medium were added. At 0, 6, 12, and 24 hpi respectively, supernatant samples were collected and used for the determination of viral $TCID_{50}$ in MDCK cells. Four independent experiments were performed with each conducted in duplicate.

2.2.3 Cell proliferation assay

A549 Cells (15000 cells/well) were seeded onto 96 well plates overnight. Cells were then infected with influenza B/Brisbane/60/2008 at an MOI of 1.0. After 1-h incubation, the cells were washed with PBS 3 times, and then replaced with fresh virus growth medium. At 6, 12, and 24 hpi, respectively, CellTiter 96 Aqueous One Solution (Promega, Madison, WI, USA) containing tetrazolium compound [3-(4,5-dimethyl-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS] was added to make MTS final concentration at 317µg/ml. After 1-hour incubation with MTS, the absorbance was detected at 490 nm using a Synergy 2 microplate reader (Biotek, Winooski, VT, USA). Two independent experiments were performed with each run in triplicate. To estimate the cell numbers according to absorbance values measured for A549 cells, the corrected absorbance values at 490nm versus cell number (from 5000 to 35000 with 5000 increments) were used to plot a regression equation with an R squared of 0.992. Cell numbers at different time points (6, 12, 24 hpi) for control cells and virus-infected cells were calculated according to this equation.

2.2.4 Next-generation sequencing

A549 cells were infected with influenza B/Brisbane/60/2008 at an MOI of 1.0. Total cellular RNA was isolated from mock-infected cells (0 hpi) and infected cells at, 6, 12, and 24 hpi, respectively, with TRIZOL followed by RNA purification with ethanol. RNA quality was tested using a Bioanalyzer 2100 (Agilent, Palo Alto, CA) and then processed for cDNA library construction by using a cDNA library prep kit (Illumina Inc., San Diego, CA) according to the manufacturer's instruction. cDNA library construction and all other procedures were conducted in Genomics Core Research Facility (GCRF) at the University of Nebraska-Lincoln (UNL). Briefly, mRNAs were purified from the total

RNA using oligo(dT) magnetic beads followed by fragmentation. The resultant mRNAs were reverse transcribed to cDNAs that were subjected to an end repair process followed by ligation to the adapters. After separation in agarose gel through electrophoresis, cDNA fragments with a size of about 200 bp were excised, extracted, and amplified by PCR using two primers that match the ends of adaptors. PCR-enriched samples were then sequenced by an Illumina Genome Analyzer IIx (GA IIx) sequencer in the GCRF at UNL. Two biological replicates were sequenced for each time point (0, 6, 12 and 24 hpi). The eight samples were barcoded and sequenced in one lane. Each of the eight samples has approximately 5-7 million 100-nucleotide single-end reads. All sequencing results passed quality control. The raw reads of the eight samples were submitted to NCBI SRA database under Biosample ID: SRS1328599.

2.2.5 Transcriptome read processing

The raw read quality was checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The 5 nucleotides at the 3' end of each read, whose sequencing qualities were low, were removed using fastx_trimmer (http://hannonlab.cshl.edu/fastx_toolkit/index.html). GSNAP was then used to map raw reads to human (version: GRCh37.p13) and Influenza B virus genomes (T. D. Wu and S. Nacu, 2010). Samtools was used to index bam files and remove PCR duplicates (H. Li et al., 2009). HTSeq was used to count the number of reads mapped to human genes defined by the Ensemble genome database and viral genes (S. Anders et al., 2015). The percentage of viral RNA was calculated by the equation viral mRNA reads/total mRNA reads.

2.2.6 RT-PCR amplification and sequencing of defective RNA genomes

A549 cells were infected with B/Brisbane/60/2008 virus at an MOI of 1.0. Total cellular RNAs were isolated from mock-infected cells and infected cells at 24 hpi, respectively, with TRIZOL followed by subsequent RNA purification with ethanol. The extracted RNAs were subjected to segment-specific RT-PCRs for the amplification of defective RNA molecules. Following electrophoresis in 1% agarose gel, amplified RT-PCR products were excised and sequenced by a commercial operator (http://www.genscript.com/) and contigs were then assembled using IBV PB1, M, and NS segment sequences. The GenBank accession numbers of these defective RNA sequences are KX092351 for PB1ΔA, KX092352 for PB1ΔB, KX092353 for MΔA, and KX092354 for MΔB. Inclusion of NS segment in RT-PCR amplification served as a positive control because of a well-known splicing event originated from the NS segment. The validation primers are shown in Table 2.1. Briefly, cDNA samples specific to PB1 or M segments were reverse transcribed from total RNA with forward primers BPB1-1For and BM-For (Table 2.1) that target only negative-sense vRNAs. The method to reverse transcribe negative-sense vRNAs was described by Kawakami E (E. Kawakami et al., 2011). Each RT-PCR product was sequenced with the same set of primers for its PCR.

To determine whether or not the observed defective genomes could be produced from the segment-derived mRNA in the absence of active viral replication, human embryonic kidney cells HEK293T (ATCC ® CRL-3216) were transfected, respectively, with PB1, M, and NS segments of influenza B/Yamanashi/166/98 reverse genetics

system (E. Hoffmann et al., 2002). At 48 h post-transfection, total RNAs were extracted and then treated with DNase to remove residual DNA. Following the inactivation of DNase by phenol/chloroform extraction and ethanol precipitation, the treated RNAs were used in subsequent RT-PCR reactions with the above primers and experimental conditions. DNA plasmids, containing PB1 or M or NS segment, harbors bidirectional promoters (Pol I and II), which allow for the synthesis of both mRNAs (RNA polymerase II) and negative-sense vRNAs (RNA polymerase I). The amplified products were separated through electrophoresis in 1% agarose gel.

Table 2.1. Validation Primers used in IBV defective genome RT-PCR.

Target	Expected Size (bp)	Forward primer (5'-3')			Reverse primer (5'-3')		
		Name	Sequence	Positions	Name	Sequence	Positions
PB1ΔA	286	BPB1-1For	ATGAAT ATAAATC CTTATTT TCTCTTC ATAGAT GTGCCC G	22-61	BPB1-Rev	CCAATAACC CCATAAACA TCTTCGAAGC TTATATGTAC CC	2269-2308
PB1ΔB	190	BPB1 ΔB_L 1	GGAACA GGAACA GGCTAC ACA	118-138	BPB1Δ B-R1	GTGAGCCAT TGCTTTCTCA A	2237-2256

MΔA	206	BM-For	CTTTCTT AAAATG TCGCTGT TTGGAG ACACAA TTGC	15-50	BM-Rev	GCATAGTAA GAAATACAG TAAAATTGA ATTTAATGC	1094-1129
MΔB	163	BMΔ B_L1	GTTGGTT TGGTGG GAAAGA A	113-132	BM-Rev	GCATAGTAA GAAATACAG TAAAATTGA ATTTAATGC	1094-1129
NS2	410	BNS2-For	GCAGAA GCAGAG GATTTGT TTAGTCA CTGGCA AAC	2-36	BNS-Rev	CAAGAGGAT TTTTATTTTA AATTCACAA GCACTG	1060-1093

2.2.7 Quantitative measurement of relative abundances of defective genomes to their parental full-length segments.

To measure the ratios of MΔA to full-length M segment and of PB1ΔA to full-length PB1 segment, A549 cells were infected with 1.0 MOI of IBV and both clarified supernatants and cell lysates were collected at 6, 12, and 24 hpi. Viral RNAs from supernatant and total cellular RNAs from cell lysates were then extracted using Trizol Reagent (Invitrogen, CA, USA) according to the manufacturer's protocol. RNA concentration and purity were determined by NanoDrop 1000 spectrometer (Thermo Scientific) and 200 ng RNA was used for reverse transcription with vRNA-specific

primers using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA). Droplets were made and analyzed on the Bio-Rad QX200 Droplet Digital PCR (ddPCR) system according to the Droplet Digital PCR applications guide. The ddPCR reaction consisted of 10 ul 2x Supermix for Probes (Bio-Rad), 900 nM primers, 25 nM probes and 8 ul cDNA into a final volume of 20ul. The primers and probes were specifically designed to query both the truncated (via targeting junction regions) and full-length genomes (Table 2.2). No-template controls (NTC) were included in every run. Three separate experiments were performed with each assaying sample in triplicate.

Table 2.2. Primers and probes used for ddPCR.

Target genes	Primers/probes name	Primers/probes sequence (5'-3')
PB1ΔA	qPB1A-For	GCTATGTTGACCCACTGGCTTC
	qPB1A -Rev	GGCAGCAATTTCAACAACATTC
	qPB1A -Probe	6FAM-TGCACTGTTATGGGAATAA-MGBNFQ
PB1	qPB1-FL-For	CCACCTTTGTCGGCTCCAT
	qPB1-FL-Rev	AACGGCACTGAACACAACAATAA
	qPB1-FL-Probe	VIC-ATCATTCAACCTAAAAGAG-MGBNFQ
MΔA	qMA-For	GCAGAAAGCCCCTCAATTATTATGT
	qMA-Rev	GACAGAAGATGGAGAAGGCAAAG
	qMA-Probe	6FAM-TCATTCAATACCTCAGTTCT-MGBNFQ
M	qM-FL-For	TATGAGCCCTGTGTGAATGTGAT
	qM-FL-Rev	GCAAGTAAAACTAGGAACGCTCTGT
	qM-FL-Probe	VIC-CTTGTTTCTCGCATAAAG-MGBNFQ

2.2.8 Inhibition of IBV and IAV replication by defective RNAs.

293T and MDCK cells were co-cultured at approximately 70% confluence in 12-well plates and transfected with each of indicated defective RNA plasmids (0.25 μg each) in the context of pCAGEN vector (pCAG, positive-sense mRNA production) or pHW2000 vector (pHW, both positive- and negative-sense RNA production), together with an 8-plasmid IBV RGS system (0.25 μg each plasmid). Transfection system was prepared in Trans-LT1 (Mirus) and Opti-MEM I medium according to the manufacturer's instructions. Culture supernatants were collected at 72 hpi and titrated in MDCK cells to determine end-point titers ($TCID_{50}$/ml). The dose-dependent effects of PB1ΔA and MΔA in interference with IBV replication were investigated under similar experimental conditions by decreasing the amount of the defective genome plasmid used for transfection while maintaining the same concentration of the corresponding parental segment. Three relative ratios of the defective RNA and its parental segment plasmids: 1x, 1:1 (0.25 to 0.25 μg); 0.2x, 1:5 (0.25 to 1.25 μg); and 0.1x, 1:10 (0.25 to 2.5 μg), were used together with seven other IBV RGS plasmids (0.25 μg each plasmid) in the transfection/infection experiment. Three separate experiments were performed with each conducted in triplicate.

A similar approach was used to investigate whether expression of IBV-derived PB1ΔA and MΔA defective genomes could interfere with IAV replication. Co-culture of HEF293T and MDCK cells was transfected with an 8-plasmid IAV A/WSN/33 RGS (0.25 μg each plasmid) in the absence or presence of different amounts of PB1ΔA and MΔA. The amount of defective RNA plasmids used was 0.25 μg, 1.25 μg and 2.5 μg,

which resulted in three different dose groups (based on the relative ratios of the defective RNAs and its parental segment plasmids): 1x, 1:1 (0.25 to 0.25 μg); 5x, 5:1 (1.25 to 0.25 μg); and 10x, 10:1 (2.5 to 0.25 μg). Culture supernatants were collected at 24 h post-transfection and viral $TCID_{50}$/ml were measured at 6, 12, and 24 h following infection of the MDCK cell line, respectively. DNA amount in each transfection reaction here and above was made equivalent by adding pHW empty vector. Three separate experiments were performed with each conducted in triplicate.

2.3 Results

2.3.1 Analysis of viral gene transcription by Next-Generation Sequencing (NGS)

To investigate IBV mRNA transcription profile and the host response to IBV infection, we infected human lung epithelial A549 cells in duplicate with human IBV strain (B/Brisbane/60/2008, Victoria lineage) at a multiplicity of infection (MOI) of 1.0. IBV growth kinetics in A549 cells showed that the viral replication was detected at 6 hours post infection (hpi) at an appreciable level ($10^{1.5}$ $TCID_{50}$/ml) and increased in a time-dependent manner over a 24-hour period (Fig. 2.1A). At 24 hpi, approximately 98% of cells expressed IBV-specific antigens as determined by flow cytometry-based assay (data not shown), suggesting nearly complete infection was achieved. IBV-infected cells and uninfected controls exhibited similar cell proliferation profiles at 6 and 12 hpi, but a statistical significant difference in cell proliferation was observed at 24 hpi (Fig. 2.1B), indicating virus replication exerted a time-dependent effect on the cell proliferation.
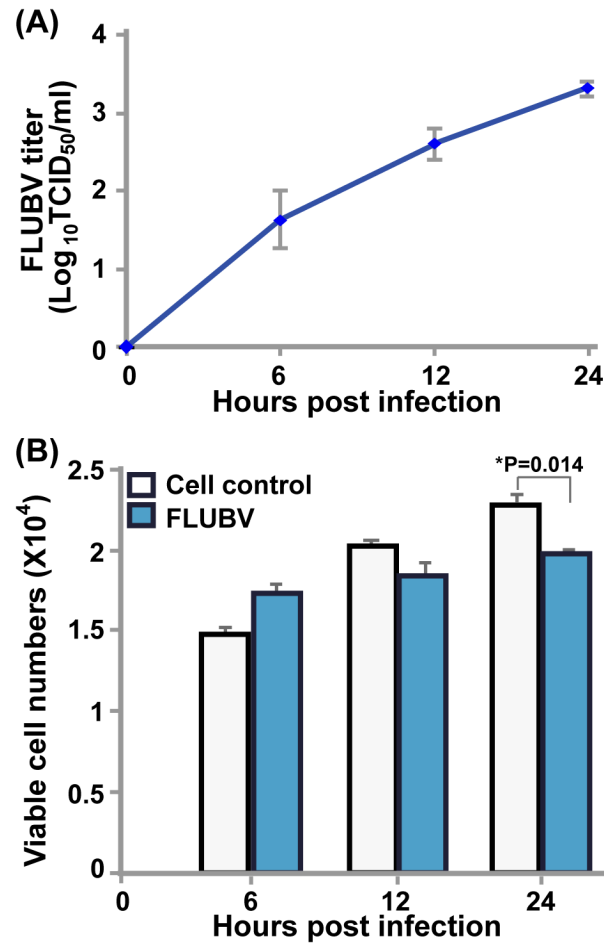
Figure 2.1. Viral replication kinetics and proliferation dynamics of IBV infected A549 cells. (A) Viral titer in A549 cells infected with B/Brisbane/60/2008 virus. Supernatant samples of A549 cells infected by 1.0 MOI of IBV were collected at 0, 6, 12, and 24 hpi. The $TCID_{50}$ of each sample was evaluated in MDCK cells. Mean $TCID_{50}$ ($\pm$standard deviations) from 3 independent experiments in duplicate are shown. (B) A549 cell growth in the first 24 h of infection. MTS assay in 96-well plate format was performed at 0, 6, 12, and 24 hpi. The bars represent mean cell numbers (+standard deviations) from 3 independent experiments in triplicate. A '*' indicates a significant difference using student's t-test ($P<0.05$).

Total mRNAs were isolated from infected cells at 0, 6, 12, and 24 hpi and the corresponding cDNA libraries were sequenced using Illumina-based NGS technology. The NGS reads were then aligned to the eight segments of IBV as well as the human genome. To measure similarities of gene expression profiles at all time points, we calculated Pearson correlation coefficients (Pcc) distance for all pairs of samples and constructed an Unweighted Pair Group Method with Arithmetic mean (UPGMA) dendrogram (Fig. 2.2A). The dendrogram showed that the replicate samples of all time points clustered together first, suggesting the variations of expression profiles between repeats are lower than those displayed between time points. The Pcc distance of gene expression profiles between mock cell and other time points increased over time, suggesting cellular response to IBV infection was significantly changed. In addition, principle component analysis was used to further calculate the variations between gene expression profiles (Fig. 2.2B). The analysis showed that there was little variation between repeats of the same time point, suggesting a high reproducibility of the experiments. The gene expression profiles of 6, 12, and 24 hpi showed gradual divergence from that of mock cell samples along principal component 1 (PC1). The 6 and 12 hpi profiles were also different from 0 hpi along PC2 while the profiles of 24 hpi have PC2 scores similar to 0 hpi, implying the expression of some genes was altered at 6 and 12 hpi and then gradually returned to basal levels at 24 hpi.

For each sample, we obtained on average over 6 million 100-nucleotide (nt) reads mapping to either IBV or human genome. At 6 hpi, around 2.5% of the total reads

mapped to the viral genomic segments, while at 12 hpi, this proportion increased to 24%.

At 24 hpi, nearly 30% of the total mapped reads were derived from IBV segments (Fig. 2.2C). The observed increase in viral reads over time correlated well with a step-wise increase in viral titers as shown in the viral replication kinetics experiment (Fig. 2.1A). During infection, M segment-derived viral transcripts were predominant as measured by Reads Per Kilobase of transcript per Million mapped reads (RPKM) at 12 and 24 hpi (Fig. 2.2D), followed by those of HA, NP, NA, and NS segments. The viral transcripts of the polymerase segments (PA, PB1, and PB2) were detected with the lowest abundance. Intriguingly, the expression of M segment-related transcripts continued to increase to 24 hpi, which was in a marked contrast to viral transcripts of all other seven segments. For the latter, the numbers of segment-specific viral transcripts either reached a plateau at 12 hpi with negligible increase for PA, PB1, PB2, and NS segments; or slight decrease for HA, NP, and NA segments at 24 hpi, respectively.
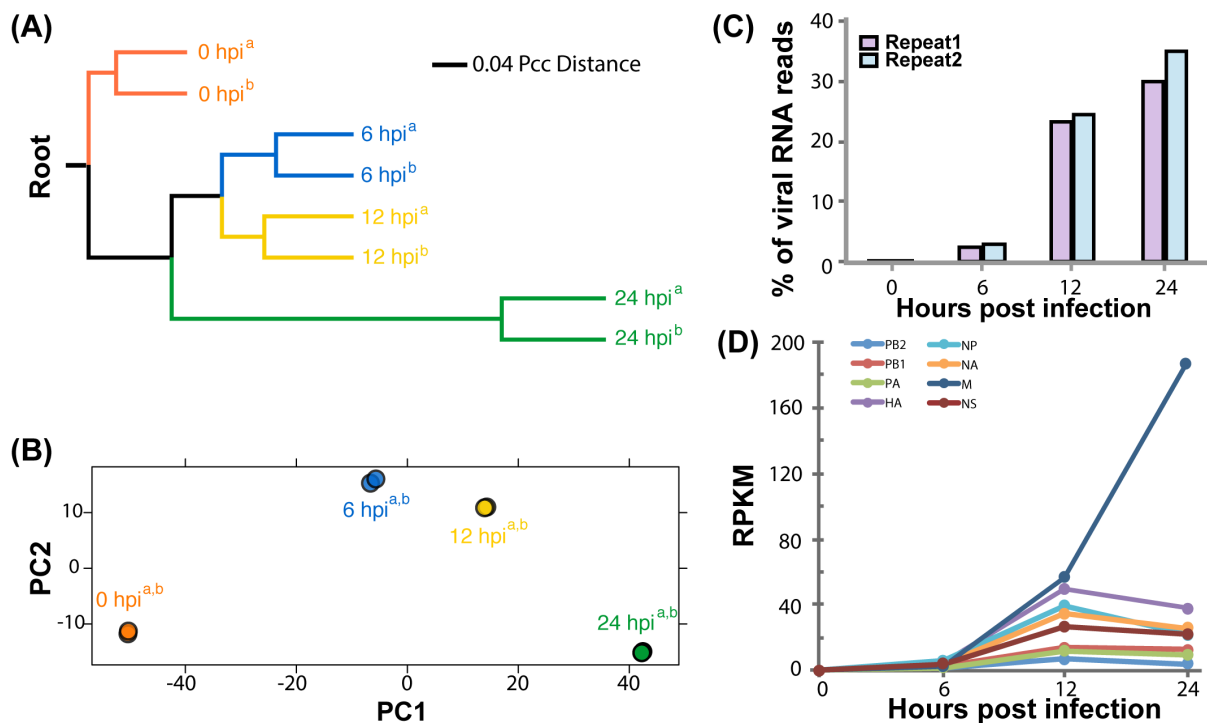
Figure 2.2. Analysis of viral gene transcription by NGS. (A) Hierarchical clustering of expression profiles. The UPGMA tree was rerooted using RNA-seq read from the two control samples as root. The divergence of expression profiles at 0, 6, 12, and 24 hpi is shown in the UPGMA dendrogram. Two repeats (indicated with superscript 'a' and 'b') at each time point are shown separately. (B) Principal components analysis of host gene expression profile divergence. IBV-induced expression differences are shown using the first two components (PC1 and PC2) of Principal components analysis. The two repeats at each time point are partially (6 hpi) or almost completely (0, 12, and 24 hpi) overlapping with each other for both PC1 and PC2. Two repeats (indicated with superscript 'a' and 'b') are shown separately. (C) The percentage of viral reads in the NGS samples. Genome-wide total gene expression was quantified using RNA-seq. Two total RNA samples at each time point were extracted from A549 cells infected with 1.0 MOI

of IBV. Percentage of viral genome reads was shown, which represents the ratio of the reads of IBV genome to total reads of genomes of Homo sapiens and IBV. (D) Expression level changes of the eight IBV segments over time. The expression level was displayed using reads per kilobases per million reads (RPKM).

2.3.2 Identification of defective RNA genomes from M and PB1 segments

Closer examination of the position-dependent read coverage across IBV genome showed that M segment and, to a lesser extent, PB1 segment had a large proportion of reads mapping to the 5' and 3' ends of segments, with a significant reduction in the number of reads mapping to the central regions (Fig. 2.3).
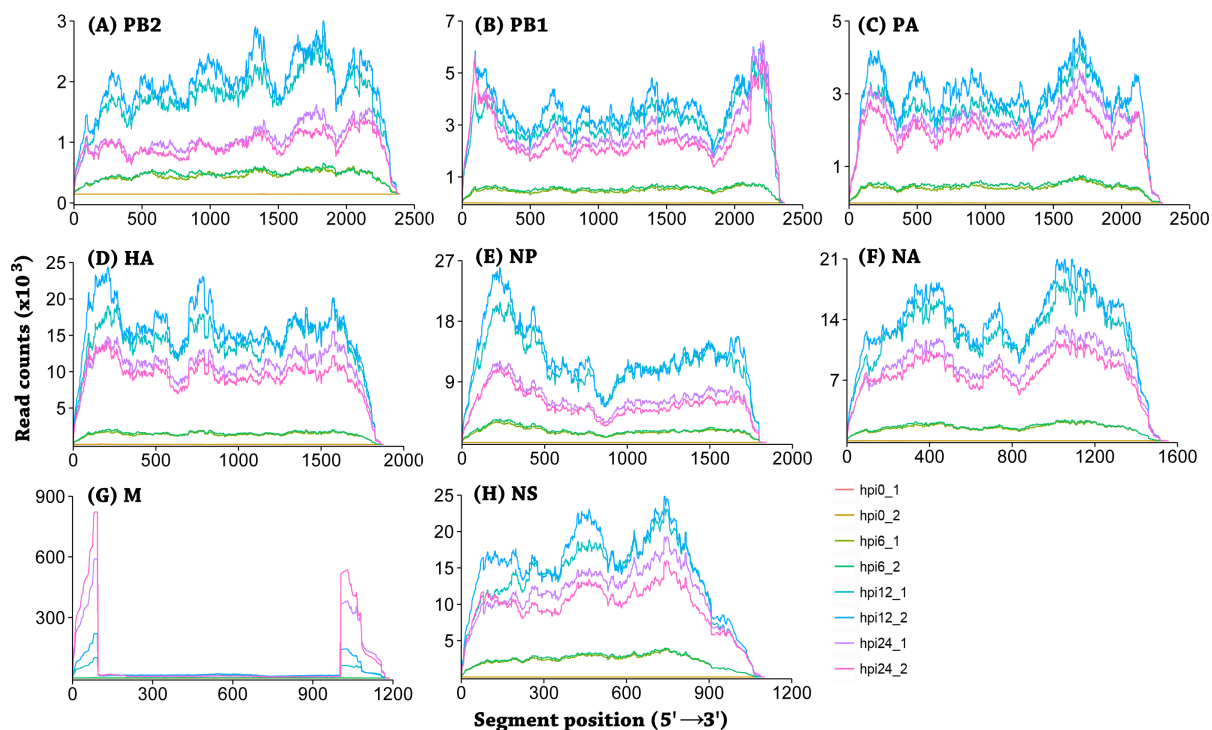
Figure 2.3. Positional sequencing depth of IBV gene segments over time. Panels A-H show positional sequencing depth for segments PB2, PB1, PA, HA, NP, NA, M, and NS, respectively. The sequencing depth for each segment is shown as RNA-seq read counts at each position (5'→3'). Two replicates from each time point (0, 6, 12 and 24 hpi) are shown separately.

The overrepresentation of sequencing reads in the 5' and 3' ends of the segments suggested the presence of internal deletions in transcripts of these genes. This observation prompted us to search for junction reads that were mapped to two discontinuous regions of each segment toward better understanding of the origin of these RNA molecules. Analysis of NS segment-associated junction reads was included in this study because NS segment is known to utilize a splicing mechanism to produce NS2 mRNA. To remove sequencing noises, we only focused on junction sites that were supported by more than 200 reads. In the NS segment, we observed a splicing junction site between genomic positions 80 and 763 (nts 45-80 joined to 763-1071) with sequencing coverage of more than 1000 reads, matching the previously reported splicing donor and acceptor sites for the NS2 transcript (D. J. Briedis and R. A. Lamb, 1982) (Fig. 2.4A). Further analysis showed that this splicing event utilized a consensus dinucleotide motif GT/AG: GT present at the extreme 5' (donor) and AG at the extreme 3' (acceptor) ends of the intron. In addition, we identified two splicing-like junction sites in the PB1 segment that joined genomic positions 116 with 2118, and 240 with 2190, respectively, which resulted in the production of two previously uncharacterized small transcripts, PB1ΔA and PB1ΔB (Fig. 2.4B). Moreover, two splicing-like junction sites in the M segment that joined genomic

positions 95 with 1005, and 199 with 1054, respectively, were also observed (Fig. 2.4C). Further analysis indicated that none of the junction sites observed in PB1 and M segments conformed to the canonical splicing donor and acceptor dinucleotide motif GT/AG that are present in the splicing sites of human and influenza viral genes (N. Sheth et al., 2006). Follow-up Reverse transcription polymerase chain reaction (RT-PCR) and Sanger sequencing experiments confirmed the presence of NS2, PB1ΔA, PB1ΔB, MΔA, and MΔB mRNAs in virus-infected A549 cells (Fig. 2.4D). Specifically, with primers corresponding to terminal sequences of the segment (Table 2.1), we amplified nearly full-length PB1ΔA and MΔA with the identical junction site sequences as determined in RNA-Seq (Figs. 2.4B-C). For PB1ΔB and MΔB, following the initial failure in PCR amplification with these terminal primers (Table 2.1), the revised PCRs with forward primer sequences (Table 2.1) specific to the junction regions of PB1ΔB and MΔB successfully amplified PB1ΔB and MΔB (Fig. 2.4D). The junction site position for MΔB determined by RT-PCR coupled with Sanger sequencing was identical to that measured in RNA-Seq experiment. We have only noted one discordance in the junction site position for PB1ΔB determined between NGS and conventional RT-PCR and Sanger sequencing methods. In RNA-Seq, the junction site position for PB1ΔB was 240/2190, while RT-PCR and Sanger sequencing determined the junction sites for PB1ΔB to be 240/2189, which was chosen for cloning of PB1ΔB plasmids that were used in functional studies below.

Based on the well-documented literature on IAV DI RNAs (A. R. Davis et al., 1980; A. R. Davis and D. P. Nayak, 1979; N. J. Dimmock and A. J. Easton, 2015; J.

Michael Janda et al., 1979; D. P. Nayak et al., 1982; N. Sheth et al., 2006) plus that none of the detected junction sites for the four small transcripts is consistent with the GT/AG splicing motif, we speculated that these small transcripts were not derived from mRNA splicing, instead, they are likely to be viral RNA polymerase-dependent products. To test the hypothesis, we transfected each of the three segments (PB1, M, and NS) to HEK293T cells separately using a vector containing a CMV promoter, and extracted total RNAs from transfected cells. We then used RT-PCR to detect PB1ΔA and MΔA with same sets of PCR primers and PCR conditions we used to detect them in virus-infected cells (Fig. 2.4D). The NS2 transcript was detected, indicating that the NS2 splicing event occurred independent of viral replication, a bona fide splicing-derived transcript (Fig. 2.4E). In a marked contrast, neither of the two transcripts (PB1ΔA and MΔA) was detected, indicating that their synthesis needs active viral replication. These data support the speculation that these PB1 segment- and M segment-specific small transcripts are produced by viral RNA polymerase-driven transcription.
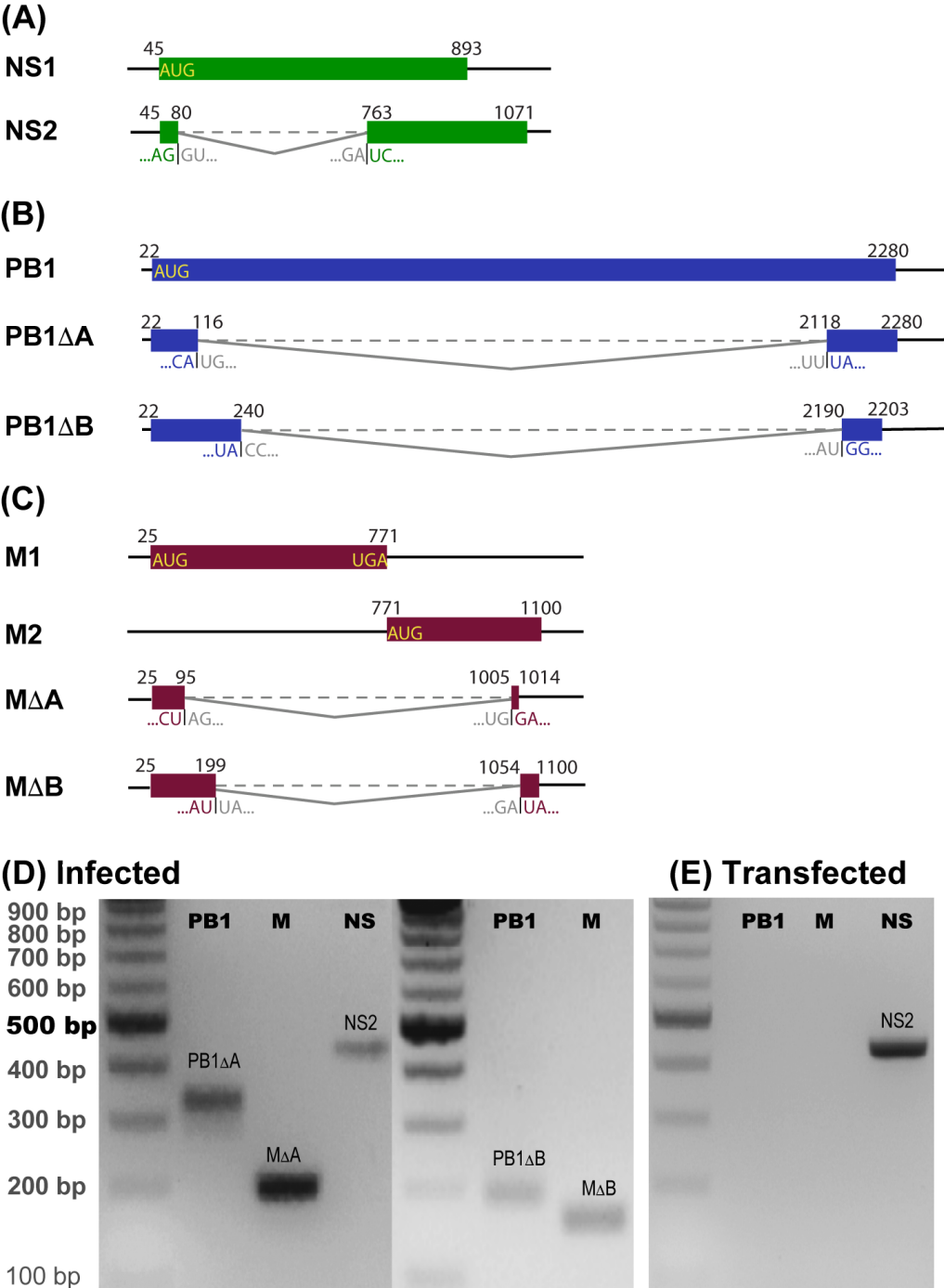
Figure 2.4. Alternative splicing or junction sites and defective RNA genomes. (A) NS segment splicing. Diagram showing the known splicing site in IBV NS segment that generates NS2 mRNA. (B) PB1 segment-derived defective RNAs. Schematic diagram

shows the two newly identified defective RNAs from PB1 segment termed as PB1ΔA, and PB1ΔB. (C) M segment-derived defective RNAs. Schematic diagram displays the two newly identified defective RNAs in the M segment termed as MΔA and MΔB. The initiation, junction, and termination positions (5'→3') of an mRNA's ORF are labeled with position numbers above. An mRNA's two junction sequences are shown under ORFs. (D) Detection of PB1ΔA, MΔA, PB1ΔB, MΔB and NS2 RNAs in IBV infected A549 cells. Segment-specific primers were used to amplify full-length PB1ΔA, MΔA, and NS2 by RT-PCR. Junction-specific primers (forward) and segmental primers (reverse) were used to amplify partial PB1ΔB and MΔB transcripts. The cDNA was generated from total mRNA extracted from IBV-infected A549 cells. (E) Absence of PB1ΔA and MΔA RNAs in HEK293T cells transfected with plasmid DNA encoding IBV PB1 and M segment. Segment-specific primers were used to amplify PB1ΔA and MΔA by RT-PCR. NS2-specific RT-PCR was performed as a splicing control. The cDNA was generated from total mRNA extracted from transfected HEK293T cells with IBV PB1, M, or NS segments.

2.3.3 Abundances of Defective RNAs

We then used the RNA-Seq data to determine the transcription dynamics of defective RNAs in influenza B virus-infected A549 cells (Fig. 2.5A). We estimated the relative transcription levels of defective RNAs by calculating the number of junction reads per million mapped reads. The analysis showed that the levels of PB1ΔA and PB1ΔB were very low at 6 hpi, and reached peak levels at 24 hpi. Similar to PB1ΔA and PB1ΔB, MΔA and MΔB levels reached the highest at 24 hpi (Fig. 2.5A). Interestingly,

the transcription level of MΔA was over 100-fold higher than MΔB. The positional read coverage at the MΔA junction sites was about 70-fold higher than that in the central region of M segment, which is largely from M1 and M2 transcripts. This suggested that the transcription level of MΔA was significantly higher than these of M1 and M2.

We next developed vRNA-sense reverse transcription (E. Kawakami et al., 2011) coupled with two-color probe Droplet Digital PCR (ddPCR) assay that enabled us to measure the relative abundance of the observed defective vRNAs to their parental segments in both intracellular and extracellular (virus particles) environments. Such an assay also allowed us to detect transcripts derived only from vRNA-sense, not from mRNA and cRNA forms. We selected MΔA and PB1ΔA as focused defective species here and during the rest of the study. As shown in Figs. 2.5B and 2.5C, our quantitative data showed that MΔA (Fig. 2.5B) tended to be much more abundant relative to its parental segment than PB1ΔA (Fig. 2.5C) in both infected cells and virus particles at different time points following viral infection. MΔA was nearly 200-fold and 300-fold more abundant than the full-length M segment in infected cells and purified virus particles at 24 hpi, respectively (Fig. 2.5B). In marked contrast, PB1ΔA only exhibited about 1.25 to 1.75-fold increase in copy numbers over the full-length PB1 in both infected cells and virus particles at this time point (Fig. 2.5C). Significantly, our quantitative analysis on abundance of IBV defective vRNA transcripts correlated well with our RNA-Seq analysis of the mRNA profile (Figs. 2.2D and 2.5A). The results of these experiments provided evidence for the existence and the abundance of the negative-sense defective genomes, which are correlated with the level of virus replication.
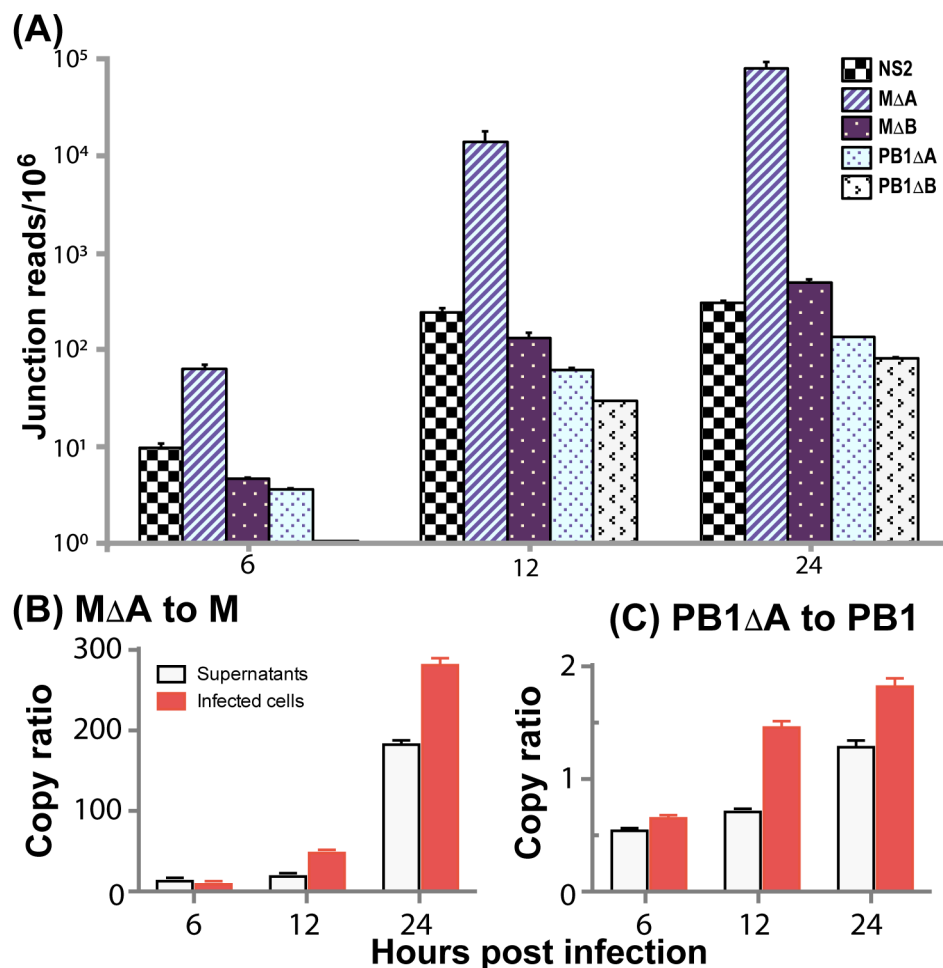
Figure 2.5. Transcription profiles of defective RNAs and their relative abundances.
(A) Transcription profiles of defective mRNAs. RNA-seq junction reads for the four defective RNAs (MΔA, MΔB, PB1ΔA, and PB1ΔB) and NS2 are shown as junction reads per million total reads at indicated time points following IBV infection of A549 cells. (B) Ratio of MΔA vRNAs relative to full-length M vRNAs. (C) Ratio of PB1ΔA vRNAs relative to full-length PB1 vRNAs. Ratios of MΔA to full-length M segment and of PB1ΔA to full-length PB1 segment were measured respectively by ddPCR. Both clarified supernatants and cell lysates of A549 cells infected with 1.0 MOI of IBV were collected at 6, 12, and 24 hpi. Total RNAs were then extracted and 200 ng RNA was used for

reverse transcription with vRNA-specific primers. ddPCRs were conducted with primers and probe specific to amplify defective RNAs (the junction regions) and full-length segments. No-template controls (NTC) were included in every run. The ratios of defective genomes to their corresponding full-length segments were normalized and shown as mean±SEM (standard error of the mean) from three separate experiments with each assaying sample in triplicate.

It should be noted that we have also investigated whether the viral inoculum contained these defective vRNAs. Using the same vRNA-sense RT-PCR coupled with the Sanger sequencing method, we found that viral inoculum contained PB1ΔA and MΔA, which is in good agreement with the above data, showing that PB1ΔA and MΔA vRNAs can be packaged into budding viral particles. These results demonstrated that defective vRNAs such as MΔA and PB1ΔA required active IBV replication. The presence of both MΔA and PB1ΔA in viral particles suggested that these short forms were incorporated into budding virions during viral morphogenesis.

2.3.4 Inhibition of IBV and IAV replication by defective vRNAs

It is known that vRNA-sense DI genomes of IAV are required to inhibit virus replication (Y. Boergeling et al., 2015; T. Odagiri and M. Tashiro, 1997). To determine whether the four defective genomes (MΔA, MΔB, PB1ΔA, and PB1ΔB) were able to interfere with the replication of IBV, we placed each individual defective sequence into the pHW2000 (pHW) vector with bidirectional promoters (Pol I and II) that allow both RNA polymerase I transcription of vRNAs and RNA polymerase II-driven transcription

of mRNAs. As such, we generated a set of control plasmids in the pCAGEN (pCAG) vector that just contains a pol II promoter (chicken ß-actin) so only mRNAs are transcribed. To test the effects of vRNA-sense defective genome transcription on virus replication, we employed a transfection-based reverse genetics system (RGS) to generate B/Yamanashi/166/98 virus (E. Hoffmann et al., 2002) in the presence or absence of defective genomes (1x, 1:1 ratio of defective genome and the corresponding parental segment plasmids), viral infectivity was then determined to measure the defective RNAs' effects. As shown in Fig. 2.6B, IBV replication was completely inhibited in the presence of either defective genome derived from the pHW vector, as robust virus replication was only detected in the empty vector control (mock). In contrast to the halted replications by the pHW series, the replication levels were not much affected by the pCAG series (Fig. 2.6). The only noticeable inhibition by the pCAG series was approximately 1-log reduction in the pCAG-PB1ΔA (Fig. 2.6A). This moderate inhibition by the transcriptions of PB1ΔA mRNA may indicate a role of a novel polypeptide from this specific defective transcript in viral replication, which should be investigated in future study. Overall, our activity data suggested that the presence of each of the four defective transcripts in vRNA form, but not in mRNA form, is sufficient to inhibit IBV viral replication.

To further investigate the dose-dependent effects of the defective vRNA-induced interference with IBV replication, we decreased the amount of the pHW-vectored plasmids used for transfection while maintaining the same concentration of the corresponding parental segment. Three relative ratios of the plasmids carrying defective

genome sequences to the plasmids carrying parental segment sequences (1x, 1:1; 0.2x, 1:5; and 0.1x, 1:10) were used together with seven other IBV RGS plasmids in the transfection/infection experiment. Dose-dependent suppression of IBV replication by PB1ΔA or MΔA was nicely demonstrated in Figs. 2.6C and 2.6D. Specifically, PB1ΔA or MΔA, when used in equal amount as its parent segment plasmid (1x group), completely inhibited IBV replication (Figs. 2.6C and 2.6D), which was similar to our previous data (Fig. 2.6B). Reducing the amount of DI plasmid by 80% relative to the corresponding full-length segment plasmid (0.2x group) resulted in a moderate reduction (around 1.0 - log) of viral infectivity when compared to controls (pHW vector and mock transfections). Furthermore, there was no detectable reduction in viral infectivity observed in 0.1x group where the PB1ΔA (Fig. 2.6C) or MΔA DI (Fig. 2.6D) plasmid for transfection was further reduced by 90% relative to the full-length segment plasmid.
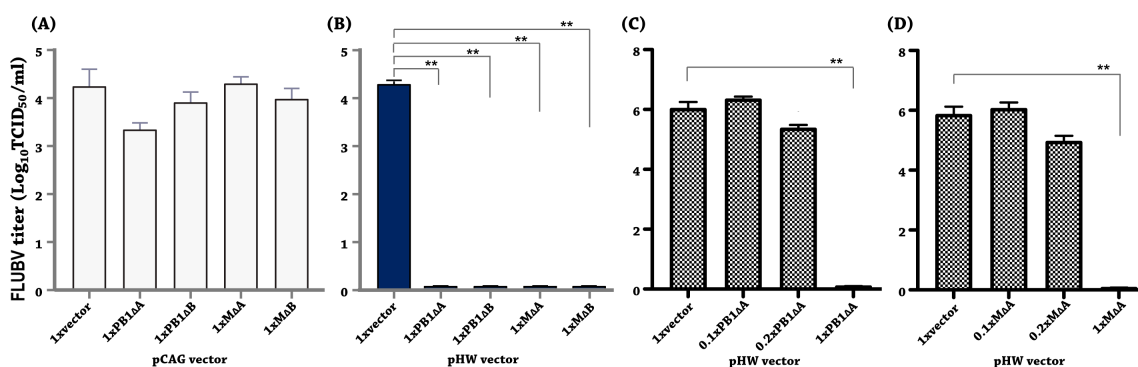
Figure 2.6. Inhibition of IBV replication by defective RNAs. Co-culture of HEK293T and MDCK cells were transfected with each of indicated defective genome plasmids (0.25 µg each) in the context of (A) pCAG vector (positive-sense mRNA production) or (B) pHW vector (both positive and negative-sense RNA production), together with an 8-

plasmid IBV RGS system (0.25 μg each plasmid). Culture supernatants collected at 24 hpi were titrated in MDCK cells to determine end-point titers (TCID$_{50}$/ml). The dose-dependent effects of PB1ΔA and MΔA genomes in interference with IBV replication (Panels C and D) were investigated under similar experimental conditions by decreasing the amount of the defective genome plasmid used for transfection while maintaining the same concentration of the corresponding parental segment. Three relative ratios of the defective genome and parental segment plasmids: 1x, 1:1 (0.25 to 0.25 μg); 0.2x, 1:5 (0.25 to 1.25 μg); and 0.1x, 1:10 (0.25 to 2.5 μg), were used together with seven other IBV RGS plasmids (0.25 μg each plasmid) in the transfection/infection experiment. The data are showed as mean±SEM (standard error of the mean) from three separate experiments in triplicate each. One-way ANOVA is used to calculate significance with one star indicating $p < 0.05$ and two stars indicating $P < 0.01$, respectively.

Next, we investigated whether expression of IBV-derived PB1ΔA and MΔA defective genomes could interfere with IAV replication using a similar approach as described above. For determining viral infectivity, indicator of defective RNAs' effects, we measured viral TCID$_{50}$/ml at 6, 12, and 24 h following infection of the MDCK cell line, respectively. Similarly, three different dose groups (based on the relative ratios of the defective genome and its parental segment plasmids) were used in this experiment, which were 1x, 1:1 (0.25 to 0.25 μg); 5x, 5:1 (1.25 to 0.25 μg); and 10x, 10:1 (2.5 to 0.25 μg). As demonstrated in Fig. 2.7, IAV replication was not overall substantially inhibited by either MΔA or PB1ΔA when used in the same amount of plasmid DNA as the corresponding full-length M or PB1 segment (1x group). Interestingly, significant

inhibitions of IAV replication were achieved when PB1ΔA plasmid was used 5 (5x group) or 10 times (10x group) higher than its parental PB1 segment (Fig. 2.7A). For example, in 10x group, no detectable viruses were found in the presence of PB1ΔA at 6, 12 and 24 hpi. In the 5x group, introduction of PB1ΔA resulted in a complete inhibition of IAV replication at 6 hpi followed by nearly 3-log and around 1-log reductions relative to mock vector control at 12 and 24 hpi, respectively (Fig. 2.7A). In a marked contrast, increasing the quantity of transfected MΔA plasmid showed no significant inhibitory activity against IAV (i.e., 5x and 10x group) over the course of this experiment (Fig. 2.7B).
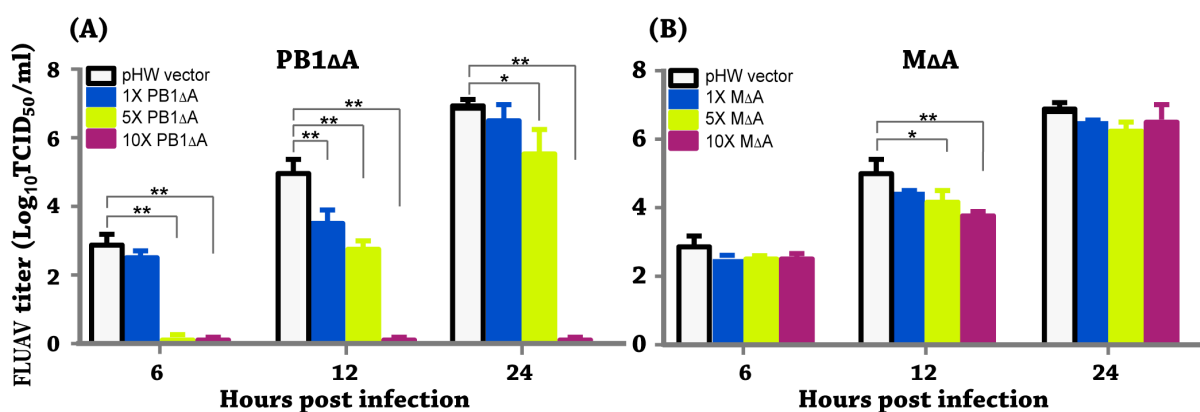
Figure 2.7. Inhibition of IAV replication by defective RNAs. Co-culture of HEF293T and MDCK cells was transfected with an 8-plasmid A/WSN/33 RGS in the absence or presence of different amounts of PB1ΔA (Panel A) and MΔA (Panel B). Note 1x: 0.25 μg defective RNA plasmid and 0.25 μg of each of RGS plasmids, 5x: 1.25 μg defective RNA plasmid and 0.25 μg of each of RGS plasmids, and 10x: 2.5 μg defective RNA plasmid and 0.25 μg of each of RGS plasmids. DNA amount in each transfection reaction was made equivalent by adding pHW empty vector. Culture supernatants collected at 6,

12, and 24 hpi were titrated for determining viral infectivity ($TCID_{50}$/ml) in MDCK cells. The data are showed as mean±SEM (standard error of the mean) from three separate experiments in triplicate each. One-way ANOVA is used to calculate significance with one star indicating $p<0.05$ and two stars indicating $P<0.01$, respectively.

In summary, the above experimental data demonstrated that all four defective genomes (PB1ΔA, PB1ΔB, MΔA and MΔB) identified from our NGS data potently inhibited IBV replication. Further investigation of PB1ΔA and MΔA showed that PB1ΔA, not MΔA, was able to interfere with IAV replication but the degree of inhibition was less pronounced when compared to that observed for IBV.

2.3.5 Cell type-independent phenomenon of IBV defective genomes production.

Finally, we were interested in determining whether the cell type played a decisive role in the production of the observed PB1ΔA, PB1ΔB, MΔA, and MΔB defective genomes (Fig. 2.8). In addition to the A549 cell line, we included Madin Darby Canine Kidney (MDCK) and human epithelial lung Calu-3 due to the following considerations. Firstly, MDCK cell line has been a constant source to study DI genomes of IAV. Secondly, Calu-3 is another relevant epithelia cell line to investigate the biology of influenza viruses. Total RNAs were extracted from viral supernatants collected at 24 hpi with 1.0 MOI of influenza B/Brisbane/60/2008, which was then followed by defective genome-specific RT-PCR amplification. After the agarose gel electrophoresis of RT-PCR products, the DNA bands of correct sizes were excised, purified and sequenced toward the identification of junction sites. As demonstrated in Fig. 2.8, all four defective

genomes with the expected correct junction site were produced in the three tested cell lines, which suggested a cell type-independent mechanism for IBV defective genomes production.
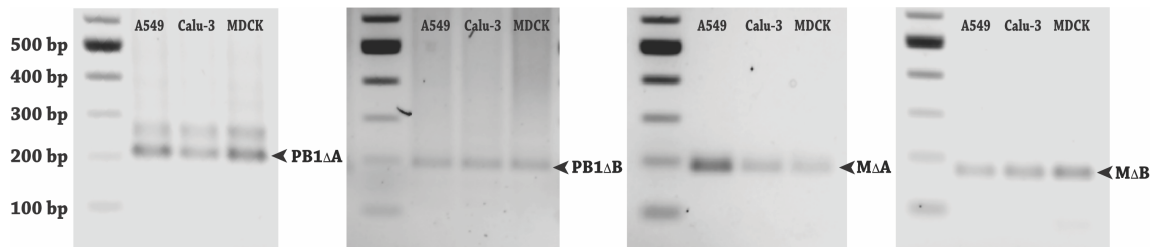
Figure 2.8. Cell type-independent production of IBV defective genomes. A549, Calu-3, and MDCK cells were infected, respectively, with 1.0 MOI of B/Brisbane/60/2008. Total RNAs were extracted from viral supernatants collected at 24 hpi, which was then followed by defective genome-specific RT-PCR amplification. After the agarose gel electrophoresis of RT-PCR products, the DNA bands of correct sizes were excised, purified and sequenced through the Sanger method toward the identification of junction sites. The data shown are representative of three independent experiments.

2.4 Discussion

DI genomes have been well described in natural IAV infections (A. R. Davis et al., 1980; A. R. Davis and D. P. Nayak, 1979; N. J. Dimmock and A. J. Easton, 2014, 2015; N. J. Dimmock et al., 2008; J. Michael Janda et al., 1979; D. P. Nayak et al., 1982; K. Saira et al., 2013). Despite earlier studies showing that DI genomes can arise from any of the eight segments during viral replication, it is generally agreed upon that the three

largest polymerase segments (PA, PB1, and PB2) are the primary donors generating these DI RNAs in cell cultures as well as in infected animals (N. J. Dimmock and A. J. Easton, 2014, 2015; K. Saira et al., 2013). The transmission of DI genomes has occurred between human patients infected with influenza A (H1N1) pdm09 virus (K. Saira et al., 2013). Because the position of the internal deletion is highly variable in the segment, a heterogeneous population of segment-specific DI genomes has been frequently observed in IAV-infected cells and animals (N. J. Dimmock et al., 2008; K. Saira et al., 2013). To date, more than 50 different species of DI RNAs have been described in IAV-infected cells (N. J. Dimmock and A. J. Easton, 2015; S. Noble and N. J. Dimmock, 1995). Although IBV DI phenomena has been reported in 1954 (P. Von Magnus, 1954), and the presence of substantial amount of putative IBV DIs in addition to IAV DIs in live attenuated influenza vaccine has been described recently (P. S. Gould et al., 2017), their roles in moderating viral replication and establishing host protective immunity is not clear.

Here we demonstrated that IBV, like IAV, is capable of producing defective genomes during viral replication. Specifically, by NGS coupled with RT-PCR and traditional Sanger sequencing, we have confirmed the presence of four defective genomes in A549 human lung epithilial cells that had been infected with B/Brisbane/60/2008. The stringent threshold (i.e., above 200 junction reads) employed in our sequence analysis may result in the detection of a relatively low number of defective genomes in IBV-infected A549 cells. An alternative explanation is our experimental conditions where no continuous virus passage or high infectious dose was employed (only 1 MOI used in this

study). Regardless of a precise cause, our data indicated that both PB1 and M segments likely serve as a major source for the production of defective RNAs upon IBV infection.

Each of the identified four defective genomes (PB1ΔA, PB1ΔB, MΔA, and MΔB), when transcribed into vRNA-sense, was able to potently inhibited the replication of IBV. The interfering activities of the defective genomes we had seen are similar to these of the well-characterized IAV DI genomes. Therefore, we propose that the four defective genomes should be categorized as DI genomes. Despite their unknown mechanism of action, by analogy with IAV, one can envision that the defective vRNAs inhibit the growth of IBV by competing with their corresponding full-length segment for the engagement of the viral polymerase complex during replication of vRNA into cRNA as well as for the final packaging into budding virus particles. Interestingly, only PB1ΔA had the ability to exert a moderate and dose-dependent interference effect, which was much less pronounced than that seen in the interference of cognate IBV replication (Figs. 2.6 and 2.7). Because PB1 is the most conserved vRNA segment between IAV and IBV, we speculate that the PB1 segment shares more similar features such as the panhandle structure (vRNA promoter for viral RNA polymerase complex formed by the 5' and 3' termini of each segment) and the packaging signal sequence (containing the 5' and 3' non-coding regions and the terminal coding sequences of each segment) than the M segment between IAV and IBV. Previous studies have shown that the panhandle structure and the packaging sequence play a decisive role in DI-mediated interference in viral replication (A. Baum et al., 2010; T. Frensing et al., 2014; J. M. Ngunjiri et al., 2013; M. Perez-Cidoncha et al., 2014). The sequence and structural differences between IBV PB1

and M segments, probably subtle, warrant further investigation toward addressing the mechanism by which DI RNA from IBV PB1, not M segment, interfere with IAV replication.

Presence of the highly abundant MΔA (i.e., 200-fold and 300-fold higher than the parental M segment, Fig. 2.5B) in both virus-infected cell and virus particles raised several interesting questions. First, the observed high abundance of MΔA in virus particles provides evidence that the majority of influenza B virions lacks the full-length parental M segment and are thus incapable of conducting autonomous replication. Overproduction of IBV M segment-derived DI genome may represent a major source that derives high particle-to-PFU (plaque formation unit) ratio as normally seen in animal viruses such as influenza A virus. Second, the presence of a predominant M segment-derived defective genome during active IBV replication may further discriminate IBV from IAV in defective genome production and M segment-centered expression strategy. Numerous studies of IAV have shown that the three polymerase segments, not the M segment, are the major forces driving the generation of DI genomes (A. R. Davis et al., 1980; Philip A. Jennings et al., 1983; S. Noble and N. J. Dimmock, 1995; K. Saira et al., 2013). Moreover, the M2 protein of IAV is translated from M2 mRNA, which is generated via alternative splicing (R A Lamb et al., 1981). However, IBV's M segment depends on a translational stop-start strategy to synthesize BM2 protein (C. M. Horvath et al., 1990). To our knowledge, no splicing event has been previously observed in IBV M segment. Intriguingly, the M segment in IAV is favored for splicing because three independent splicing events have been identified, while its counterpart in IBV is largely

selected for producing defective genomes. Further dissection of this notable difference may help us better appreciate the differences between these two closely related influenza viruses in biology and evolution. Third, why does IBV replication give rise to such abundant MΔA species? It can be assumed that MΔA is a perfect vRNA template that can be recognized and replicated efficiently by the viral RNA polymerase complex. This situation, if true, may consume substantial energy and resources, which, the viral replication machinery could otherwise save for manufacturing more infectious viral particles for subsequent infection. Alternatively, it can be proposed that MΔA abundance is intentionally required for efficient IBV replication and transmission. For example, it may behave as an accessory factor to facilitate viral replication or help IBV-infected cells recover from apoptosis or damage. Paradoxically, this model should reconcile the prevailing dogma that DI is a negative regulator of viral replication by competing out its full-length segment. An emerging role for DI such as antiviral innate immune response inducer should be also taken into consideration to address the potential opposing effects of DI in modulation of IBV replication *in vitro* and *in vivo* (A. Baum et al., 2010; T. Frensing et al., 2014; J. M. Ngunjiri et al., 2013; M. Perez-Cidoncha et al., 2014). Furthermore, excessive production of the defective genomes such as MΔA may allow IBV to employ it as a decoy molecule in misleading the anti-IBV innate response to help IBV evade immune recognition toward efficient replication and transmission in humans, which shall be invested in further studies.

Our studies also demonstrated that all four defective genomes initially found in infected A549 cells were also present in two other cell types, MDCK and Calu-3. This

result indicated that cell type might not have a direct role in IBV-specific defective genomes diversity. Despite decades of research, the molecular mechanism underlying DI RNA formation has not been fully elucidated. Current studies suggest that a viral RNA polymerase slippage-based faulty replication process is a likely mechanism (N. J. Dimmock and A. J. Easton, 2014). Heterologous populations of DI present in IAV and various species of segment-derived DIs in terms of the junction sites support this model. Our observation of over-abundant MΔA in IBV, however, may challenge this model. It will be interesting to address whether defective RNA production especially MΔA, is a specific mechanism or due to the viral RNA polymerase slippage-based faulty replication that often occurs randomly. In summary, we identified four DI genomes specific to IBV replication, with M segment-derived MΔA being the most abundant. The negative-sense defective vRNAs potently inhibit the replication of B/Yamanashi/166/98. PB1ΔA was able to modestly interfere with IAV replication. In addition, the productions of the four DIs are independent of cell types. Further characterization of these DI genomes should advance our understanding of the biology and evolution of IBV.

2.5 Acknowledgements

BIBLIOGRAPHY

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Baum, A., Sachidanandam, R., and Garcia-Sastre, A. (2010). Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. Proceedings of the National Academy of Sciences of the United States of America *107*, 16303-16308.

Bean, W. J., Kawaoka, Y., Wood, J. M., Pearson, J. E., and Webster, R. G. (1985). Characterization of virulent and avirulent A/chicken/Pennsylvania/83 influenza A viruses: potential role of defective interfering RNAs in nature. J Virol *54*, 151-160.

Boergeling, Y., Rozhdestvensky, T. S., Schmolke, M., Resa-Infante, P., Robeck, T., Randau, G., Wolff, T., Gabriel, G., Brosius, J., and Ludwig, S. (2015). Evidence for a Novel Mechanism of Influenza Virus-Induced Type I Interferon Expression by a Defective RNA-Encoded Protein. PLoS pathogens *11*, e1004924.

Briedis, D. J., and Lamb, R. A. (1982). Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J Virol *42*, 186-193.

Chambers, T. M., and Webster, R. G. (1987). Defective interfering virus associated with A/Chicken/Pennsylvania/83 influenza virus. J Virol *61*, 1517-1523.

Cox, N. J., and Subbarao, K. (2000). Global epidemiology of influenza: past and present. Annu Rev Med *51*, 407-421.

Davis, A. R., Hiti, A. L., and Nayak, D. P. (1980). Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proceedings of the National Academy of Sciences of the United States of America *77*, 215-219.

Davis, A. R., and Nayak, D. P. (1979). Sequence relationships among defective interfering influenza viral RNAs. Proceedings of the National Academy of Sciences of the United States of America *76*, 3092-3096.

Dimmock, N. J., and Easton, A. J. (2014). Defective interfering influenza virus RNAs: time to reevaluate their clinical potential as broad-spectrum antivirals? J Virol *88*, 5217-5227.

Dimmock, N. J., and Easton, A. J. (2015). Cloned Defective Interfering Influenza RNA and a Possible Pan-Specific Treatment of Respiratory Virus Diseases. Viruses *7*, 3768-3788.

Dimmock, N. J., Rainsford, E. W., Scott, P. D., and Marriott, A. C. (2008). Influenza virus protecting RNA: an effective prophylactic and therapeutic antiviral. J Virol *82*, 8570-8578.

Dubois, Julia, Terrier, Olivier, and Rosa-Calatrava, Manuel (2014). Influenza Viruses and mRNA Splicing: Doing More with Less. mBio *5*.

Easton, A. J., Scott, P. D., Edworthy, N. L., Meng, B., Marriott, A. C., and Dimmock, N. J. (2011). A novel broad-spectrum treatment for respiratory virus infections: influenza-based defective interfering virus provides protection against pneumovirus infection in vivo. Vaccine *29*, 2777-2784.

Frensing, T., Pflugmacher, A., Bachmann, M., Peschel, B., and Reichl, U. (2014). Impact of defective interfering particles on virus replication and antiviral host response in cell

culture-based influenza vaccine production. Applied microbiology and biotechnology *98*, 8999-9008.

Glezen, W. P. (2014). Editorial commentary: Changing epidemiology of influenza B virus. Clin Infect Dis *59*, 1525-1526.

Gould, P. S., Easton, A. J., and Dimmock, N. J. (2017). Live Attenuated Influenza Vaccine contains Substantial and Unexpected Amounts of Defective Viral Genomic RNA. Viruses *9*.

Gutierrez-Pizarraya, A., Perez-Romero, P., Alvarez, R., Aydillo, T. A., Osorio-Gomez, G., Milara-Ibanez, C., Sanchez, M., Pachon, J., and Cordero, E. (2012). Unexpected severity of cases of influenza B infection in patients that required hospitalization during the first postpandemic wave. The Journal of infection *65*, 423-430.

Hai, R., Schmolke, M., Varga, Z. T., Manicassamy, B., Wang, T. T., Belser, J. A., Pearce, M. B., Garcia-Sastre, A., Tumpey, T. M., and Palese, P. (2010). PB1-F2 expression by the 2009 pandemic H1N1 influenza virus has minimal impact on virulence in animal models. J Virol *84*, 4442-4450.

Hause, B. M., Collin, E. A., Liu, R., Huang, B., Sheng, Z., Lu, W., Wang, D., Nelson, E. A., and Li, F. (2014). Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. mBio *5*, e00031-00014.

Hause, B. M., Ducatez, M., Collin, E. A., Ran, Z., Liu, R., Sheng, Z., Armien, A., Kaplan, B., Chakravarty, S., Hoppe, A. D., Webby, R. J., Simonson, R. R., and Li, F. (2013). Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens *9*, e1003176.

Hoffmann, E., Mahmood, K., Yang, C. F., Webster, R. G., Greenberg, H. B., and Kemble, G. (2002). Rescue of influenza B virus from eight plasmids. Proceedings of the National Academy of Sciences of the United States of America *99*, 11411-11416.

Horvath, C. M., Williams, M. A., and Lamb, R. A. (1990). Eukaryotic coupled translation of tandem cistrons: identification of the influenza B virus BM2 polypeptide. The EMBO Journal *9*, 2639-2647.

Janda, J. Michael, Davis, Alan R., Nayak, Debi P., and De, Barun K. (1979). Diversity and generation of defective interfering influenza virus particles. Virology *95*, 48-58.

Jennings, Philip A., Finch, John T., Winter, Greg, and Robertson, James S. (1983). Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? Cell *34*, 619-627.

Kawakami, E., Watanabe, T., Fujii, K., Goto, H., Watanabe, S., Noda, T., and Kawaoka, Y. (2011). Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. J Virol Methods *173*, 1-6.

Lamb, R A, Lai, C J, and Choppin, P W (1981). Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proceedings of the National Academy of Sciences *78*, 4170-4174.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Lin, J. H., Chiu, S. C., Shaw, M. W., Lin, Y. C., Lee, C. H., Chen, H. Y., and Klimov, A. (2007). Characterization of the epidemic influenza B viruses isolated during 2004-2005 season in Taiwan. Virus Res *124*, 204-211.

McCullers, J. A., Facchini, S., Chesney, P. J., and Webster, R. G. (1999a). Influenza B virus encephalitis. Clin Infect Dis *28*, 898-900.

McCullers, J. A., and Hayden, F. G. (2012). Fatal influenza B infections: time to reexamine influenza research priorities. The Journal of infectious diseases *205*, 870-872.

McCullers, J. A., Saito, T., and Iverson, A. R. (2004). Multiple genotypes of influenza B virus circulated between 1979 and 2003. J Virol *78*, 12817-12828.

McCullers, J. A., Wang, G. C., He, S., and Webster, R. G. (1999b). Reassortment and insertion-deletion are strategies for the evolution of influenza B viruses in nature. J Virol *73*, 7343-7348.

Monto, A. S. (2008). Epidemiology of influenza. Vaccine *26 Suppl 4*, D45-48.

Nakagawa, N., Higashi, N., and Nakagawa, T. (2009). Cocirculation of antigenic variants and the vaccine-type virus during the 2004-2005 influenza B virus epidemics in Japan. J Clin Microbiol *47*, 352-357.

Nayak, D. P., Sivasubramanian, N., Davis, A. R., Cortini, R., and Sung, J. (1982). Complete sequence analyses show that two defective interfering influenza viral RNAs contain a single internal deletion of a polymerase gene. Proceedings of the National Academy of Sciences of the United States of America *79*, 2216-2220.

Nerome, R., Hiromoto, Y., Sugita, S., Tanabe, N., Ishida, M., Matsumoto, M., Lindstrom, S. E., Takahashi, T., and Nerome, K. (1998). Evolutionary characteristics of influenza B virus since its first isolation in 1940: dynamic circulation of deletion and insertion mechanism. Arch Virol *143*, 1569-1583.

Ngunjiri, J. M., Buchek, G. M., Mohni, K. N., Sekellick, M. J., and Marcus, P. I. (2013). Influenza virus subpopulations: exchange of lethal H5N1 virus NS for H1N1 virus NS

triggers de novo generation of defective-interfering particles and enhances interferon-inducing particle efficiency. Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research *33*, 99-107.

Noble, S., and Dimmock, N. J. (1995). Characterization of Putative Defective Interfering (DI) A/WSN RNAs Isolated from the Lungs of Mice Protected from an Otherwise Lethal Respiratory Infection with Influenza Virus A/WSN (H1N1): A Subset of the Inoculum DI RNAs. Virology *210*, 9-19.

Odagiri, T., and Tashiro, M. (1997). Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step. J Virol *71*, 2138-2145.

Osterhaus, A. D., Rimmelzwaan, G. F., Martina, B. E., Bestebroer, T. M., and Fouchier, R. A. (2000). Influenza B virus in seals. Science *288*, 1051-1053.

Paddock, C. D., Liu, L., Denison, A. M., Bartlett, J. H., Holman, R. C., Deleon-Carnes, M., Emery, S. L., Drew, C. P., Shieh, W. J., Uyeki, T. M., and Zaki, S. R. (2012). Myocardial injury and bacterial pneumonia contribute to the pathogenesis of fatal influenza B virus infection. The Journal of infectious diseases *205*, 895-905.

Palese, P., and Shaw, M. (2007). Orthomyxoviridae: The viruses and their replication. In Field's Virology, B.N. Fields, D.M. Knipe, and P.M. Howley, eds. (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins), pp. 1647-1689.

Paul Glezen, W., Schmier, J. K., Kuehn, C. M., Ryan, K. J., and Oxford, J. (2013). The burden of influenza B: a structured literature review. American journal of public health *103*, e43-51.

Perez-Cidoncha, M., Killip, M. J., Oliveros, J. C., Asensio, V. J., Fernandez, Y., Bengoechea, J. A., Randall, R. E., and Ortin, J. (2014). An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. J Virol *88*, 4632-4646.

Ran, Z., Shen, H., Lang, Y., Kolb, E. A., Turan, N., Zhu, L., Ma, J., Bawa, B., Liu, Q., Liu, H., Quast, M., Sexton, G., Krammer, F., Hause, B. M., Christopher-Hennings, J., Nelson, E. A., Richt, J., Li, F., and Ma, W. (2015). Domestic pigs are susceptible to infection with influenza B viruses. J Virol *89*, 4818-4826.

Robb, N. C., and Fodor, E. (2012). The accumulation of influenza A virus segment 7 spliced mRNAs is regulated by the NS1 protein. The Journal of general virology *93*, 113-118.

Saira, K., Lin, X., DePasse, J. V., Halpin, R., Twaddle, A., Stockwell, T., Angus, B., Cozzi-Lepri, A., Delfino, M., Dugan, V., Dwyer, D. E., Freiberg, M., Horban, A., Losso, M., Lynfield, R., Wentworth, D. N., Holmes, E. C., Davey, R., Wentworth, D. E., and Ghedin, E. (2013). Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. J Virol *87*, 8064-8074.

Scott, P. D., Meng, B., Marriott, A. C., Easton, A. J., and Dimmock, N. J. (2011). Defective interfering influenza A virus protects in vivo against disease caused by a heterologous influenza B virus. The Journal of general virology *92*, 2122-2132.

Shen, J., Kirk, B. D., Ma, J., and Wang, Q. (2009). Diversifying selective pressure on influenza B virus hemagglutinin. J Med Virol *81*, 114-124.

Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A. R., and Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. Nucleic acids research *34*, 3955-3967.

Vijaykrishna, D., Holmes, E. C., Joseph, U., Fourment, M., Su, Y. C., Halpin, R., Lee, R. T., Deng, Y. M., Gunalan, V., Lin, X., Stockwell, T. B., Fedorova, N. B., Zhou, B., Spirason, N., Kuhnert, D., Boskova, V., Stadler, T., Costa, A. M., Dwyer, D. E., Huang, Q. S., Jennings, L. C., Rawlinson, W., Sullivan, S. G., Hurt, A. C., Maurer-Stroh, S., Wentworth, D. E., Smith, G. J., and Barr, I. G. (2015). The contrasting phylodynamics of human influenza B viruses. eLife *4*, e05055.

Von Magnus, P. (1954). Incomplete forms of influenza virus. Advances in virus research *2*, 59-79.

Wise, H. M., Foeglein, A., Sun, J., Dalton, R. M., Patel, S., Howard, W., Anderson, E. C., Barclay, W. S., and Digard, P. (2009). A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. J Virol *83*, 8021-8031.

Wu, P., Goldstein, E., Ho, L. M., Yang, L., Nishiura, H., Wu, J. T., Ip, D. K., Chuang, S. K., Tsang, T., and Cowling, B. J. (2012). Excess mortality associated with influenza A and B virus in Hong Kong, 1998-2009. The Journal of infectious diseases *206*, 1862-1871.

Wu, T. D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics *26*, 873-881.

Chapter 3. Influenza D virus diverges from its related influenza C virus in the recognition of 9-O-acetylated N-acetyl- or N-glycolyl- neuraminic acid-containing glycan receptors

ABSTRACT

Novel influenza D virus (IDV) utilizes cattle as a primary reservoir with periodically spillover to other mammalian hosts including pigs and horses. Viral attachment to terminal sialic acids (SAs) of sialoglycans exposed on the cell surface is a determinant of tissue tropism and host range. Using the hemagglutination assay, we demonstrated that IDV, like its related influenza C virus (ICV), utilized 9-O-acetylated N-acetylneuraminic acid ($Neu5,9Ac_2$) for attachment to Turkey red blood cells (RBCs) and 9-O-acetylation group was a critical SA receptor determinant of IDV. Furthermore, with the sialylated glycan microarray (SGM), we found that IDV also interacted with 9-O-acetylated N-glycolylneuraminic acid (Neu5Gc9Ac) equally well, while ICV preferred $Neu5,9Ac_2$ over Neu5Gc9Ac. Finally, the glycan array data were supported by functional studies showing that both $Neu5,9Ac_2$ and Neu5Gc9Ac receptor analogs exhibited a dose-dependent inhibition of IDV replication as well as its ability to agglutinate RBCs. Similar results were obtained for ICV but the inhibition levels were much less pronounced with relatively more inhibition of ICV replication and attachment to RBCs by $Neu5,9Ac_2$ than Neu5Gc9Ac. Neu5Gc9Ac is different from $Neu5,9Ac_2$ only by an additional oxygen atom at the C5 position. The results of our experiments reveal that IDV and ICV diverge in communicating with both O-acetyl group at the C9 position and acetyl/glycolyl groups at the C5 position in terminal 9-carbon SAs. These findings will provide a framework for

further investigation towards better understanding of how newly found multiple-species-infecting IDV exploits natural 9-O-acetylated SA variations to expand its host range.

IMPORTANCE

Influenza D virus (IDV) utilizes cattle as a primary reservoir. Recently increased IDV outbreaks in pigs and serological evidence of IDV infection in humans have raised a concern about the potential of IDV adapting to humans. Viral attachment to terminal sialic acids on the cell surface serves as a first determinant of tissue tropism and host range. Our studies demonstrate that IDV is more efficient in recognizing both human $Neu5,9Ac_2$ and non-human Neu5Gc9Ac receptors than ICV, a ubiquitous human pathogen. ICV seems to strongly prefer human $Neu5,9Ac_2$ over non-human Neu5Gc9Ac. Humans cannot make Neu5Gc9Ac due to frame-shift mutations occurring in the CMP-Neu5Ac hydroxylase (CMAH) that synthesizes Neu5GC, the substrate of the Neu5Gc9Ac. Our findings provide evidence that IDV has acquired the unique ability to infect and transmit among agricultural animals that are enriched in Neu5Gc9Ac, in addition to pose a zoonotic risk to humans only expressing $Neu5,9Ac_2$.

3.1 Introduction

The *Orthomyxoviridae* family has three influenza genera, A, B, and C, which are classified according to antigenic differences in their nucleoprotein (NP) and matrix (M) proteins. The fourth genus of influenza, named influenza D, has been recently described (https://www.cdc.gov/flu/about/viruses/types.htm). Influenza D (IDV) represents a novel

type of virus, which is more closely related to influenza C (ICV) than influenza A (IAV) or influenza B (IBV) (B. M. Hause et al., 2013). IDV is unique among four influenza types in that it utilizes bovine as a primary reservoir and amplification host with periodically spillover to other mammalian hosts (M. F. Ducatez et al., 2015; L. Ferguson et al., 2015; E. Foni et al., 2017; B. M. Hause et al., 2014; B. M. Hause et al., 2013; M. Quast et al., 2015; E. Salem et al., 2017; S. K. White et al., 2016; S. L. Zhai et al., 2017). Since its first isolation from a pig with influenza-like symptoms in Oklahoma of the United States in 2011, IDVs have been found in cattle and swine populations in North America and Eurasia (M. F. Ducatez et al., 2015; L. Ferguson et al., 2015; E. Foni et al., 2017; B. M. Hause et al., 2014; B. M. Hause et al., 2013; W. M. Jiang et al., 2014; M. Quast et al., 2015; E. Salem et al., 2017; S. K. White et al., 2016; S. L. Zhai et al., 2017). Susceptibility to infection by this novel virus has also been demonstrated in sheep, goats, horses, camelids, guinea pigs, and ferrets (M. F. Ducatez et al., 2015; L. Ferguson et al., 2015; E. Foni et al., 2017; B. M. Hause et al., 2014; B. M. Hause et al., 2013; W. M. Jiang et al., 2014; M. Quast et al., 2015; E. Salem et al., 2017; C. Sreenivasan et al., 2015; S. K. White et al., 2016; S. L. Zhai et al., 2017). Of public health importance, serological evidence of IDV infections in humans has been demonstrated (B. M. Hause et al., 2013; S. K. White et al., 2016), and increased IDV outbreaks in pigs have been recently observed in China and Italy respectively (E. Foni et al., 2017; S. L. Zhai et al., 2017).

To date, IDVs have evolved into two genetic and antigenic lineages, namely D/OK and D/660, represented by well-characterized swine IDV (sIDV) D/OK (D/swine/Oklahoma/1334/2011) and bovine IDV (bIDV) D/660

(D/bovine/Oklahoma/660/2013) isolates (E. A. Collin et al., 2015). The recognition of terminal Sialic acids (SAs) of sialoglycans exposed on the cell surface is a first and critical step of the influenza virus lifecycle. Using a fluorescent dye labeled recombinant Hemagglutinin-Esterase-Fusion (HEF) protein of sIDV D/OK, a recent study found that sIDV HEF protein interacted with both α2,3 and α2,6-linked 9-O-acetylated SAs (H. Song et al., 2016), suggesting that sIDV likely utilizes them as its receptors for virus entry. Furthermore, this study resolved the sIDV HEF crystal structure, and showed that sIDV differed from its related ICV in that its HEF protein possessed an open receptor-binding cavity (H. Song et al., 2016), which may allow IDV to accommodate more diverse glycan moieties containing a 9-O-acetly group and as a result increase its tissue and species tropisms. ICV utilizes 9-O-acetylated N-acetylneuraminic acid (Neu5,9Ac$_2$) with either α2,3 or α2,6 linkage as a functional receptor (P. B. Rosenthal et al., 1998). ICV has been thought to have a narrow tissue tropism and a limited host range (Y. Matsuzaki et al., 2016).

Despite the progress, little is known about the functional relevance of identified 9-O-acetylated Neu5Ac (Neu5,9Ac$_2$) and Neu5Gc (Neu5Gc9Ac) (H. Song et al., 2016), putative candidate receptors of IDV, in the context of viral infection. Receptor biology of the bovine D/660 lineage has remained uncharacterized. As such, further investigation into the receptor biology of IDV and biological relevance of various SA forms carrying a 9-O-acetyl group in IDV infection is needed because it will offer novel insights into how influenza viruses such as IDV originating from a novel bovine reservoir frequently spill over and transmit to new hosts including pigs and humans.

3.2 Materials and methods

3.2.1 Viruses and cells

Madin-Darby Canine Kidney (MDCK) cells (ATCC) were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (FBS) at 37 °C with 5% CO_2. IAV (H1N1 A/WSN/1933 virus), IBV (B/Brisbane/60/2008), ICV (C/Johannesburg/1/66), IDVs (D/swine/Oklahoma/1334/2011 and D/bovine/Oklahoma/660/2013), used for the Hemagglutiniation or the cell-based inhibition assay, were propagated in MDCK cells with DMEM containing 1ug/ml Tosyl phenylalanyl chloromethyl ketone (TPCK)-trypsin (Sigma-Aldrich). ICV (C/Johannesburg/1/66) and bIDV (D/bovine/Oklahoma/660/2013) used for the glycan array were propagated in 9 day-old embryonated specific-pathogen-free (SPF) chicken eggs.

3.2.2 Sialic acids removal assay

1 ml 5% Turkey red blood cells (RBCs) (Lampire) were treated with 200 milliunits (mU) Neuraminidase from *Clostridium perfringens* (Sigma-Aldrich) at 37 °C for 1 hour (h). After 1 h treatment, RBCs were washed with PBS for 3 times and then diluted to a final 1% concentration in PBS for the hemagglutination (HA) assay. 25 μl of each virus with four HA units were mixed with same amount of Neuraminidase-treated RBCs, followed by 1 h incubation at 4 °C before observing results. Non-treated RBCs were used as control.

3.2.3 Hemagglutination (HA) assay-based competitive inhibition assay

Bovine submaxillary mucin (BSM) (Sigma-Aldrich), synthetic 5-N-acetyl-9-O-acetyl neuraminic acid (Neu5,9Ac$_2$) and 5-N-acetyl-4-O-acetyl neuraminic acid (Neu4,5Ac$_2$) (Applied Biotech, Austria) were added to different types of viruses containing four HA units each. Mixtures were incubated for 30 min at room temperature. 25 µl of Turkey RBCs were added to the mixtures followed by reading results after 30 min incubation at room temperature.

3.2.4 9-O acetylated group removal assay

20 ug pure, synthetic Neu5,9Ac$_2$ (Applied Biotech) were treated with 5 mU sialate-9-O-acetylesterase (9-O-SE) (Applied Biotech) at 37 °C for 3 h. Four HA units of different types of viruses were incubated with 9-O-SE treated or non-treated Neu5,9Ac$_2$ for 30 min at room temperature. Then 25ul of Turkey RBCs were added and results were observed after 30 min incubation at room temperature. All HA assays were performed in three independent assays with each in triplicate.

3.2.5 Cell-based inhibition assay by receptor analogs (Digital Droplet PCR)

25 µl IDV and ICV containing 100 TCID50 were respectively pretreated with an equal volume of Neu5,9Ac2, Neu5Gc9Ac and Neu5Ac receptor analogs at 20, 80 and 320 ng/µl concentrations for 30 min at 4 °C. MDCK cells in 96-well plates were washed once with PBS and then infected with virus-receptor analogue mixtures for 1h. Cells were washed three times and further incubated in DMEM containing 1 µg/ml TPCK-trypsin for 6 h. After 6 h post infection (hpi), supernatants were removed and cells in each well were lysed with 200 µl Trizol (Life Technologies). Total RNAs were extracted according

to manufacturer's instructions, which were then followed by reverse transcription reactions with oligo(dT) primer and High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). NP segment-derived mRNA molecules of ICV and IDV were selected as our target for determining the effects of various receptor analogs on viral replication, while the mRNA of TATA-Box binding protein (TBP) of canine was used as the reference gene for PCR data normalization. The detailed information for primers and probes is provided in Table 3.5. The ddPCR reaction consisted of 10 µl 2x Supermix for Probes (Bio-Rad), 900 nM primers, 25 nM probes and 8 µl undiluted cDNA into a final volume of 20 µl. All samples were tested in three independent experiments with each assayed in triplicate. Non- template controls (NTC) were included in every run. Viral mRNA copies were normalized with TBP references and results were reported as NP mRNA copies per million TBP.

Table 3.5. Primers and probes used for digital droplet PCR.

Target genes	Primers/probes name	Primers/probes sequence
	qD660NP-For	5'-TGCCGATTGGTGGAGTCAA-3'
IDV NP	qD660NP-Rev	5'-TTTCAGTGCCATTCCCAATCT-3'
	qD660NP-Probe	5'-6FAM-AGCTGGGAAATGTAGTGC-MGBNFQ-3'
	qJHBNP-For	5'-TGAAGCCTACATTGCCATTTGT-3'
ICV NP	qJHBNP-Rev	5'-GCCATTTTCCAGGATCAACATT-3'
	qJHBNP-Probe	5'-6FAM-AGGAAGTGGGCCTTAA-MGBNFQ-3'
	Canine-TBP-For	5'-AGGATGATCAAACCCAGAATTGTT-3'
Canine TBP	Canine-TBP-Rev	5'-GCCCTTTAGAATAGGGTAGATGTTTTC-3'
	Canine-TBP-Probe	5'-VIC-TTGTACTAACAGGTGCTAAAG-MGBNFQ-3'

3.2.6 Virus labeling and sialylated glycan microarray (SGM)

IDV and ICV were propagated in 9 day-old embryonated SPF chicken eggs. Allantonic fluid were harvest at 3 dpi followed by a brief centrifugation at 2,000 x g for 20 min at 4°C. The clarified supernatant was layered over a 20% sucrose cushion in PBS buffer and then ultracentrifuged at 100,000 x g for 2h at 4 °C in a SW28 rotor (Beckman Coulter). Pellets were resuspended in PBS and further purified over a 30–60% cold sucrose gradient at 28,000 rpm for 4h at 4 °C in an SW 28 rotor. Fractions (1.0 ml) were collected from the top and all fractions showing greater or equal to 6 log 2 HA units were pooled together, followed by a final ultracentrifugation through a 20% sucrose cushion (w/v) in PBS at 100,000 g for 2h at 4 °C. The pellets were resuspended in CMS buffer (0.15M NaCl, 0.25mM CaCl2, 0.8 mM MgCl2, PH 7.4) and diluted to a solution containing 1.0×10^5 HAU (HA unit)/ml.

We employed the standard protocol to generate dye-labeled IDV and ICV virions as described previously (X. Song et al., 2011). In brief, 10 μl of 1.0 M sodium bicarbonate (pH 9.0) was added into 100 μl purified viruses containing ~1.0×10^4 HA units) followed by addition of Alexa Fluor-488 succinimidyl ester (Molecular Probes) in a ratio of 0.005 μg Alexa per HAU. This ration was determined by the HA titration experiment to give maximal labeling without loss of binding activity. After stirring for 1 h at room temperature in the dark, the labeling reaction mixtures were dialyzed (Slide-A-Lyzer Mini Dialysis Units 7000 MWCO, Pierce) in CaMgS (0.25 mM CaCl2, 0.8 mM MgCl$_2$ in borate buffered saline, pH 7.2) at 4°C overnight. After functional evaluation (HA activity), labeled viruses were ready for glycan receptor binding experiment.

To determine the glycan-binding specificities of labeled IDV and ICV, a glycan screen array was performed at the Consortium for Functional Glycomics. The printed glycan array consists of 77 sialylated glycans incorporating 16 modified sialic acids in α2–3 and α2–6 linkages to different underlying structures (X. Song et al., 2011). Briefly, fluorescently labeled virus was incubated on a glycan microarray slide under a coverslip at 4°C for 1 h, washed to remove unbound virus, and scanned using a ProScan Array (Perkin-Elmer Life sciences) equipped with multiple lasers. The data were processed using the manufacturer's software as described previously, which provided raw values in relative fluorescence units (RFU) from each spot in an Excel spreadsheet. The intensity of binding to each of the 77 glycans on the array was graphed and shown as values representing means ± S.D.s of six replicate samples.

3.3 Results

3.3.1 9-O-acetylation group is a critical sialic acid receptor determinant of IDV.

We first employed the traditional hemagglutination (HA) assay-based competitive inhibition approach to determine the glycan receptor of IDV. To determine whether sialic acids serve as a component of the receptor for IDV, we treated Turkey RBCs with neuraminidase from *Clostridium perfringens*. As summarized in Table 3.1, pretreatment with neuraminidase resulted in a complete loss of RBC agglutination mediated by IDV as well as by IAV A/WSN/1933 (H1N1) or IBV (B/Brisbane/60/2008) or ICV (C/Johannesburg/1/66) (4 HA units used for each virus per reaction). This result indicated that sialic acids are involved in IDV-mediated RBC agglutination.

Table 3.1. The effects of Neuraminidase (NA) treatment on viral ability in agglutination of Turkey red blood cells.

Virus	Untreated RBCs	NA-treated RBCs
IAV A/WSN/1933 (H1N1)	-[a]	+[b]
IBV B/Brisbane/60/2008	-	+
ICV C/Johannesburg/1/66	-	+
IDV D/bovine/Oklahoma/660/2013	-	+

a indicates evident hemagglutination (no inhibitory effect) in the wells.

b denotes non-hemagglutination (inhibitory effect) in the wells.

Agglutination of RBCs by IDV was also inhibited in the presence of a range of concentration of bovine submaxillary mucin (BSM) (from 6.25 to 400 ng), which is rich in 9-O-acetylayed Neu5Ac (Neu5,9Ac2) and Neu5Gc (Neu5Gc9Ac) (Table 3.2) (M. A. Langereis et al., 2015). Intriguingly, the level and pattern of inhibition of IDV-mediated RBC agglutination by BSM were similar to that observed in ICV that is known to utilize $Neu5,9Ac_2$-containing glycan as a receptor for viral entry (P. B. Rosenthal et al., 1998), thereby suggesting that IDV likely utilizes $Neu5,9Ac_2$ molecule as a receptor for attachment.

Table 3.2. The effects of bovine submaxillary mucin (BSM) treatment on viral ability in agglutination of Turkey red blood cells.

Virus	BSM (ng)				
	PBS	6.25	25	100	400
IAV	-b	-	-	-	-
IBV	-	-	-	-	-
ICV	-	-	+c	+	+
IDV	-	-	+	+	+
DMEM	-	-	-	-	-

a viruses used in this experiment were IAV IAV A/WSN/1933 (H1N1), IBV (B/Brisbane/60/2008), ICV (C/Johannesburg/1/66), and IDV (D/bovine/Oklahoma/660/2013) (4 HA units used for each virus per reaction).

b indicates evident hemagglutination (no inhibitory effect) in the wells.

c denotes non-hemagglutination (inhibitory effect) in the wells.

To further confirm these results, we performed the HA assay involving a synthetic Neu5,9Ac$_2$. For a control, Neu4,5Ac$_2$ was used, which has O-acetyl group in the C4, not C9 position. RBC agglutination by IDV and ICV was not inhibited by Neu4,5Ac$_2$ (up to 102 µg/per well, the maximum concentration used). In contrast, agglutination started to be completely lost for IDV at the presence of 1.6 µg Neu5,9Ac$_2$ per well and for ICV at 6.4 µg Neu5,9Ac$_2$ per well (Table 3.3). This result further demonstrated that Neu5,9Ac$_2$ is a sialic acid receptor of IDV. It is interesting to note that the minimal concentration to abolish ICV-mediated RBC agglutination was 4 times higher than that needed for IDV.

Table 3.3. The effects of synthetic Neu5,9Ac$_2$ and Neu4.5Ac$_2$ receptor analogs on viral ability in agglutination of Turkey red blood cells.

Virus[a]	Neu5,9Ac$_2$ (ng/ul)					
	PBS	5	20	80	320	1280
IAV	-[b]	-	-	-	-	-
IBV	-	-	-	-	-	-
ICV	-	-	-	+[c]	+	+
IDV	-	-	+	+	+	+
DMEM	-	-	-	-	-	-

Virus[a]	Neu4,5Ac$_2$ (ng/ul)					
	PBS	5	20	80	320	1280
IAV	-[b]	-	-	-	-	-
IBV	-	-	-	-	-	-
ICV	-	-	-	-	-	-
IDV	-	-	-	-	-	-
DMEM	-	-	-	-	-	-

[a] viruses used in this experiment were IAV IAV A/WSN/1933 (H1N1), IBV (B/Brisbane/60/2008), ICV (C/Johannesburg/1/66), and IDV (D/bovine/Oklahoma/660/2013) (4 HA units used for each virus per reaction).

[b] indicates evident hemagglutination (no inhibitory effect) in the wells.

[c] denotes non-hemagglutination (inhibitory effect) in the wells.

Finally, pretreatment of the synthetic Neu5,9Ac$_2$ with recombinant sialate-9-O-acetylesterase (9-O-SE) (e.g., removes 9-O-acetylated group) abolished its inhibitory effect of IDV- or ICV- mediated RBC agglutination, whereas binding and agglutination of RBCs by IAV were not affected (Table 3.4). In summary, the results of our qualitative HA-based experiments suggest that IDV uses Neu5,9Ac$_2$ as a receptor for infection.

Table 3.4. The effects of synthetic Neu5,9Ac2 receptor analog and recombinant sialate-9-O-acetylesterase (9-O-SE) pretreated Neu5,9Ac2 on viral ability in agglutination of Turkey red blood cells.

Virus	PBS	Neu5,9Ac2	Neu5,9Ac2 with 9-O-SE	9-O-SE with PBS
IAV	-b	-	-	-
ICV	-	+c	-	-
IDV	-	+	-	-

a viruses used in this experiment were IAV IAV A/WSN/1933 (H1N1), ICV (C/Johannesburg/1/66), and IDV (D/bovine/Oklahoma/660/2013) (4 HA units used for each virus per reaction).

b indicates evident hemagglutination (no inhibitory effect) in the wells. c denotes non-hemagglutination (inhibitory effect) in the wells.

3.3.2 Receptor binding characteristics of labeled IDV and ICV on a sialylated glycan microarray (SGM).

In parallel, we utilized sialylated glycan microarray technology to examine the specific glycans that serve as potential receptors for IDV and ICV. The SGM provided by the Consortium for Functional Glycomics consists of 77 sialylated glycans incorporating 16 modified sialic acids in α2,3 and α2,6 linkages to different underlying structures (X. Song et al., 2011). Under an identical condition, we found that IDV preferred to bind glycans

terminated in either Neu5,9Ac2 (IDs#38 and 52) or Neu5Gc9Ac (IDs#39 and 53) (Fig.

3.1A). Other forms of glycans failed to achieve significant interactions with IDV,

indicating a specific binding directed largely by the 9-O-acetyl group. Binding of IDV to

9-O-acetylated glycans was not dependent on the specific linkage (α2,3 or α2,6) (IDs#39

Vs. 53 and 38 Vs. 52) (Fig. 3.1C). ICV array did not give clear data (low signal/noise

ratio). Nevertheless, only #38 glycan terminated with Neu5,9Ac2 (α2,6 linkage) showed

consistent interactions with ICV (Fig. 3.1B) in our repeated SGM experiments.

Neu5,9Ac2 has been shown previously as a receptor determinant of ICV infection (G. N.

Rogers et al., 1986). In summary, the glycan array-based data are consistent with the

observations of the HA-based experiments, thereby further demonstrating that IDV

utilize 9-O-acetylated SAs-containing glycans as receptors for viral entry and infection.

(A)

SialyI84S 01635 5/9/2015
Block-10 Alexa488 D/bovine/Oklahoma/660/2013 (IDV) (CFG3181)

(B)

SialyI84S 01635 5/9/2015
Block-10 Alexa488 C/Johanneburg/1/66 (ICV) (CFG3181)
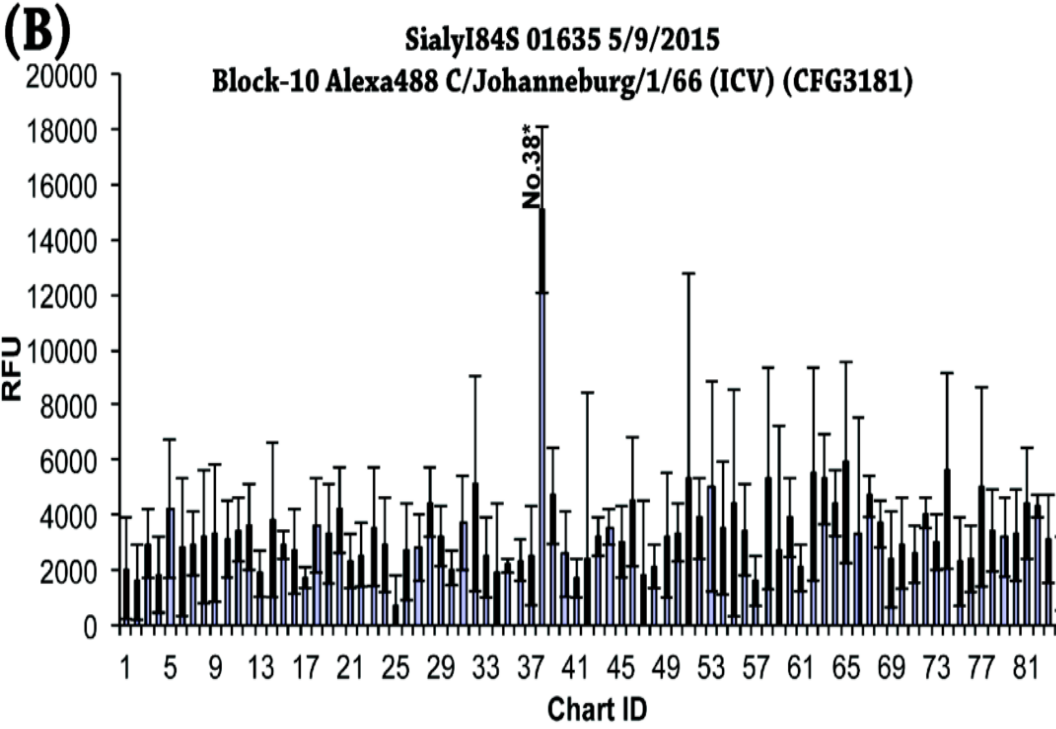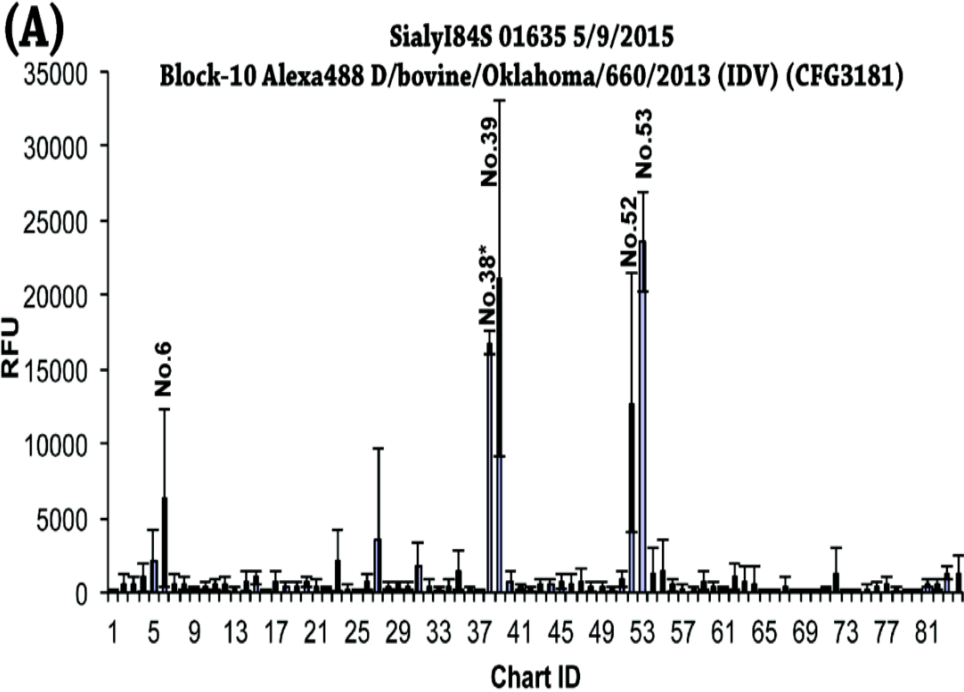
(C) CFG3181 glycan chart IDs and structures

ID	Structure
53	Neu5Gc9Acα3Galβ4GlcNAcβ2Manα3(Neu5Gc9Acα3Galβ4GlcNAcβ2Manα6)Manβ4GlcNAcβ4GlcNAcitol-AEAB
39	Neu5Gc9Acα6Galβ4GlcNAcβ2Manα3(Neu5Gc9Acα6Galβ4GlcNAcβ2Manα6)Manβ4GlcNAcβ4GlcNAcitol-AEAB
38*	Neu5,9Ac2α6Galβ4GlcNAcβ2Manα3(Neu5,9Ac2α6Galβ4GlcNAcβ2Manα6)Manβ4GlcNAcβ4GlcNAcitol-AEAB
52	Neu5,9Ac2α3Galβ4GlcNAcβ2Manα3(Galβ4GlcNAcβ2Manα6)Manβ4GlcNAcβ4GlcNAcitol-AEAB
6	Neu5,9Gc2α6Galβ4GlcNAcβ3Galβ4Glcitol-AEAB

Figure 3.1. Identification of 9-O-acetylated sialic acid receptors of influenza D virus. Glycan binding analyses as measured by relative fluorescence intensity (RFU) are presented for labeled IDV (A) and ICV (B) respectively. The structures of the sialylated glycans showing a significant interaction with IDV are listed in (C) with a "*" sign indicating a glycan that binds to both IDV and ICV. In brief, Alexa488 fluorescence-labeled purified IDV or ICV was incubated at 4°C for 1 h on the printed glycan array that consists of 77 sialylated glycans incorporating 16 modified sialic acids in α2,3 and α2,6 linkages to different underlying structures. Following washing six times to remove unbound viruses, the slides were scanned using a ProScan Array (Perkin-Elmer Life sciences) equipped with four lasers covering an excitation range from 488 to 637 nm. For Alexa488 fluorescence, 495 nm (excitation) and 510 nm (emission) were used. The data were processed using the manufacturer's software, which provided raw values in relative fluorescence units (RFU) from each spot in an Excel spreadsheet. The intensity of

binding to each of the 77 glycans on the array was graphed and shown as values representing means ± S.D.s of the four replicates. Note that an equal amount of purified IDV or ICV (500 HA units) was used for the glycan array analysis. The data presented in this figure are representative of two independent glycan microarray experiments performed in the four replicates.

3.3.3 Functional studies of the roles of Neu5,9Ac$_2$- or Neu5Gc9Ac-containing glycans in IDV and ICV infection. To determine the functional relevance of these receptor candidates, we first tested four synthetic glycans and examined their ability to disrupt the agglutination of Turkey RBCs by IDV and ICV in hemagglutination inhibition format where we replaced antibody with receptor analog with four hemagglutination (HA) unit-working concentration used for each virus per reaction. These four glycans are Neu5Ac-α2-3LNnT-beta-Propyl N3, Neu5Gc-α2-3LNnT-beta-Propyl N3, Neu5Ac9Ac-α2-3LNnT-beta-Propyl N3, and Neu5Gc9Ac-α2-3LNnT-beta-Propyl N3**.** As shown in Fig. 3.2A, the presence of both Neu5,9Ac$_2$ and Neu5Gc9Ac at a concentration of 20 ng/ul and above resulted in a complete loss of RBC agglutination by IDV. In contrast, neither Neu5Ac nor Neu5Gc affected IDV-mediated RBC agglutination, thereby confirming our glycan array data. Interestingly, we noticed that ICV needed 4 and 16 times higher concentrations of Neu5,9Ac$_2$ and Neu5Gc9Ac, respectively, than IDV for completely losing RBC agglutination. This result suggests that IDV binds both 9-O-acetylated SAs with an affinity higher than ICV, and ICV prefers Neu5,9Ac$_2$ over Neu5Gc9Ac.

In parallel, four synthetic glycans were examined for their ability to inhibit the replication of ICV (Fig. 3.2B) and IDV (Fig. 3.2C) in a MDCK cell-based multiple round replication assay format. After pretreatment for 30 min at 4 ºC, MDCK cells were infected with 100 $TCID_{50}$ of IDV or ICV, followed by additional 6h incubation. Extracted cellular RNAs were subjected to reverse transcription with oligo (dT) followed by Bio-Rad QX200 Droplet Digital ddPCR (targeting NP mRNA). Copy numbers of NP-specific mRNA in each reaction were normalized to internal reference gene encoding TATA box binding protein (TBP) to derive relative copy number of viral NP mRNA. The detailed information for primers and probes is provided in Table 3.5. As shown in Fig. 3.2C, both $Neu5,9Ac_2$ and Neu5Gc9Ac exhibited a dose-dependent inhibition of IDV replication. For example, IDV infectivity was reduced approximately by 50% with $Neu5,9Ac_2$ and by 40% with Neu5Gc9Ac at 64 ng/µl, respectively, whereas the infectivity was reduced by 90% with either treatment at 1024 ng/µl. Similarly, a dose-dependent inhibition of ICV was observed by the above receptor analogs but the inhibition levels were less pronounced (Fig. 3.2B). ICV replication was inhibited significantly more by $Neu5,9Ac_2$ than Neu5Gc9Ac (Fig. 3.2B). The results from cell-based and HA-based inhibition assays were in good agreement with each other, thereby further validating our findings from glycan array experiments. In summary, our data suggest that IDV binds both $Neu5,9Ac_2$ and Neu5Gc9Ac equally well, while ICV prefers $Neu5,9Ac_2$ over Neu5Gc9AC. Our results also indicate that IDV is more efficient in recognizing $Neu5,9Ac_2$ and Neu5Gc9Ac receptors than ICV.

(A)

Virus	Neu5Ac (ng/μl)					Neu5Gc (ng/μl)					Neu5,9Ac2 (ng/μl)					Neu5Gc9Ac (ng/μl)				
	1.25	5	20	80	320	1.25	5	20	80	320	1.25	5	20	80	320	1.25	5	20	80	320
ICV	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	+
IDV	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+	+	+
DMEM	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

(B)

(C)
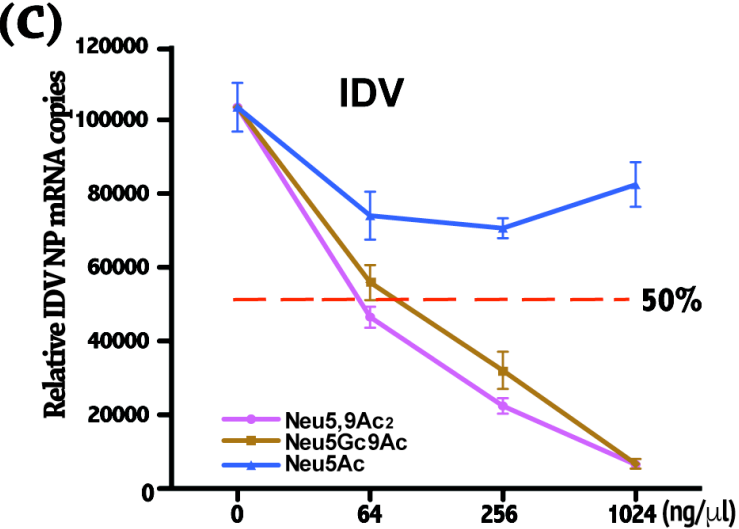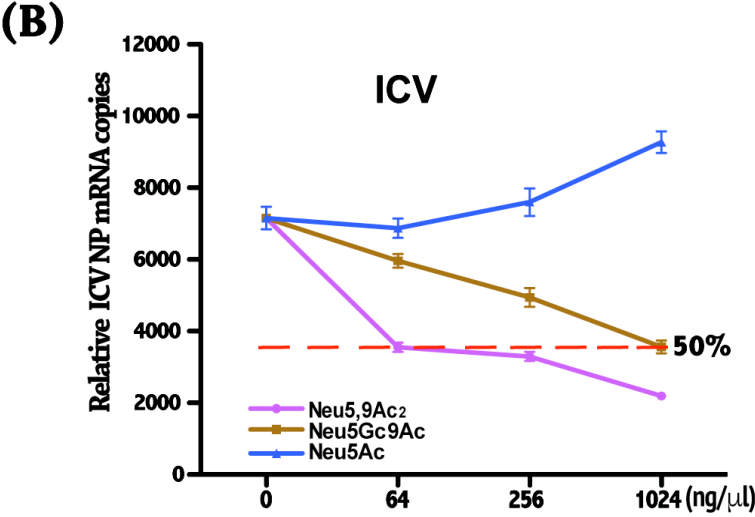
Figure 3.2. Inhibition of viral hemagglutination and infection by receptor analogs.
(A) A loss of Turkey RBC agglutination by ICV or IDV at the presence of synthetic receptor analog. Indicated receptor analogs at various concentrations (1.25, 5, 20, 80, 320 ng/ul) were added to IDV or ICV containing four HA units each. Mixtures were incubated for 30 min at room temperature. 25 µl of Turkey RBCs were then added to the mixtures followed by reading results after 30 min incubation at room temperature. Normal medium DMEM was used as a negative control. Note that "+" sign indicates no hemagglutination, while "-" denotes evident hemagglutination. The data presented in panel A are representative of four independent experiments performed in duplicate. Inhibition of ICV (B) and IDV (C) replication in an MDCK-based replication assay by receptor analogs as determined by Digital Droplet PCR. 25 µl IDV and ICV containing 100 TCID50 were respectively pretreated with an equal volume of Neu5,9AC$_2$ or Neu5Gc9Ac or Neu5Ac receptor analogs at 20, 80 and 320 ng/µl concentrations for 30 min at 4 °C. MDCK cells in 96-well plates were washed once with PBS and then infected with virus-receptor mixtures for 1h. Cells were washed three times and further incubated in DMEM containing 1 µg/ml TPCK-trypsin for 6 h. After 6 h post infection (hpi), supernatants were removed and cells in each well were lysed with 200 µl Trizol. Total RNAs were extracted according to manufacturer's instructions, which were then followed by reverse transcription reactions with oligo(dT) primer and High-Capacity cDNA Reverse Transcription Kit. NP segment-derived mRNA molecules of ICV and IDV were selected as our target for determining the effects of various receptor analogs on viral replication, while the mRNA of TATA-Box binding protein (TBP) of canine was used as the reference gene for PCR data normalization. The detailed information for primers and

probes is provided in Table 3.5. The ddPCR reaction consisted of 10 µl 2x Supermix for Probes (Bio-Rad), 900 nM primers, 25 nM probes and 8 µl undiluted cDNA into a final volume of 20 µl. Non- template controls (NTC) were included in every run. Viral mRNA copies were normalized with TBP references and results were reported as NP mRNA copies per million TBP. Receptor analog-mediated inhibitory effects were shown for ICV in panel B and for IDV in panel C as values representing means ± S.D.s of three independent experiments performed in triplicate.

3.4 Discussion

Neu5Gc and Neu5Gc9Ac are generally abundant in many agricultural animals such as cattle and swine (A. N. Samraj et al., 2015), but are absent in humans and ferrets due to frame-shift mutations of CMAH that converts Neu5Ac to Neu5Gc (A. Irie et al., 1998; P. S. Ng et al., 2014), the substrate of the Neu5Gc9Ac. The efficient usage of either Neu5,9Ac$_2$ or Neu5Gc9Ac as a receptor likely gives IDV an ecological niche to infect multiple agricultural animals with abundant expression of Neu5Gc9Ac as well as humans only expressing Neu5,9Ac$_2$. The zoonotic risk to humans by IDV is further supported by our previous study showing that humans-like ferrets are susceptible to IDV infection (C. Sreenivasan et al., 2015). It should be noted that despite the genetic inability to synthesize Neu5Gc and its derivative Neu5Gc9Ac, humans still possess the cellular 9-O-acetylation machinery for the synthesis of 9-O-acetylated sialic acids. In this regard, humans can take Neu5Gc from dietary sources, convert it into Neu5Gc9Ac, and express it on the cell surfaces of various cell types (M. Bardor et al., 2005). It has been shown previously that Neu5Gc and possible its 9-O-acetylated form (Neu5Gc9Ac) can accumulate in human cancer cells with an unknown mechanism (S. Inoue et al., 2010). In

light of the fact that exogenous expression non-human Neu5Gc can make humans vulnerable to pathogens utilizing Neu5GC as a receptor for infection (E. Byres et al., 2008), it will be interesting to address in the near future whether humans or ferrets expressing exogenous Neu5Gc9Ac under certain conditions are more susceptible to IDV infection.

Humans are thought to be the primary host and reservoir of ICV, although this virus has been identified in other hosts probably after reverse zoonotic transmission from humans (Y. Matsuzaki et al., 2016). Our observation that ICV preferentially uses Neu5,9Ac$_2$ over Neu5Gc9Ac appears to support this theory. First, as discussed above, the evolutionary loss of CMAH gene for conversion of Neu5Ac to Neu5Gc may make humans exclusively and abundantly express Neu5Ac and Neu5,9Ac$_2$ (A. Irie et al., 1998). As such humans become a perfect host for ICV that selectively prefers Neu5,9Ac$_2$ for viral attachment and infection. Second, agricultural animals such as cattle and pigs that are rich in Neu5Gc can express high levels of Neu5Gc9Ac (A. N. Samraj et al., 2015). Because a significant portion of Neu5Ac has been converted to Neu5Gc in these animals, the number of Neu5Ac molecule available for Neu5,9Ac$_2$ synthesis are substantially reduced. Reduced expression of the Neu5,9Ac$_2$ receptor may render agricultural animals become less susceptible to ICV infection when compared to humans. In case that these animals would be infected by ICV after reverse zoonotic transmission from humans, it is questionable that ICV transmission can be sustained among them. On the other hand, IDV differs from ICV in that it effectively engages with no clear preference both Neu5,9Ac$_2$ and Neu5Gc9Ac receptors. One can envision that the differential expression

levels of these two 9-O-acetylayed sialic acids in animals and humans may not substantially affect IDV replication and spread. And therefore, this unique receptor binding property should enable IDV effectively infect and transmit among different mammalian hosts including humans.

9-O-acetylated derivatives (Neu5,9Ac$_2$ and Neu5Gc9Ac) and their precursors (Neu5Ac and Neu5Gc) are most common sialic acids in nature (X. Song et al., 2011). Neu5Gc9Ac is different from Neu5,9Ac$_2$ only by an additional oxygen atom at the C5 position (X. Song et al., 2011). The differential usage of these two nearly identical 9-O-acetylated SAs between IDV and ICV implies that two seven-segmented influenza viruses diverge in communicating with both O-acetyl group at the C9 position and acetyl/glycolyl groups at the C5 position in terminal 9-carbon sialic acids. We interpret our experimental data presented here that IDV in general may have high binding affinity and broad specificity, while ICV may have lower binding affinity and narrow selectivity for 9-O-acetylated SAs. Such qualitative and quantitative differences in the virus-glycan receptor interaction that may ultimately discriminate IDV from its related ICV in infectious landscape and ecology, which clearly warrant further investigation.

3.5 Acknowledgments

BIBLIOGRAPHY

Bardor, M., Nguyen, D. H., Diaz, S., and Varki, A. (2005). Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. The Journal of biological chemistry *280*, 4228-4237.

Byres, E., Paton, A. W., Paton, J. C., Lofling, J. C., Smith, D. F., Wilce, M. C., Talbot, U. M., Chong, D. C., Yu, H., Huang, S., Chen, X., Varki, N. M., Varki, A., Rossjohn, J., and Beddoe, T. (2008). Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. Nature *456*, 648-652.

Collin, E. A., Sheng, Z., Lang, Y., Ma, W., Hause, B. M., and Li, F. (2015). Cocirculation of two distinct genetic and antigenic lineages of proposed influenza D virus in cattle. Journal of virology *89*, 1036-1042.

Ducatez, M. F., Pelletier, C., and Meyer, G. (2015). Influenza D virus in cattle, France, 2011-2014. Emerging infectious diseases *21*, 368-371.

Ferguson, L., Eckard, L., Epperson, W. B., Long, L. P., Smith, D., Huston, C., Genova, S., Webby, R., and Wan, X. F. (2015). Influenza D virus infection in Mississippi beef cattle. Virology *486*, 28-34.

Foni, E., Chiapponi, C., Baioni, L., Zanni, I., Merenda, M., Rosignoli, C., Kyriakis, C. S., Luini, M. V., Mandola, M. L., Bolzoni, L., Nigrelli, A. D., and Faccini, S. (2017). Influenza D in Italy: towards a better understanding of an emerging viral infection in swine. Sci Rep *7*, 11660.

Hause, B. M., Collin, E. A., Liu, R., Huang, B., Sheng, Z., Lu, W., Wang, D., Nelson, E. A., and Li, F. (2014). Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. mBio *5*, e00031-00014.

Hause, B. M., Ducatez, M., Collin, E. A., Ran, Z., Liu, R., Sheng, Z., Armien, A., Kaplan, B., Chakravarty, S., Hoppe, A. D., Webby, R. J., Simonson, R. R., and Li, F. (2013). Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens *9*, e1003176.

Inoue, S., Sato, C., and Kitajima, K. (2010). Extensive enrichment of N-glycolylneuraminic acid in extracellular sialoglycoproteins abundantly synthesized and secreted by human cancer cells. Glycobiology *20*, 752-762.

Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., and Suzuki, A. (1998). The molecular basis for the absence of N-glycolylneuraminic acid in humans. The Journal of biological chemistry *273*, 15866-15871.

Jiang, W. M., Wang, S. C., Peng, C., Yu, J. M., Zhuang, Q. Y., Hou, G. Y., Liu, S., Li, J. P., and Chen, J. M. (2014). Identification of a potential novel type of influenza virus in Bovine in China. Virus genes *49*, 493-496.

Langereis, M. A., Bakkers, M. J., Deng, L., Padler-Karavani, V., Vervoort, S. J., Hulswit, R. J., van Vliet, A. L., Gerwig, G. J., de Poot, S. A., Boot, W., van Ederen, A. M., Heesters, B. A., van der Loos, C. M., van Kuppeveld, F. J., Yu, H., Huizinga, E. G., Chen, X., Varki, A., Kamerling, J. P., and de Groot, R. J. (2015). Complexity and Diversity of the Mammalian Sialome Revealed by Nidovirus Virolectins. Cell reports *11*, 1966-1978.

Matsuzaki, Y., Sugawara, K., Furuse, Y., Shimotai, Y., Hongo, S., Oshitani, H., Mizuta, K., and Nishimura, H. (2016). Genetic Lineage and Reassortment of Influenza C Viruses Circulating between 1947 and 2014. Journal of virology *90*, 8251-8265.

Ng, P. S., Bohm, R., Hartley-Tassell, L. E., Steen, J. A., Wang, H., Lukowski, S. W., Hawthorne, P. L., Trezise, A. E., Coloe, P. J., Grimmond, S. M., Haselhorst, T., von Itzstein, M., Paton, A. W., Paton, J. C., and Jennings, M. P. (2014). Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. Nature communications *5*, 5750.

Quast, M., Sreenivasan, C., Sexton, G., Nedland, H., Singrey, A., Fawcett, L., Miller, G., Lauer, D., Voss, S., Pollock, S., Cunha, C. W., Christopher-Hennings, J., Nelson, E., and Li, F. (2015). Serological evidence for the presence of influenza D virus in small ruminants. Veterinary microbiology *180*, 281-285.

Rogers, G. N., Herrler, G., Paulson, J. C., and Klenk, H. D. (1986). Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. The Journal of biological chemistry *261*, 5947-5951.

Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C. H., Meier-Ewert, H., Skehel, J. J., and Wiley, D. C. (1998). Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature *396*, 92-96.

Salem, E., Cook, E. A. J., Lbacha, H. A., Oliva, J., Awoume, F., Aplogan, G. L., Hymann, E. C., Muloi, D., Deem, S. L., Alali, S., Zouagui, Z., Fevre, E. M., Meyer, G., and Ducatez, M. F. (2017). Serologic Evidence for Influenza C and D Virus among Ruminants and Camelids, Africa, 1991-2015. Emerging infectious diseases *23*, 1556-1559.

Samraj, A. N., Pearce, O. M., Laubli, H., Crittenden, A. N., Bergfeld, A. K., Banda, K., Gregg, C. J., Bingman, A. E., Secrest, P., Diaz, S. L., Varki, N. M., and Varki, A. (2015).

A red meat-derived glycan promotes inflammation and cancer progression. Proceedings of the National Academy of Sciences of the United States of America *112*, 542-547.

Song, H., Qi, J., Khedri, Z., Diaz, S., Yu, H., Chen, X., Varki, A., Shi, Y., and Gao, G. F. (2016). An Open Receptor-Binding Cavity of Hemagglutinin-Esterase-Fusion Glycoprotein from Newly-Identified Influenza D Virus: Basis for Its Broad Cell Tropism. PLoS pathogens *12*, e1005411.

Song, X., Yu, H., Chen, X., Lasanajak, Y., Tappert, M. M., Air, G. M., Tiwari, V. K., Cao, H., Chokhawala, H. A., Zheng, H., Cummings, R. D., and Smith, D. F. (2011). A sialylated glycan microarray reveals novel interactions of modified sialic acids with proteins and viruses. The Journal of biological chemistry *286*, 31610-31622.

Sreenivasan, C., Thomas, M., Sheng, Z., Hause, B. M., Collin, E. A., Knudsen, D. E., Pillatzki, A., Nelson, E., Wang, D., Kaushik, R. S., and Li, F. (2015). Replication and Transmission of the Novel Bovine Influenza D Virus in a Guinea Pig Model. Journal of virology *89*, 11990-12001.

White, S. K., Ma, W., McDaniel, C. J., Gray, G. C., and Lednicky, J. A. (2016). Serologic evidence of exposure to influenza D virus among persons with occupational contact with cattle. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology *81*, 31-33.

Zhai, S. L., Zhang, H., Chen, S. N., Zhou, X., Lin, T., Liu, R., Lv, D. H., Wen, X. H., Wei, W. K., Wang, D., and Li, F. (2017). Influenza D Virus in Animal Species in Guangdong Province, Southern China. Emerging infectious diseases *23*, 1392-1396.

Chapter 4. Genesis, antigenic evolution, and temperature-dependent replication of a recent human influenza C virus clinical isolate

Abstract: Unlike influenza A and B viruses that infect humans and cause severe diseases in seasonal epidemics, influenza C virus (ICV) is a ubiquitous childhood pathogen typically causing mild respiratory symptoms. ICV infections are rarely diagnosed and less research has been performed on it despite the virus being capable of causing severe disease in infants. Here we report on the isolation of a human ICV from a child with acute respiratory disease, provisionally designated C/Victoria/2/2012 (C/Vic). The full-length genome sequence and phylogenetic analysis revealed that the hemagglutinin-esterase-fusion (HEF) gene of C/Vic was derived from C/Sao Paulo lineage, PB2, PB1, M and NS of C/Vic were classified into C/Yamagata-related lineage, while P3 and NP were divided into C/Mississippi-related lineage. Furthermore, antigenic analysis using the HI assay found that 1947 C/Taylor virus (C/Taylor lineage) was antigenically more divergent from 1966 C/Johannesburg (C/Aichi lineage) than from 2012 C/Vic. Structure modeling of the HEF protein identified two mutations in the 170-loop of the HEF protein around the receptor binding pocket as a possible antigenic determinant responsible for the discrepant HI results. Finally, C/Vic was found to replicate more efficiently at the cool temperature found in the nasal cavity (33 °C) than at the core body temperature (37 °C) in a panel of epithelial cell lines from human, swine and canine. Taken together, results of our studies reveal novel insights into the genetic and antigenic evolution of ICV and provide a framework for further investigation of the molecular determinants of temperature-dependent growth and antigenic property.

4.1 Introduction

The *Orthomyxoviridae* family has four genera of influenza viruses, influenza A, influenza B, influenza C, and influenza D. Among them, influenza A, B, and C are known to cause moderate to severe respiratory diseases in humans. Specifically, human infections by influenza A (IAV) and B (IBV) viruses can lead to severe respiratory diseases, while influenza C virus (ICV) usually causes mild upper respiratory diseases in humans, although it has the ability in causing severe lower respiratory illness in children less than 2 years of age (S. Katagiri et al., 1983; Y. Matsuzaki et al., 2006; H. Moriuchi et al., 1991). ICV is distributed worldwide (A. C. Dykes et al., 1980; Y. Matsuzaki et al., 2016; H. Nishimura et al., 1987) and multiple genetic lineages co-circulate globally (D. A. Buonagurio et al., 1986; Y. Furuse et al., 2016; Y. Matsuzaki et al., 2000; Y. Matsuzaki et al., 1994). In addition to three genera of influenza viruses that all infect humans, a new group of influenza viruses with cattle as a primary reservoir has been recently described and these new viruses are classified into influenza D genus due to its distinctness from other influenza genera (E. A. Collin et al., 2015; B. M. Hause et al., 2014; B. M. Hause et al., 2013; Z. Sheng et al., 2014). Influenza D viruses (IDV) are thought to primarily infect and cause respiratory diseases in cattle and to some extent in pigs. Nevertheless, IDV-specific antibodies had been found in humans, especially those who had a previous history of contact with cattle.

Among human influenza viruses, ICV infections are rarely diagnosed and less research has been performed on it despite the virus having the ability to cause severe diseases in newborn infants. Previous analyses of the hemagglutinin-esterase-fusion (HEF) gene, encoding the antigenic determinants of viruses, have divided ICVs into six genetic and antigenic lineages, designated C/Taylor, C/Mississippi, C/Aichi, C/Yamagata, C/Kanagawa, and C/Sao Paulo (S. Hachinohe et al., 1989; Y. Matsuzaki et al., 2003; Y. Matsuzaki et al., 1994; Y. Muraki et al., 1996; K. Sugawara et al., 1988; K. Sugawara et al., 1993). ICV is thought to evolve relatively slow compared to IAV or IBV to a lesser extent (P. Chakraverty, 1974; Y. Furuse et al., 2016; H. Kawamura et al., 1986; Y. Muraki et al., 1996). However, frequent reassortments among ICVs have always occurred in nature and most of the circulating ICV are generated due to various reassortantments involving multiple ICV lineages (Y. Matsuzaki et al., 2003). Co-circulation of several ICV lineages in humans has been well documented (Y. Matsuzaki et al., 1994). Genetic diversity and the associated antigenic drift caused by reassortment play an important role in driving ICV's periodical recurrence in humans. To date, the majority of published research on human ICV utilized historic reference virus strains isolated between the 1950s and 1960s. As such, more studies of currently circulating strains of ICV are critically needed in order to better understand the epidemiology, genetic and antigenic evolution, and biology of this group of human influenza viruses, which can pose a significant risk to worldwide infant population.

Here we described the isolation of a contemporary influenza C virus - C/Victoria/2/2012 (C/Vic) - from a diseased child with acute respiratory symptoms in 2012. Phylogenetic

analysis indicated that viral HEF gene was derived from C/Sao Paulo lineage, which is consistent with the finding that the dominant antigenic group is C/Sao Paulo lineage from 2006 to 2016 (Y. Matsuzaki et al., 2014; T. Odagiri et al., 2015; S. Tanaka et al., 2015; T. Yano et al., 2014). PB2, PB1, M and NS of C/Vic were classified into C/Yamagata-related lineage, while P3 and NP were divided into C/Mississippi-related lineage. Furthermore, antigenic analysis using the HI assay found 1947 C/Taylor virus (C/Taylor lineage) was antigenically more divergent from1966 C/Johannesburg (C/Aichi lineage) than from C/Vic. Structure modeling of the HEF protein identified two mutations in the 170-loop of the HEF protein around the receptor binding pocket as a possible antigenic determinant responsible for the discrepant HI results. Growth kinetics studies conducted at both cool (33 °C) and core body (37 °C) temperature demonstrated a temperature-dependent replication property for this contemporary ICV isolate. Information obtained through this study shall provide novel insights into genesis, antigenic evolution and replication biology of human influenza C virus.

4.2 Materials and Methods

4.2.1 Cell and virus cultures

Madin-Darby canine kidney (MDCK), human lung adenocarcinoma A549 (A549), human lung adenocarcinoma Calu-3 (Calu-3), Swine Testicle (ST), Swine tracheal and bronchial epithelial cell line-MK1-OSU (MK1-OSU) and SD-PJEC cells, a subclone of the IPEC-J2 cell line, originally derived from newborn piglet jejunum, were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) at 37 °C with 5% CO_2. The infectious culture medium is serum-free DMEM containing a

final concentration of 0.0005% trypsin (Sigma-Aldrich, St Louis, USA). Virus cultures were incubated at 33 °C for up to 7 days or development of cytopathic effect (CPE).

4.2.2 Genome sequencing and phylogenetic analysis

Viral RNA was isolated using a MagMAX™ Viral RNA Isolation kit. Full genome amplification was performed as previously described except that the primers were modified to match the non-coding regions of ICV: "ICV3'", 5'-ACGCGTGATCGCATAAGCAG-3' and "ICV5'", 5'-ACGCGTGATCAGCAGTAGCAAG-3'. Amplicons were used for library preparation using the NEBNext® Fast DNA Library Prep Set for Ion Torrent™. Libraries were sequenced using an Ion Torrent Personal Genome Machine with manufacturer's reagents and protocols. Contigs were assembled with C/Ann Arbor/1/50 (accession numbers NC_006306-NC_006312) as templates, which were analyzed by the SeqMan NGen module from DNAStar. Nucleotide sequences were edited and compiled using the Lasergene 8 software package (DNASTAR, Inc.). The genome sequence of C/Vic was submitted to GenBank under accession numbers KM504277-KM504283.

The genome sequences of C/Vic were determined in our laboratory. The sequences of other ICV and IDV strains were obtained from GenBank. To fully understand the evolutionary history of ICVs, the six strains (C/Taylor, C/Mississippi, C/Aichi, C/Yamagata, C/Kanagawa, and C/Sao Paulo) represented the six genetic and antigenic lineages were included in the phylogenetic analysis of each RNA segment. Reference viruses included historic ICVs such as C/Ann Arbor/1/50 and C/Taylor/1233/1947, as well as contemporary viruses. Identical sequences were removed and only the ones with

the earliest isolation time were used. Incomplete sequences were removed. The complete coding regions of each segment was used to construct phylogenetic trees.

Sequences of each segment were aligned at the amino acid level using Muscle (ref), and the aligned sequences were translated back to generate alignments of nucleotide sequences, which were used for phylogenetic tree construction using MEGA6 (ref). The best substitution model for each segment was estimated using MEGA6. The best models predicted were T92+G for NS and GTR+G+I for the other six segments. The phylogenetic trees were generated using Maximum-likelihood method in MEGA6 with 1000 bootstrap replicates to estimate the confidence of the tree topologies. The trees were rooted using the ancestor branches of IDV strains.

4.2.3 Hemagglutination inhibition assay

Antibody cross-reactivity was determined using a panel of reference viruses and antisera in hemagglutination inhibition (HI) assays. Chicken antiserum against C/ Taylor/1233/47 (C/Taylor lineage) and C/Taylor/1233/47 virus were provided by NIAID biodefense and Emerging Infections Research Resources Repository (BEI Resources). Antiserum against C/Victoria/2/2012 (C/Sao Paulo lineage) was generated in rabbits, while C/Johannesburg/1/66 virus (C/Aichi lineage) was provided by Peter Palese at Mt. Sinai Medical School, New York. All sera were heat inactivated at 56°C for 30 min prior to use. HI assays were performed following standard procedures. In brief, sera were treated with a receptor-destroying enzyme for 24 h at 37°C and then adsorbed with a 20% suspension of turkey erythrocytes in phosphate-buffered saline (PBS) for 30 min at room temperature.

Virus suspensions containing 4 to 8 HA units of the virus were incubated for 1 h with serial 2-fold dilutions of antiserum, and the HI titer was determined as the reciprocal of the highest dilution that showed complete inhibition of hemagglutination using 0.5% washed turkey erythrocytes (Lampire biological laboratories, Pipersville, PA). All viruses were assayed in triplicate. Mean HI titers and standard deviations were calculated from triplicate data. Heterologous mean HI titers were normalized to mean homologous HI titers (mean heterologous titers/mean homologous titers). Relative HI titers were reported to account for differences in homologous HI titers between C/Taylor and C/Victoria/2/2012.

4.2.4 Temperature-dependent virus replication kinetics

MDCK, ST, and MK1-OSU cells were infected with C/Vic at a multiplicity of infection (MOI) of 0.01 while A549, Calu-3, and SD-PJEC cells were infected with an MOI of 1.0. The virus was allowed to adsorb to cells incubated at 33 °C or 37 °C for 1 hour (h). The inoculum was then removed and the cells were washed twice with PBS. Cells were maintained with DMEM containing tosyl phenylalanyl chloromethyl ketone (TPCK) treated trypsin (1 µg/ml TPCK-Trypsin for MK1-OSU and MDCK, 0.1 µg/ml TPCK-Trypsin for SD-PJEC, A549, Calu-3 and ST). Cells culture supernatants were collected every 24 hours. Virus titration was performed on MDCK cells. Results were obtained from two independent experiments.

4.2.5 Structure modeling and sequence alignment

The viral HEF sequences (C/Vic and C/Taylor) were modeled with Modeller (N. Eswar et al., 2007). The structure of C/Johannesburg/1/66 (C/JHB/1/66) was used as template (sequence identity: 95-96%). Sequences were aligned with Muscle (R. C. Edgar, 2004). Domain boundaries and the trimeric interface were determined according to those of C/JHB/1/66 (P. B. Rosenthal et al., 1998). Variant residues of the HEF protein among three viruses were colored pink.

4.3 Results and Discussion

4.3.1 Virus isolation and full-length genome analysis

In 2012, a nasopharyngeal swab from child patient exhibiting clinical symptoms of acute respiratory infection was submitted for virus isolation to Victorian Infectious Diseases Reference Laboratory, Melbourne, Australia. After cultivation in MDCK cells at 33 °C together with RT-PCR diagnosis and sequence confirmation, human influenza C virus was isolated from the patient-derived nasopharyngeal swab sample, provisionally designated C/Victoria/2/2012. The virus cultured in MDCK cells was then sequenced on an Ion Torrent Personal Genome Machine and De novo genome assembly was employed to compile viral full-length genome. The full-length sequences of all segments were determined and used for phylogenetic analysis.

4.3.2 Phylogenetic analysis

To determine the evolutionary pathway of C/Vic, we performed phylogenetic analysis on all seven segments of C/Vic and reference ICVs. Nucleotide sequences of the complete coding region for the seven RNA segments were used for phylogenetic tree construction

(Fig. 4.1). Reference viruses including historic ICVs such as C/Taylor/1947 and C/Ann Arbor/50, as well as the recent isolates including C/Yamagata/16/2014 and C/Eastern-India/1202/2011 were used in our analysis (T. Roy Mukherjee et al., 2013). We included all influenza D virus (IDV) strains isolated from 2011 to 2016 to better represent the evolutionary history of both influenza genotypes. The trees were rooted using the ancestor branches of IDV strains.

Previous studies of the HEF segment, the primary determinant of host range and target of virus-neutralizing antibodies, have classified ICVs into six genetic and antigenic lineages, designated C/Taylor, C/Mississippi, C/Aichi, C/Yamagata, C/Kanagawa, and C/Sao Paulo. As shown in Fig. 4.1, the HEF-based tree recaptured the currently defined lineage classification for ICVs, thereby validating our analytical approach. Phylogenetic analysis of the HEF gene placed C/Vic in C/Sao Paulo lineage, which also includes most recent ICV isolates from 2006 to 2015. C/Vic is closely clustered together with C/Yamagata/33/2014 from Japan and two strains C/Biliran/1/2013 and C/Leyte/2/2011 from Philippines. Interestingly several recent ICV isolates, including C/Yamagata/7/2012 and C/Miyagi/2/2014, belonged to C/Kanagawa lineage. This result supports the notion that multiple ICV lineages have co-circulated in global human populations in recent years. As shown in Fig. 4.1, in contrast to a HEF-based tree, the topologies of internal branches of the phylogenetic trees of other segments have low statistical supports, which may be due to their high sequence similarities. This prevented us from drawing any precise lineage classification of ICVs. Yoko Matsuzaki *et al* divided the genetic lineages of the internal genes into two major lineages, the C/Mississippi-related lineage and the

C/Yamagata-related lineage (Y. Matsuzaki et al., 2016). From this aspect, PB2, PB1, M and NS of C/Vic were classified into C/Yamagata-related lineage, while P3 and NP were divided into C/Mississippi-related lineage. A close examination of the PB2 tree revealed that C/Vic formed a subgroup with C/pig/Beijing/115/81, a porcine ICV isolated from pigs with influenza-like symptoms in China in 1981. A similar finding was also observed in the M gene-derived tree in that C/Vic was closely related to C/pig/Beijing/115/81. These results indicated the potential that recent ICV strains including C/Vic may replicate and transmit among pigs. Analysis of PB1 gene-derive tree demonstrated that C/Vic were closely related to C/Miyagi/5/2014 and C/Tokyo/4/2014. Similar to HEF segment, C/Vic NS segment was closely clustered together with C/Tokyo/4/2014, C/Biliran/1/2013 and C/Leyte/2/2011. These results suggested that these viruses were circulating in Australia, Philippines and Japan from 2011 to 2014. Phylogenetic analysis of the P3 and NP genes form a subgroup with historical ICV isolates including C/Mississippi/80, C/Nara/82 and C/Greece/1/79, were classified into C/Mississippi-related lineage. Taken together, the results of our phylogenetic studies demonstrated that C/Vic is a reassortant virus composed of segments derived from multiple ICV lineages or strains, which evolved independently.

The trees were rooted using the ancestor branches of influenza D virus strains. Our previous study showed IDVs isolated in United States were classified into two distinct cocirculating lineages represented by D/swine/Oklahoma/1334/2011 (D/OK) and D/bovine/Oklahoma/660/2013 (D/660) (E. A. Collin et al., 2015). However, the recently reported D/bovine/Ibaraki/7768/2016 was not clustered into either lineage except M

segment (S. Murakami et al., 2016). We included all IDV isolates in our phylogenetic trees to better understand the evolution between ICVs and IDVs. The phylogenetic analysis showed two distinct lineages of IDVs in PB2 and PB1, M and NS segments. In P3, HEF and NP gene-derived trees, D/France/2986/2012 and D/Ibaraki/7768/2016 were not fell into either lineage. Interestingly, several American isolates such as D/bovine/Minnesota/628/2013, D/bovine/Kansas/13-21/2012 and D/bovine/Texas/3-13/2011 in P3-derived tree and three isolates from Shandong, China in NP-derived tree were not clustered into any lineage either. The results demonstrated that more than two lineages of IDVs are co-circulating all over the world and frequently reassorted with one another.

HEF

NP

C/Yamagata/9/2006
C/Tokyo/3/2010
C/Fukuoka/1/2012
C/Yamagata/2/2010
C/Sendai/TU1/2012
C/Miyagi/6/2014
C/Aichi/1/99
C/Leyte/2/2011
C/Biliran/1/2013
C/Yamagata/33/2014
◆ C/Victoria/2/2012
C/Sao Paulo/378/82

C/Sao Paulo

C/Johannesburg/1/66
C/Aichi/1/81
C/Miyagi/5/91
C/Yamagata/4/92

C/Aichi

C/Ann Arbor/1/50
C/Paris/1/67
C/Taylor/1233/1947

C/Taylor

C/Yamagata/7/88
C/Yamagata/26/81
C/Yamagata/8/96
C/Yamagata/15/2004
C/England/892/1983
C/pig/Beijing/115/81

C/Yamagata

C/Miyagi/77
C/Kanagawa/1/76
C/Fukuoka/2/2006
C/Miyagi/11/2002
C/Yamagata/14/2004
C/Yamagata/2/2005
C/Yamagata/1/2006
C/Fukushima/1/2004
C/Miyagi/28/2002
C/Singapore/DSO-070193/2006
C/Yamagata/7/2012
C/Miyagi/2/2014

C/Kanagawa

C/Yamagata/26/2004
C/Miyagi/8/96
C/Yamagata/3/2000
C/Miyagi/3/93
C/Mississippi/80

C/Mississippi

D/bovine/Mexico/S8/2015
D/bovine/Mexico/S7/2015
D/bovine/Mexico/S56/2015
D/bovine/Kansas/13-21/2012
D/bovine/Nebraska/9-5/2012
D/bovine/Texas/3-13/2011
D/bovine/Kansas/11-8/2012
D/bovine/Oklahoma/660/2013
D/bovine/Mississippi/C00014N/2014
D/bovine/Mississippi/C00013N/2014

D/660

D/bovine/France/2986/2012
D/bovine/Ibaraki/7768/2016
D/bovine/Italy/46484/2015
D/swine/Italy/199724-3/2015
D/bovine/Italy/1/2014
D/bovine/Kansas/1-35/2010
D/bovine/Mexico/S62/2015
D/swine/Oklahoma/1334/2011
D/bovine/Mississippi/C00030P/2014
D/bovine/Mississippi/C00046N/2014
D/bovine/Kansas/14-22/2012
D/bovine/Minnesota/729/2013
D/bovine/Minnesota/628/2013
D/bovine/Shandong/Y125/2014
D/bovine/Shandong/Y127/2014
D/bovine/Shandong/Y217/2014

D/OK

2

C/Miyagi/5/2014
◆ C/Victoria/2/2012
C/India/P135047/2013
C/Yamagata/16/2014
C/Yamagata/2/2005
C/Miyagi/9/96
C/Fukuoka/2/2006
C/Miyagi/12/2004
C/India/P119564/2011
C/Yamagata/6/2000
C/Yamagata/8/96
C/Nara/82
C/Greece/79
C/Mississippi/80
C/Miyagi/5/91
C/Yamagata/4/92
C/Aichi/1/81
C/Yamagata/26/81
C/Yamagata/7/81
C/Shizuoka/79
C/Sao Paulo/378/82
C/Kanagawa/1/76
C/Miyagi/77
C/Ann Arbor/1/50
C/Taylor/1233/1947
C/Johannesburg/1/66
C/Paris/1/67
C/Kyoto/1/79
C/Miyagi/1/93
C/pig/Beijing/115/81
C/England/892/1983
C/Miyagi/7/93
C/Miyagi/4/2006
C/Yamagata/16/2008
C/Tokyo/3/2010
C/Yamagata/14/2012
C/Yamagata/29/2012
C/Miyagi/2/2008
C/Fukuoka/1/2012
C/Yamagata/2/2010
D/bovine/Ibaraki/7768/2016
D/bovine/Shandong/Y125/2014
D/bovine/Shandong/Y127/2014
D/bovine/Shandong/Y217/2014
D/bovine/France/2986/2012
D/bovine/Italy/46484/2015
D/swine/Italy/199724-3/2015
D/bovine/Italy/1/2014
D/swine/Oklahoma/1334/2011
D/bovine/Mississippi/C00046N/2014
D/bovine/Kansas/14-22/2012
D/bovine/Oklahoma/660/2013
D/bovine/Kansas/11-8/2012
D/bovine/Kansas/1-35/2010
D/bovine/Texas/3-13/2011
D/bovine/Nebraska/9-5/2012
D/bovine/Minnesota/628/2013
D/bovine/Minnesota/729/2013
D/bovine/Kansas/13-21/2012
D/bovine/Mexico/S62/2015
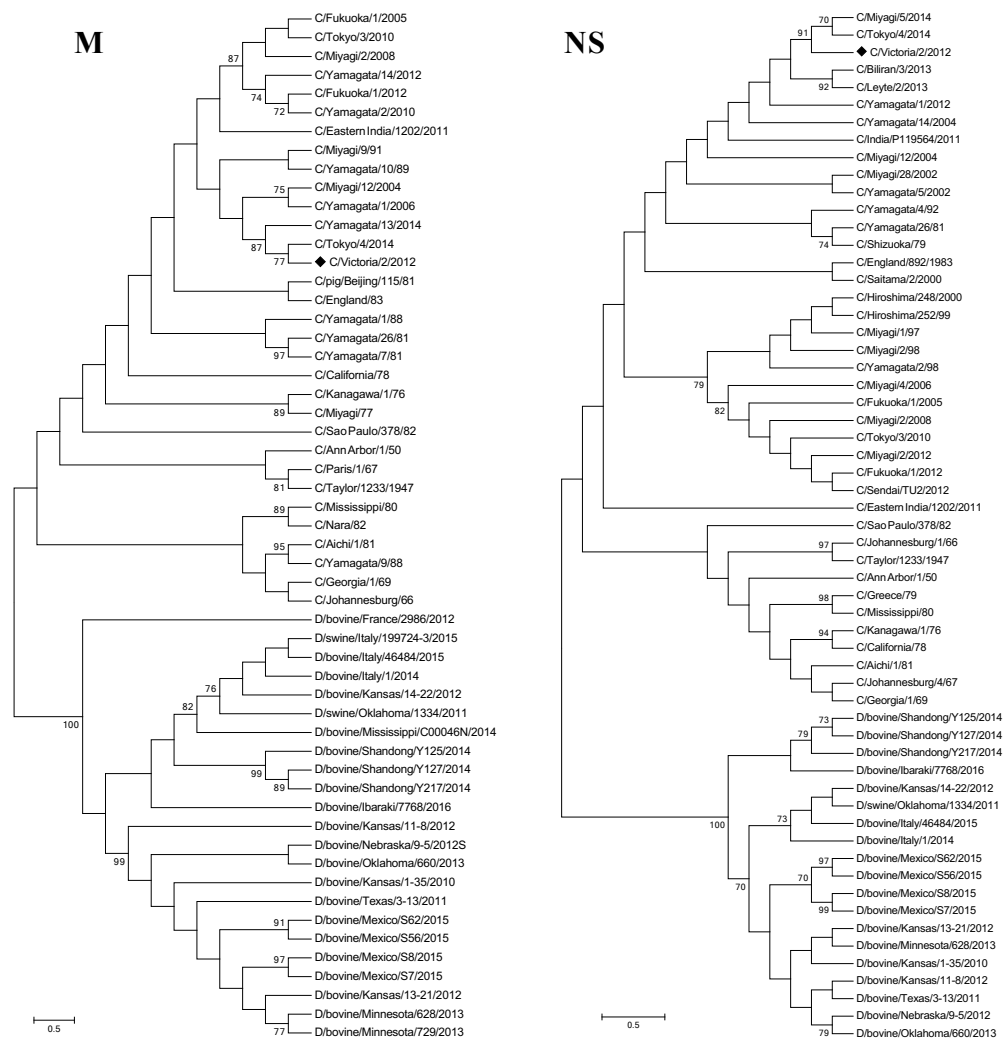D/bovine/Mexico/S56/2015
D/bovine/Mexico/S8/2015
D/bovine/Mexico/S7/2015

5

Figure 4.1. Phylogenetic trees of the seven genomic segments of influenza C virus.
Sequences of each segment of human influenza C viruses were aligned and then used for phylogenetic tree construction through the MEGA7 program. The best substitution model for each segment was estimated using MEGA7. The phylogenetic trees were generated using Maximum-likelihood method of MEGA7 with 1000 bootstrap replicates to verify the topology. The trees were rooted using the ancestor branches of influenza D virus strains. C/Victoria/2/2012 strain is marked as diamond (◆).

4.3.3 Antigenic evolution

To investigate the antigenic property of C/Vic, HI assays were performed using polyclonal antiserum generated against 1947 C/Taylor (C/Taylor lineage) and 2012 C/Vic (C/Sao Paulo lineage). Historical 1966 C/Johannesburg (C/JHB) (C/Aichi lineage) virus was also included together with 1947 C/Taylor and 2012 C/Vic in this antigenic study. Homologous HI titers for C/Taylor and C/Vic were 5120 and 1280, respectively. Heterologous mean HI titers against C/Taylor and C/Vic antisera were normalized to homologous C/Taylor and C/Vic titers. Interestingly, the three lineage viruses spanning over 60 years showed equivalent cross-reactivity with C/Vic antisera with relative HI titers 1.0 (Fig. 4.2). In contrast, cross-reactivity profile with C/Taylor antisera discriminated three viruses clearly. Specifically, C/Taylor virus reacted most strongly with homologous C/Taylor antisera (relative HI titer 1.0) followed by C/Vic (relative HI titer 0.5) and by C/JHB (relative HI titer 0.25) (Fig. 4.2). These data suggested that 1947 C/Taylor was more antigenically related to 2012 C/Vic than to 1966 C/JHB. The observed antigenic variations were in good agreement with HEF phylogeny of these ICV strains (Fig. 4.1).
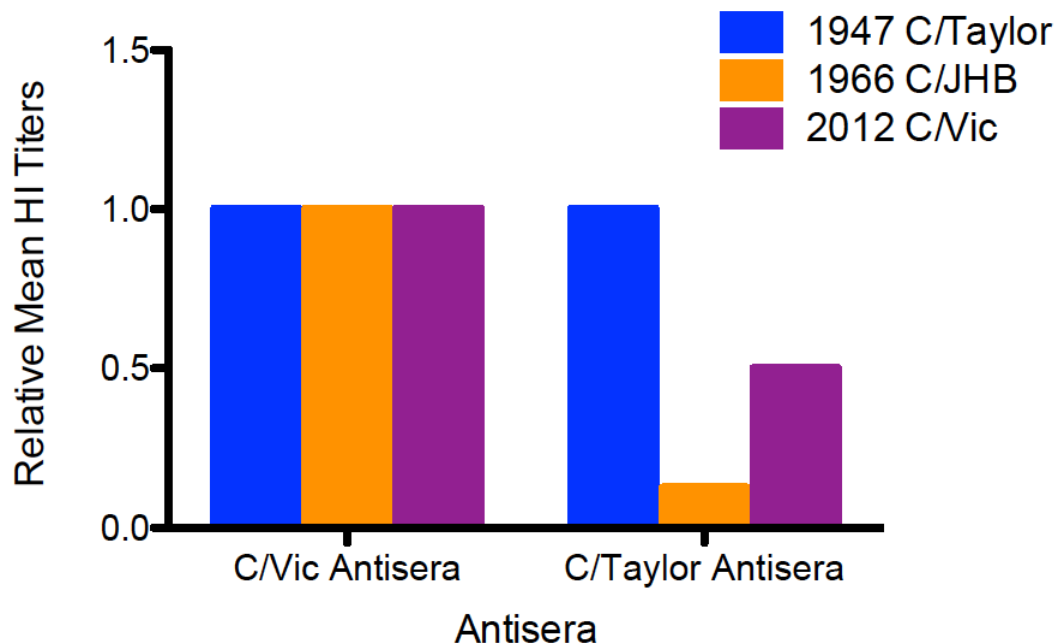
Figure 4.2. Relative mean hemagglutination inhibition titers for 3 influenza C viruses from triplicate data. Relative mean HI titers were calculated by normalizing mean heterologous HI titers to mean homologous HI titers (mean heterologous HI titer/mean homologous HI titer).

4.3.4 Structure-basis of antigenic variation

To identify critical residues of the HEF protein among these three ICVs contributing most to the observed antigenic drift (i.e., HA titer change), pairwise comparisons of viral HEF structures were pursued. The resolved crystal structure of C/JHB HEF (PDB ID: 1FLC) was chosen as a model template to model the HEF structures of the C/Vic and C/Taylor strains. We focused on four secondary elements (the 170-loop, 230-helix, 270-loop, and 290-loop) constituting the HEF receptor-binding pocket (RBP) for our investigation (Fig. 4.3A). Structure modeling showed that three ICVs exhibited two

amino acid substitutions in the 170-loop proximal to the RBP of the HEF, which likely have a decisive role in the degree of antigenic distance observed among three ICV strains (Fig. 4.3). Specifically, the more antigenically similar virus pair C/Vic-C/JHB or C/Vic-C/Taylor only had one mutation in the 170-loop. K172G or T170V substitution occurred between C/Vic and C/JHB (Fig. 4.3B) or C/Vic and C/Taylor (Fig. 4.3C) (listed in an arrangement as the C/Vic amino acid residue, HEF position, and C/JHB or C/Taylor amino acid), respectively. In contrast, the more antigenically divergent virus pair C/Taylor-C/JHB (Fig. 4.3D) acquired these two mutations V170T and K172G of the 170-loop. These structure-based analyses seemed to indicate a link between the level of antigenic variation among three strains of ICV and the number of the amino acid changes in the proximity of the HEF receptor-binding pocket. In addition, multiple amino acid substitutions were also identified among three viruses in various regions of the HEF including HEF1, HEF2, and the trimeric interface. Considering that all those mutations were distant from the RBP of the HEF protein, we speculate that these amino changes may not modulate directly antigenicity of influenza C viruses.
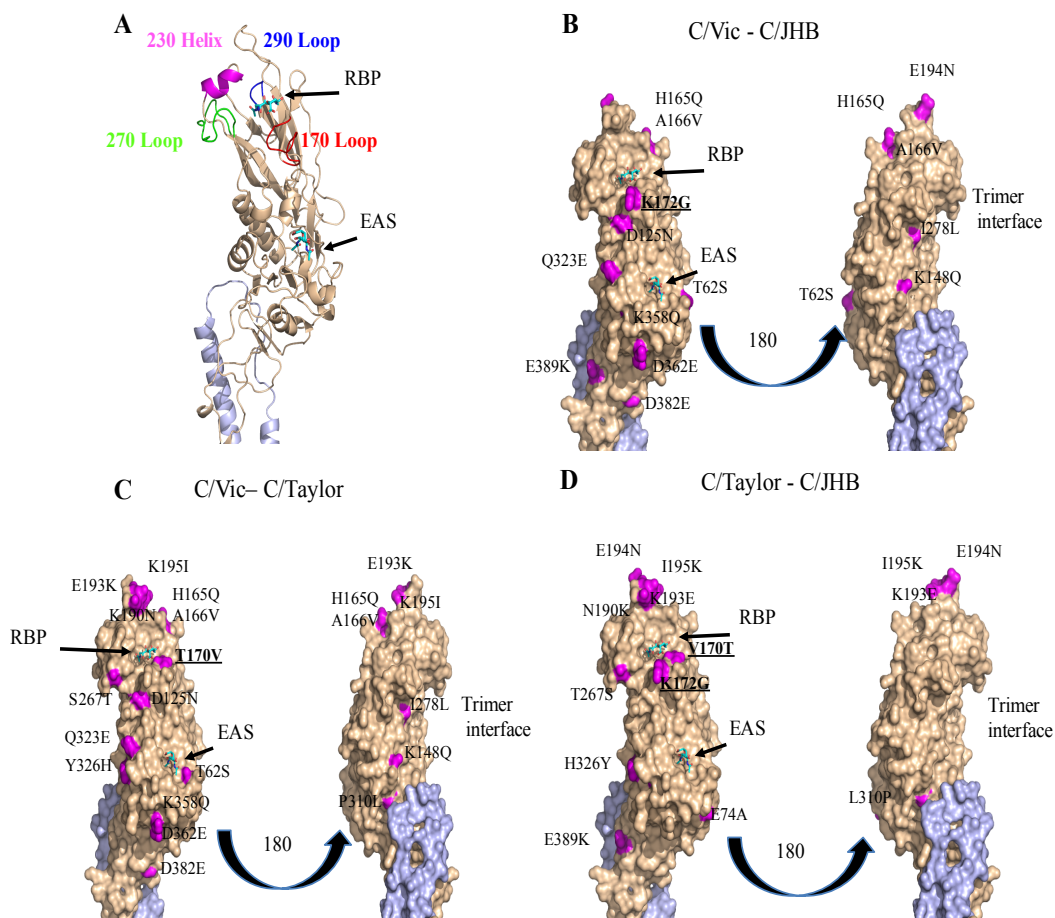
Figure 4.3. Structure-basis of antigenic variation. (A) Cartoon view of modeled HEF1 domain of C/Vic (colored Wheat), while HEF2 was colored light blue. Salic acids in receptor binding pocket (RBP) and esterase active site (EAS) were colored cyan. Four secondary elements (the 170-loop, 230-helix, 270-loop, and 290-loop) constituting the HEF receptor-binding pocket were indicated with different colors. (B) (C) (D) HEF1 residues different among three virus pairs were mapped onto the modeled HEF1 domain of C/Vic. Note the critical amino acid changes in the 170-loop of the HEF protein around the receptor-binding pocket were highlighted by bold and underlined words.

4.3.5 Temperature-dependent replication

To investigate whether the replication robustness of this recent ICV C/Vic isolate is dependent on temperature and its *in vitro* cellular tropism, we determined the growth kinetics of C/Vic in a panel of human, swine, and canine cell lines at the cool temperature (33 °C) as well as at the core body temperature (37 °C). Among cell lines, MDCK cell line is commonly used for animal and human influenza virus replication, while two human lung epithelial cell lines A549 and Calu-3 are often employed to study the replication of human influenza viruses. Including two swine cell lines, ST and MK1-OSU, allowed us to determine whether human C/Vic was capable of replicating in swine cells, especially in swine airway epithelial cells (MK1-OSU). An influenza C virus has been previously isolated from diseased pigs and our phylogenetic analysis showed that C/Vic is more closely related to this porcine ICV isolate C/pig/Beijing/115/81 in the two internal genes (PB2 and M) than its human counterparts (Fig. 4.1).

As demonstrated in Fig. 4.4, C/Vic replicated more robustly at the cool temperature (33 °C) than at the core body temperature (37 °C) in tested cell lines. This result was in good agreement with those previously reported for other ICVs. Interestingly, the two human airway cells lines A549 and Calu-3 showed dramatic differences in their ability to support C/Vic replication. C/Vic reached its peak titer of 4.2 $\log_{10}$TCID$_{50}$/ml at 33 °C and 3.7 $\log_{10}$TCID$_{50}$/ml at 37 °C in Calu-3, respectively, while the virus only reached to 1.0 $\log_{10}$TCID$_{50}$/ml in A549 at both 33 °C and 37 °C. The observation that these two very similar human lung-derived cell lines behaved differently in support of C/Vic replication is interesting. We speculate that A549 cells may lack the cellular protein(s) that is critical

for the efficient replication of C/Vic. Alternatively, we can hypothesize that A549 cells may produce the novel cellular inhibitor(s) that suppresses the replication of C/Vic. Future experiments are needed to address these hypotheses. Furthermore, C/Vic was found to replicate well in the swine cell lines ST and MK1-OSU at 33 °C. ICV was previously isolated from pigs (Y. J. Guo et al., 1983) and shown to transmit between human and swine (H. Kimura et al., 1997). A novel influenza D virus distantly related to human ICV was also isolated in swine in 2011 (B. M. Hause et al., 2013). These data seemed to suggest that C/Vic has the potential to replicate in pigs. In contrast to ST and MK1-OSU, the third swine cell line, SD-PJEC, minimally supported C/Vic replication as the virus was not detected at 37 °C and titers reached only 1.0 $\log_{10}TCID_{50}$/ml at 33 °C at 72 and 96 hours post infection. Likewise, SD-PJEC also poorly supported human IBV replication (Z. Sun et al., 2012). It is worth mentioning that SD-PJEC is very susceptible to IAV infection. As such, this primary cell line can be useful to further elucidate restriction factors that affect the replication of ICV and IBV, which will be a focus of our future investigation.
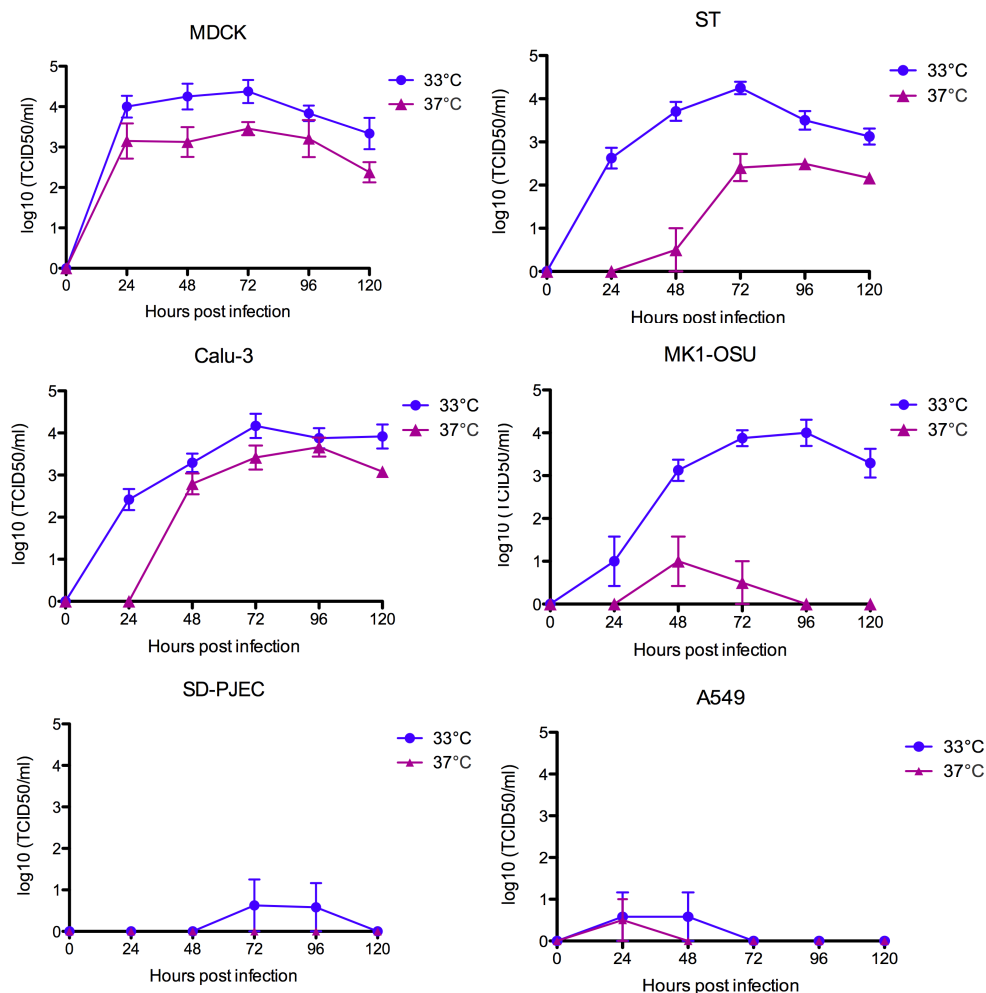
Figure 4.4. Growth kinetics of C/Victoria/2/2012 in different cell lines at 33 °C and 37 °C. MDCK, ST, MK1-OSU and SD-PJEC cells were infected with C/Victoria/2/2012 at an MOI of 0.01, while A549 and Calu-3 cells were infected with C/Victoria/2/2012 at an MOI of 1.0. Viral titers were determined in MDCK cells. The results presented are the mean values from three replicates with error bars indicated by SEM.

In summary, we have determined the genesis of a contemporary human C/Vic virus and characterized its evolutionary pathway. We also showed that C/Vic isolate replicated

more robustly at 33 °C than at 37 °C, which should be further investigated toward elucidating the molecular determinants of temperature-dependent growth. Finally, structural modeling work presented here has pinpointed two critical residues in the 170-loop of the HEF protein that are likely responsible for the observed antigenic differences among three ICV strains. The information described on this contemporary ICV here, as a whole shall aid in the further investigation of biology, evolution, and pathogenesis of ICV.

4.4 Acknowledgement

BIBLIOGRAPHY

Buonagurio, D. A., Nakada, S., Fitch, W. M., and Palese, P. (1986). Epidemiology of influenza C virus in man: multiple evolutionary lineages and low rate of change. Virology *153*, 12-21.

Chakraverty, P. (1974). The detection and multiplication of influenza C virus in tissue culture. The Journal of general virology *25*, 421-425.

Collin, E. A., Sheng, Z., Lang, Y., Ma, W., Hause, B. M., and Li, F. (2015). Cocirculation of two distinct genetic and antigenic lineages of proposed influenza D virus in cattle. Journal of virology *89*, 1036-1042.

Dykes, A. C., Cherry, J. D., and Nolan, C. E. (1980). A clinical, epidemiologic, serologic, and virologic study of influenza C virus infection. Archives of internal medicine *140*, 1295-1298.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research *32*, 1792-1797.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2007). Comparative protein structure modeling using MODELLER. Current protocols in protein science / editorial board, John E Coligan [et al] *Chapter 2*, Unit 2 9.

Furuse, Y., Matsuzaki, Y., Nishimura, H., and Oshitani, H. (2016). Analyses of Evolutionary Characteristics of the Hemagglutinin-Esterase Gene of Influenza C Virus during a Period of 68 Years Reveals Evolutionary Patterns Different from Influenza A and B Viruses. Viruses *8*.

Guo, Y. J., Jin, F. G., Wang, P., Wang, M., and Zhu, J. M. (1983). Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus. The Journal of general virology *64 (Pt 1)*, 177-182.

Hachinohe, S., Sugawara, K., Nishimura, H., Kitame, F., and Nakamura, K. (1989). Effect of anti-haemagglutinin-esterase glycoprotein monoclonal antibodies on the receptor-destroying activity of influenza C virus. The Journal of general virology *70 (Pt 5)*, 1287-1292.

Hause, B. M., Collin, E. A., Liu, R., Huang, B., Sheng, Z., Lu, W., Wang, D., Nelson, E. A., and Li, F. (2014). Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. mBio *5*, e00031-00014.

Hause, B. M., Ducatez, M., Collin, E. A., Ran, Z., Liu, R., Sheng, Z., Armien, A., Kaplan, B., Chakravarty, S., Hoppe, A. D., Webby, R. J., Simonson, R. R., and Li, F. (2013). Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens *9*, e1003176.

Katagiri, S., Ohizumi, A., and Homma, M. (1983). An outbreak of type C influenza in a children's home. The Journal of infectious diseases *148*, 51-56.

Kawamura, H., Tashiro, M., Kitame, F., Homma, M., and Nakamura, K. (1986). Genetic variation among human strains of influenza C virus isolated in Japan. Virus research *4*, 275-288.

Kimura, H., Abiko, C., Peng, G., Muraki, Y., Sugawara, K., Hongo, S., Kitame, F., Mizuta, K., Numazaki, Y., Suzuki, H., and Nakamura, K. (1997). Interspecies transmission of influenza C virus between humans and pigs. Virus research *48*, 71-79.

Matsuzaki, Y., Katsushima, N., Nagai, Y., Shoji, M., Itagaki, T., Sakamoto, M., Kitaoka, S., Mizuta, K., and Nishimura, H. (2006). Clinical features of influenza C virus infection in children. The Journal of infectious diseases *193*, 1229-1235.

Matsuzaki, Y., Mizuta, K., Kimura, H., Sugawara, K., Tsuchiya, E., Suzuki, H., Hongo, S., and Nakamura, K. (2000). Characterization of antigenically unique influenza C virus strains isolated in Yamagata and Sendai cities, Japan, during 1992-1993. The Journal of general virology *81*, 1447-1452.

Matsuzaki, Y., Mizuta, K., Sugawara, K., Tsuchiya, E., Muraki, Y., Hongo, S., Suzuki, H., and Nishimura, H. (2003). Frequent reassortment among influenza C viruses. Journal of virology *77*, 871-881.

Matsuzaki, Y., Muraki, Y., Sugawara, K., Hongo, S., Nishimura, H., Kitame, F., Katsushima, N., Numazaki, Y., and Nakamura, K. (1994). Cocirculation of two distinct groups of influenza C virus in Yamagata City, Japan. Virology *202*, 796-802.

Matsuzaki, Y., Sugawara, K., Abiko, C., Ikeda, T., Aoki, Y., Mizuta, K., Katsushima, N., Katsushima, F., Katsushima, Y., Itagaki, T., Shimotai, Y., Hongo, S., Muraki, Y., and Nishimura, H. (2014). Epidemiological information regarding the periodic epidemics of influenza C virus in Japan (1996-2013) and the seroprevalence of antibodies to different antigenic groups. J Clin Virol *61*, 87-93.

Matsuzaki, Y., Sugawara, K., Furuse, Y., Shimotai, Y., Hongo, S., Oshitani, H., Mizuta, K., and Nishimura, H. (2016). Genetic Lineage and Reassortment of Influenza C Viruses Circulating between 1947 and 2014. Journal of virology *90*, 8251-8265.

Moriuchi, H., Katsushima, N., Nishimura, H., Nakamura, K., and Numazaki, Y. (1991). Community-acquired influenza C virus infection in children. The Journal of pediatrics *118*, 235-238.

Murakami, S., Endoh, M., Kobayashi, T., Takenaka-Uema, A., Chambers, J. K., Uchida, K., Nishihara, M., Hause, B., and Horimoto, T. (2016). Influenza D Virus Infection in Herd of Cattle, Japan. Emerging infectious diseases *22*, 1517-1519.

Muraki, Y., Hongo, S., Sugawara, K., Kitame, F., and Nakamura, K. (1996). Evolution of the haemagglutinin-esterase gene of influenza C virus. The Journal of general virology *77 (Pt 4)*, 673-679.

Nishimura, H., Sugawara, K., Kitame, F., Nakamura, K., and Sasaki, H. (1987). Prevalence of the antibody to influenza C virus in a northern Luzon Highland Village, Philippines. Microbiology and immunology *31*, 1137-1143.

Odagiri, T., Matsuzaki, Y., Okamoto, M., Suzuki, A., Saito, M., Tamaki, R., Lupisan, S. P., Sombrero, L. T., Hongo, S., and Oshitani, H. (2015). Isolation and characterization of influenza C viruses in the Philippines and Japan. Journal of clinical microbiology *53*, 847-858.

Rosenthal, P. B., Zhang, X., Formanowski, F., Fitz, W., Wong, C. H., Meier-Ewert, H., Skehel, J. J., and Wiley, D. C. (1998). Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. Nature *396*, 92-96.

Roy Mukherjee, T., Mukherjee, A., Mullick, S., and Chawla-Sarkar, M. (2013). Full genome analysis and characterization of influenza C virus identified in Eastern India. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases *16*, 419-425.

Sheng, Z., Ran, Z., Wang, D., Hoppe, A. D., Simonson, R., Chakravarty, S., Hause, B. M., and Li, F. (2014). Genomic and evolutionary characterization of a novel influenza-C-like virus from swine. Archives of virology *159*, 249-255.

Sugawara, K., Kitame, F., Nishimura, H., and Nakamura, K. (1988). Operational and topological analyses of antigenic sites on influenza C virus glycoprotein and their dependence on glycosylation. The Journal of general virology *69 (Pt 3)*, 537-547.

Sugawara, K., Nishimura, H., Hongo, S., Muraki, Y., Kitame, F., and Nakamura, K. (1993). Construction of an antigenic map of the haemagglutinin-esterase protein of influenza C virus. The Journal of general virology *74 (Pt 8)*, 1661-1666.

Sun, Z., Huber, V. C., McCormick, K., Kaushik, R. S., Boon, A. C., Zhu, L., Hause, B., Webby, R. J., and Fang, Y. (2012). Characterization of a porcine intestinal epithelial cell line for influenza virus production. The Journal of general virology *93*, 2008-2016.

Tanaka, S., Aoki, Y., Matoba, Y., Yahagi, K., Mizuta, K., Itagaki, T., Katsushima, F., Katsushima, Y., and Matsuzaki, Y. (2015). The dominant antigenic group of influenza C infections changed from c/Sao Paulo/378/82-lineage to c/Kanagawa/1/76-lineage in Yamagata, Japan, in 2014. Jpn J Infect Dis *68*, 166-168.

Yano, T., Maeda, C., Akachi, S., Matsuno, Y., Yamadera, M., Kobayashi, T., Nagai, Y., Iwade, Y., Kusuhara, H., Katayama, M., Fukuta, M., Nakagawa, Y., Naraya, S., Takahashi, H., Hiraoka, M., Yamauchi, A., Nishinaka, T., Amano, H., Yamaguchi, T., Ochiai, H., Ihara, T., and Matsuzaki, Y. (2014). Phylogenetic analysis and seroprevalence of influenza C virus in Mie Prefecture, Japan in 2012. Jpn J Infect Dis *67*, 127-131.