Computational analysis for infectious diseases surveillance and host-pathogen interactions (in the context of Influenza A viruses)

Beyene, Biruhalem Taye

2017

Beyene, B.T. (2017). Computational analysis for infectious diseases surveillance and host-pathogen interactions (in the context of Influenza A viruses). Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/72814>

<https://doi.org/10.32657/10356/72814>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

Computational analysis for infectious diseases surveillance and host-pathogen interactions (in the context of Influenza A viruses)

Biruhalem Taye Beyene

SCHOOL OF BIOLOGICAL SCIENCES

2017

Computational analysis for infectious diseases surveillance and host-pathogen interactions (in the context of Influenza A viruses)

Biruhalem Taye Beyene

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2017

Table of contents

Table of contents	iii
Acknowledgments	vii
List of tables.....	viii
List of figures.....	ix
List of abbreviations	xi
Abstract.....	xiii
Chapter 1: Introduction	1
1.1 Epidemiology of influenza A virus (IAV)	1
1.1.1 Pandemic IAV	2
1.1.2 Seasonal influenza	3
1.1.3 Sporadic or Zoonotic IAV	4
1.2 Molecular biology of IAV.....	4
1.2.1 Morphology of IAV	4
1.2.2 Structural components of IAV.....	6
1.2.3 IAV genes and their functions.....	6
1.3 IAV life cycle.....	13
1.3.1 Virus entry	14
1.3.2 Transcription and replication.....	16
1.3.3 Viral assembly and budding	17
1.4 IAV host adaptation and virulence.....	18
1.5 Host responses to IAV infections.....	21
1.5.1 Innate immune response	21
1.5.2 Adaptive immune response to IAV infections	26
1.5.3 IAV escape strategies from the immune response.....	27
1.6 Genome-wide host gene expressions in IAV infections	27
1.6.1 Host gene expressions in seasonal (H1N1 and H3N2) IAV infections.....	28
1.6.2 Host gene expressions in pandemic IAV infections	29
1.6.3 Host gene expressions in HPAI infections	30

1.7 Genome-wide siRNA screens for identification of important IAV host factors	32
1.8 IAV interactome	34
1.9 Anti-IAVs therapy	35
1.9.1 Anti-influenza virus inhibitors targeting the viral proteins	35
1.9.2. Host-targeted antiviral	38
1.10 Rational of the study	40
1.11 Hypothesis	41
1.12 Objectives	42
Chapter 2: Materials and methods.....	43
2.1 Data sources	43
2.1.1 Microarray data	43
2.1.2 Genome-wide siRNA screen data	44
2.1.3 IAV interactome data	44
2.3 Microarray data analysis	45
2.4 Functional annotations (GO BP and KEGG pathway analysis)	45
2.2 Databases and software packages used in this study	48
2.5 Integration of microarray, interactome and siRNA screening data set	49
2.6 Probes annotation and orthologs mapping	49
2.7 Benchmarking network growing tools using IHFs	51
2.8 Drug-host target interactions	53
2.9 Prioritising host targeted FDA approved drugs	53
2.10 Statistical analysis	55
Chapter 3: Identification of IAV-therapeutic plausible host targets through integration of genome-wide transcriptome, siRNA screens and Interactome data set.....	56
3.1 Introduction	56
3.2 Results	57
3.2.1 Host gene expression changes in A549 cells infected with different IAVs	57
3.2.2 Pathways and GO BP of the DEGs	59
3.2.3 Integration of the transcriptome, siRNA screens and interactome data set	62
3.2.4 Integration of the three datasets that used the same cell type (A549 cells)	66

3.2.5 Integration of the data set that used the same virus strain (H1N1/WSN virus and cell type (A549 cells)	67
3.2.6 Drugs targeting the plausible host factors	68
3.3 Discussions.....	68
3.4 Conclusions	74
Chapter 4. Comparison of gene expression signatures during IAV infections of different species cell lines	75
4.1 Introduction	75
4.2 Results	76
4.2.1 Annotation and orthologs mapping	76
4.2.2 Comparison of DEGs in A549 with MDCK and CEF cells	79
4.2.3 Functional annotation of DEGs at 10hpi	82
4.2.4 Correlation and clustering of DEGs at 10hpi	84
4.2.5 Integration of the DEGs with interactome and siRNA screens	86
4.2.6 DEGs associated with metabolic pathways	86
4.3 Discussion	90
4.4 Conclusion.....	97
Chapter 5: Systems-based approach to examine the cytokine responses in primary mouse lung macrophages infected with low pathogenic avian Influenza virus circulating in South East Asia.	98
5.1 Introduction	98
5.2 Results	99
5.2.1 Global changes in host gene expressions at 24 hours post-infection.	99
5.2.2 Functional annotation of DEGs	102
5.2.3 Host response to the IAVs infection.....	104
5.2.4 IAV induced cytokines	105
5.2.5 Interferon stimulated genes (ISGs) and cell death.....	105
5.2.6 Macrophage activation	106
5.2.7 DEGs associated with metabolic pathways	108
5.3 Discussion	108
5.4 Conclusion.....	113

Chapter 6: Benchmarking selected computational gene network growing tools in the context of virus-host interactions	114
6.1 Introduction	114
6.2 Results	115
6.2.1 The siRNA screening studies show low reproducibility.	115
6.2.2 Network growing algorithms partially recover IHFs from siRNA screens.....	117
6.2.3 The network growing tools recruit other functionally related IHFs.	119
6.2.4 Performance of IHFs detection varies between the network growing tools due to different data sources used.	120
6.2.5 Network topology analysis shows characteristic differences between networks from different tools.....	122
6.2.6 Interchanging and combination of genes didn't improve the performance of the free tools	124
6.2.7 Network growing algorithms detect GO biological process and pathways of the known IHFs	125
6.2.8 Network growing approach on existing data could be as powerful as new siRNA screens for finding relevant additional IHFs	127
6.2.9 The known and the new candidate host factors could be targeted by FDA-approved drugs	127
6.2.10 Ranking the FDA-approved drugs identified in this study.....	129
6.3 Discussion	130
6.4 Conclusion.....	132
Chapter 7: General conclusions.....	133
Chapter 8: Future work	136
8.1 Wet lab validations.....	136
8.2 Computational future work	137
References	138
Appendix.....	176
Author's publications.....	177
Posters, Awards and invited talks	180

Acknowledgments

My deepest gratitude goes to my supervisors Dr. Sebastian Maurer-Stroh and Associate Professor Richard J Sugrue for their continuous guidance, comments, support and patience during the entire period of my PhD study. I would like to thank members of my thesis advisory committee Dr. Vivek Tanavde, Associate Professor Liu Ding Xiang, Associate Professor Mu Yuguang and Dr. Peter Bond.

I am also grateful to my labmates who helped me in my bioinformatics (Lee Tze Chuen (Raphael), Dr. Vithiagarun Gunalan, Dr. Candida Vaz, Dr. Kenanov Dimitar, and Han Xiaochuan Alvin) and laboratory training (Dr. Myint Zu Myaing). I am also extending my gratitude to Dr. Vithiagarun Gunalan, Han Xiaochuan Alvin and Baker James Alexander who helped me in proofreading this thesis. Thanks are also given to all of my friends especially to Dr. Amsalu Efreem who took care of me while I was hospitalised in Singapore due to an unexpected accident.

I would also like to thank Singapore International Graduate Award (SINGA), Bioinformatics Institute A*STAR and Nanyang Technological University for their funding and giving me the opportunity to pursue my PhD.

Last but not least, I wish to express my heartfelt gratitude to my wife Rahel Kebede and my son Nahum Biruhalem for their continuous support, encouragement and patience during the whole period of my PhD journey.

List of tables

Table 1.1: IAV genomic segments, proteins and their functions.....	7
Table 1.2: IAV viral protein mutations associated with host adaptation and virulence	19
Table 2.1: The characteristics of microarray data sets used in the current study	43
Table 2.2: Summary table for siRNA screening studies.....	46
Table 2.3: Summary table for interactome studies	47
Table 2.4: Databases and software packages used in the current study.....	48
Table 4.1: Annotation of affymetrix probe sets with no ensembl IDs.....	78
Table 4.2: Human ortholog mapping of <i>Canis familiaris</i> (Canine) and <i>Gallus gallus</i> (Chicken) genes/probes.....	79
Table 4.3: Comparison of ortholog DEGs in IAV infected A549 and MDCK cells	80
Table 4.4: Comparison of ortholog DEGs in IAV infected A549 and CEF cells.....	81
Table 6.1: The top 10 FDA-approved drugs of the 343 drugs that target the new candidate and known IHFs (ranked based on different criteria).....	130

List of figures

Chapter 1

Figure 1.1: First human case reports of zoonotic and pandemic IAV infections in 20 th and 21 st century.....	3
Figure 1.2: Schematic representation of IAV structure (adapted from Nelson et al., 2007; Nature Review Genetics, copy right permission number 4140250559702) ⁴⁶	6
Figure 1.3: The schematic structure of IAV proteins	9

Figure 1.4: Schematic representation IAV lifecycle (adapted from Te Velthuis et al., 2016; Nature Reviews Microbiology, copy right number 4140240593811) ⁹³	14
--	----

Chapter 2

Figure 2.1: Overview of Affymetrix probe annotation and orthologs mapping.....	50
--	----

Chapter 3

Figure 3.1: Kinetics of up- and down-regulated genes in A549 cells infected with four IAVs compared to mock infections.....	58
Figure 3.2: DEGs in A549 cells infections at 10hpi compared to mock-infections.....	59
Figure 3.3: The common DEGs across all IAV infections and their extent of expressions in A549 cells at 10hpi compared to mock-infections.....	60
Figure 3.4: The top 20 Pathway and GO BPs of the DEGs in IAV infected A549 cells compared to mock-infections.....	61
Figure 3.5: Integration of all DEGs at 10hpi in A549 cell infections with interactome and siRNA screens.....	62
Figure 3.6: Integration of DEGs in at least 3 IAV infections in A549 cells with host factors identified by interactome and siRNA screens.....	64
Figure 3.7: Interactions of IAV viral proteins with selected host factors.....	65
Figure 3.8: Integration analysis of the three data set that were performed using A549 cells.....	66
Figure 3.9: Integration of the three data set performed using the same virus strain (H1N1/WSN) and cell type (A549 cells).....	67

Chapter 4

Figure 4.1: Functional annotation of the DEGs in A549, CEF and MDCK cell lines during IAV infections at 10hpi compared to mock infections.	83
Figure 4.2: Correlation and clustering of the DEGs.	84
Figure 4.3: The hierarchical clustering of the top 56 DEGs across the cell lines and virus infections compared to mock infections.	85
Figure 4.4: Integration of the DEGs in the three cell types during the IAV infections at 10hpi with interactome and siRNA screens data set.	87
Figure 4.5: DEGs and pathways involved in metabolic pathways.	88
Figure 4.6: Comparison of metabolic pathways between the pro-viral siRNA screens and DEGs.	89

Chapter 5

Figure 5.1: Global host gene expression in mouse PM Φ infected with IAVs at 24 hours-post infection compared to mock infections.	101
Figure 5.2: Top 20 statistically significant ($q < 0.05$) functionally enriched pathways or/and GO biological process (GBP) of the DEGs.	103
Figure 5.3: Selected host responses of PM Φ in IAV infections compared to mock infections.	104
Figure 5.4: DEGs associated with metabolic pathways.	107

Chapter 6

Figure 6.1: Meta-analysis of IHFs from siRNA screening studies.	116
Figure 6.2: The trend of network growing tools in detecting known IHFs.	118
Figure 6.3: Detection rate of network growing tools after randomization of IHFs and random human proteins.	119
Figure 6.4: Comparison grown nodes and their edges.	121
Figure 6.5: Network topology parameters, GOBP and pathways of grown genes.	123
Figure 6.6: The detection rate of STRING and GeneMANIA after combination and interchange of the 1st 50 grown genes.	124
Figure 6.7: Performance comparison of the three network growing tools in terms of KEGG pathways and GO BP.	126
Figure 6.8: Comparison of siRNA experimental studies and network growing tools in detecting known IHFs.	127
Figure 6.9: The known and the new candidate host factors could be targeted by FDA-approved drugs.	128

List of abbreviations

aa	Amino acids
A549 cells	Adenocarcinoma human alveolar basal epithelial cells
ADCC	Antibody-dependent cell cytotoxicity
AIDS	Acquired immune deficiency syndrome
CEF	Chicken embryo fibroblast cells
CTLs	Cytotoxic T lymphocytes
DAVID	The Database for Annotation, Visualization and Integrated Discovery
DCs	Dendritic cells
DEGs	Differentially expressed genes
FDA	US Food and Drug Administration
GO BP	Gene Ontology Biological Process
HA	Hemagglutinin
HGNC	HUGO Gene Nomenclature Committee
HIV	Human immunodeficiency virus
HPAI	Highly pathogenic avian IV
IAV	Influenza A virus
IFN	Interferon
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IHF	Influenza A virus required host factors
IPA	Ingenuity Pathway Analysis
ISGs	IFN-stimulated genes
KEGG	Kyoto Encyclopedia of Genes and Genomes
LPAI	Low pathogenic avian IV
MΦ	Macrophages
M1	Influenza A virus matrix protein
M1MΦ	M1 macrophages
M2	Influenza A virus ion channel protein

M2MΦ	M2 macrophages
MBL	Mannose-binding lectins
MDCK cells	Madin-Darby canine kidney cells
NA	Neuraminidase
NK cells	Natural killer cells
NOD	Nucleotide binding oligomerization domain
NP	Nucleoprotein
NS1	Non-structural protein 1
NS2	Non-structural protein 2
OMA	Orthologous MAtrix
PA	Polymerase acidic
PB1	Polymerase basic 1
PB2	Polymerase basic 2
PMΦ	Pulmonary murine lung macrophages
PPIs	Protein-protein interactions
PPRs	Pattern recognition receptors
RBS	Receptor binding site
RIG-I	Retinoic acid-inducible gene I
ROS	Reactive oxygen species
RSV	Respiratory syncytial virus
RT-PCR	Reverse-transcription polymerase chain reactions
SA	Sialic acid
siRNA	small RNA interference
SP	Surfactant protein
STRING	The Search Tool for the Retrieval of Interacting Genes
TLRs	Toll-like receptors
vRNA	Viral RNA
vRNPs	Viral ribonucleoproteins
WHO	World Health Organization

Abstract

Influenza A virus (IAV) is a major public health problem responsible for the death of half a million people every year worldwide. Zoonotic transmissions of the virus from swine and avian origin have occurred and can in the worst-case result in pandemics. Prediction of the next pandemic strain is still a major challenge as mechanisms of antigenic shift and the zoonotic nature of influenza viruses are still poorly understood. On the other hand, the current vaccination strategy treating seasonal influenza viruses, given its own problems in efficacy, is not well suited to mitigate an impending pandemic. While there are few clinically approved anti-IAV drugs for therapeutic and prophylactic use, these drugs have also been challenged by the emergence of drug resistance, toxicity, adverse effects and low efficacy. Hence, development of a broad range anti-IAV drug that could target all of the IAV strains irrespective of their source is critically important.

To discover effective IAV vaccines and alternative anti-IAV therapeutics, it is crucial to understand how IAVs adapt to different species' host ranges and vice versa, how different species hosts respond to different IAV subtypes. Additionally, it is crucial to understand IAV-host interactions and identifying critical viral or host factors that could support the replication of the virus or induce pathologic conditions. Recently, high throughput technologies such as mRNA microarray-based gene expression, genome-wide siRNA screening, and proteomic analysis are providing in-depth insights into host-pathogen interactions of IAVs. Hence, this study aimed to computationally investigate IAV-host interactions using three data set (transcriptome, genome-wide siRNA screens and interactome) and identify 1) virus and host specific responses, 2) host determinants in host adaptations and 3) host targets for IAV therapeutics.

This thesis contains four main projects. In the first project (chapter 3), we investigated the host gene expression changes of eight IAVs (H1N1/WSN, pH1N1, H5N2/F118, H5N2/F59, H5N2/F189, H5N3, H7N9 and H9N2 viruses) in A549 cells at different time points of infections. Then we integrated the differentially expressed genes (DEGs) in at least 3 viruses, with 1713 and 1780 host factors comprehensively curated from 11 siRNA and 13 interactome studies respectively. The integration of the three data set highlighted plausible influenza A virus required host factors (IHF) that could be targeted against IAVs. The up-regulated IHFs (e.g. *TRIM21*, *TRIM26*, *IRF2*, and *SAMHD1*) might support the replication of IAV through suppression of the

innate and adaptive antiviral immune responses. The other up-regulated IHFs could also enhance the replication of IAV at different stages of the virus lifecycle: endocytosis (*BTC*), prevention of apoptosis (*TIMM17A*), nuclear import (*JAK2*), and translation elongation of viral proteins (*HEXIM1*). Although several of these IHFs have been implicated in other viruses, the detailed mechanisms of how several of these up-regulated IHFs could support IAV replication require further investigation.

The second project (chapter 4), explored a comprehensive analysis of host gene expression changes in different IAV infections in different host species. We used host gene expression signatures of cell lines from three species (A549 (human), CEF (chicken), and MDCK (canine) in response to six IAVs (H1N1/WSN, H5N2/F59, H5N2/F118, H5N2/F189, H5N3 and H9N2 viruses). To compare the expression changes between the three species, we performed comprehensive probe set re-annotation and human ortholog mapping. The result showed that the expression signatures of different IAV isolates in a single species cell type are more similar to each other compared to the expression signatures of a single isolate in different species' cell types. The functional annotations (pathways) and the highly expressed cell-specific signatures indicated that IAVs up-regulated host factors could induce virus infectivity (e.g. *OSBPL1A* and *ARHGAP21*), reduce apoptosis (e.g. *MRPS27*) and increase cell proliferation (e.g. *COPS2*) in CEF cells. Conversely, increased antiviral, pro-apoptotic and inflammatory signatures have been identified in A549 cells. Except in H5N3 virus infections, generally IAV infections down-regulates genes associated with cell cycle and metabolic pathways with the strongest effect in MDCK cells, followed by A549 cells in a strain dependent manner, but not in CEF cells. Previously our group demonstrated that the replication of the examined viruses was significantly higher in CEF cells than the other cell types. Thus, we hypothesise that this could partially explain the mechanism how infectious LPAI viruses shed by chickens lack the inflammatory response and cellular disruption that may lead to disease conditions.

In the third project (chapter 5), we used a systems-based approach to investigate changes to the transcriptome of primary murine lung macrophages (PMΦ) in response to infection with the mouse-adapted H1N1/WSN virus and low pathogenic avian influenza (LPAI) viruses H5N2 and H5N3. The results showed that while all viruses induced antiviral responses, the H5N3 infection resulted in higher expression levels of cytokines and chemokines associated with inflammatory responses. Previously, our group showed that the LPAI H5N2 and H5N3 were able

to infect murine lung macrophages and together with increased expression of inflammatory mediators particularly in H5N3 virus could impose threats to human health in the future.

The IHFs identified by the IAV genome-wide siRNA screening studies could be used as potential anti-IV targets. However, these studies were not consistent (reproducible) mainly due to false negative results. Hence, in the fourth project (chapter 6), we applied computational gene network growing for discovering gene network links that could have been missed by these experimental investigations. Using the known IHFs we compared the network growing function of two free tools GeneMANIA and STRING and the commercial IPA for their performance of recovering other known IHFs previously identified from siRNA screens. The result showed that, given small (~30 genes) or medium (~150 genes) input sets, all three network growing tools detected significantly more known host factors than random human genes with STRING overall performing strongest. Extending the networks with all the three tools significantly improved the detection of GO biological processes of known host factors compared to not growing networks. Notably, the rate of identification of true host factors using computational network growing is equal or better to doing another experimental siRNA screening study which could also be true and applied to other biological pathways/processes.

Finally, using the known IHFs and the "new IHF candidates" (genes connected to the IHFs from the network growing analysis), we predicted computational drug-target interactions using MetaCore. We identified 343 US Food and Drug Administration (FDA) approved drugs that had an inhibitory effect on either the known or new candidate IHFs, of which 258 were new predictions. Furthermore, using different criteria, we computationally ranked the 343 FDA approved drugs for further experimental validation.

Chapter 1: Introduction

1.1 Epidemiology of influenza A virus (IAV)

Influenza is an infectious disease caused by an Influenza virus and is a threat to global public health and economy. The four types of influenza viruses include type A, B, C and D. Type A and B are the main cause of human infections. Influenza A virus (IAV) also affects birds, swine and other animals and spreads between different species and causes zoonotic, epizootic, and pandemic infections in human and other animals^{1,2}.

Influenza B virus (IBV) is distributed worldwide and causes seasonal epidemics along with IAV. It causes similar clinical symptoms as IAV, however more severe cases were observed in children and adolescents. IBV has 8 genomic segments coding for 11 known proteins. There are two lineages of IBV, the Victoria and Yamagata lineages. The prevalence of Victoria lineage is higher in a younger population than the Yamagata lineages (reviewed in³). Influenza C virus (ICV) also spreads worldwide causing mild upper and lower respiratory tract infections in humans^{4,5}. ICV also infects pigs and dogs, but its zoonotic transmission has not been established yet. Unlike IAV and IBV, ICV contains seven genomic segments⁵. In 2011, Influenza D virus (IDV) has been isolated in the USA from pigs. Several studies also demonstrated the spread of IDV in cattle (natural reservoir), in sheep and goats. The zoonotic potential of this virus has been demonstrated by seroconversion in human and its replication in upper respiratory tracts of ferrets⁶.

IAV transmits through droplets, aerosols⁷, and contact and has a median incubation period of 1.4 days before the onset of symptoms⁸. Its infection is characterised by a range of severity (from asymptomatic to severe viral pneumonia)¹. The main clinical symptoms in IAV infected individuals are nasal congestion, fever and cough. Other symptoms include headache, pharyngitis, dyspnea, myalgia, chest pain, and pneumonia⁹. Elderly and young, pregnant women and individuals with immunodeficiency, chronic diseases, and asthma are at higher risk of developing complications and severe form of diseases than other healthy individuals^{1,2,9,10}. Half of the hospitalized patients due to IAV infections developed lymphopenia and thrombocytopenia (reviewed in¹⁰). Clinical diagnosis is not specific and accurate due to the overlap of symptoms with other respiratory viruses. Several laboratory techniques are available for the diagnosis of

IAV infections including virus isolation in a cell or embryonated chicken eggs, immunofluorescence assays, reverse-transcription polymerase chain reactions (RT-PCR), rapid antigen diagnostic tests etc (reviewed in^{10,11}).

Since the first reports of IAV infection in 1918, more than 15 subtypes of IAVs have been reported to cause human infections (Fig. 1.1). Most of the subtypes were of avian and swine origin and mainly originated in either the USA or China (Fig. 1.1).

1.1.1 Pandemic IAV

Influenza virus pandemic is a worldwide outbreak of the virus in human. In recent history since 1918, there have been four known pandemic outbreaks of IAV reported so far, the 1918 "Spanish Flu" (H1N1), the 1958 Asian pandemic (H2N2), the Hong Kong pandemic 1968 (H3N2), and the most recent the 2009 swine-origin pandemic H1N1. The "Spanish Flu" was demonstrated to have the avian-descendent H1N1 virus. It was the most devastating example which killed approximately 50 million people worldwide. The high mortality has been reported to have an association with secondary bacterial infections or pneumonia. The virus re-emerged in 1977 and continues to cause seasonal epidemics each year in humans (reviewed in^{12,13}). The Asian pandemic flu was the result of acquiring three gene segments from unknown avian-like virus (HA (H2), NA(N2), and PB1 genes) to the circulating 1918 H1N1 virus through reassortment. It was also responsible for the death of 1-2 million people (reviewed in^{12,13}). The reassortment of the circulating 1958 H2N2 virus with the avian H3 and PB1 segments resulted in the novel virus the 1968 pandemic H3N2 virus which killed 0.5-1 million people (reviewed in^{12,13}). The 2009 pH1N1 virus was the result of the reassortment of two distinct swine viruses (the triple reassortant swine-lineage H1N2 and the European avian-like H1N1 viruses) (reviewed in¹³). It was estimated that the 2009 H1N1 pandemic virus caused approximately 300,000 deaths¹⁴.

Figure 1.1: First human case reports of zoonotic and pandemic IAV infections in 20th and 21st century.

Since 20th century (1918), there have been four IAV pandemics (highlighted light blue) in human. In addition, several zoonotic infections of IAV have been occurring mainly from avian (highlighted red) and swine (highlighted green) origin.

1.1.2 Seasonal influenza

The extensive capacity of the IAV to undergo antigenic drift and immune evasion resulted in seasonal epidemics¹⁵. The currently circulating seasonal IAV subtypes H1N1, H3N2 and H1N1pdm09 viruses were the descendants of the three pandemic viruses, the 1918 H1N1, the 1968 H3N2 and the 2009 pandemic H1N1 viruses respectively¹⁵. Seasonal epidemics are characteristic of temperate regions particularly in winter, but in tropical regions it may cause irregular outbreaks throughout the year¹. Annually seasonal influenza affects up to 10% and 30% of adults and children of the world population respectively. Worldwide, the seasonal epidemics are estimated to cause approximately 3 to 5 million severe cases and half a million deaths annually¹.

1.1.3 Sporadic or Zoonotic IAV

While zoonotic transmission is not common, human infections of IAVs from swine (H1N1 and H3N2 viruses) and avian origin have been documented (Fig. 1.1)^{16,17}. Avian IAVs have been classified either as low pathogenic (LPAI) or highly pathogenic (HPAI) viruses associated with mild and severe disease in poultry respectively. Examples of HPAI subtypes include H5 and H7 viruses (H5N1, H7N7 and H7N3), while H9N2, H7N2 and H7N10 viruses were regarded as LPAI. However, clades or strains of certain subtypes (e.g. H5N2 and H7N7) have both HPAI and LPAI virus and the mechanism of conversion of these viruses from LPAI to HPAI is not always understood¹⁸. In addition to the key role in the emergence of human pandemic IAV, sporadic human infection of both LPAI and HPAI were also reported^{19,20}. Earlier studies documented the transmission of avian virus (the so called " Fowl plague virus" in 1960's) to humans^{21,22}. In 1997, the direct transmission of H5N1 subtype virus from live poultry to human with the fatality rate of 75% was reported in Hong Kong, spreading and presenting major public health problems¹⁷. As of February 2017, the virus had spread to 16 countries with more than 856 confirmed cases and 459 deaths (WHO)²³. Subsequently, the spread of subtype H9N2 virus was reported from the same year and city²⁴. In March 2013, human IAV cases caused by a novel subtype LPAI H7N9 virus were reported in China^{25,26}. Since then, 1258 human infections of H7N9 virus have been reported to the WHO from China with seasonal dynamics and a mortality rate of 41% has been reported from the first four outbreaks²³.

Besides its public health impact, IAV can affect economy as well. This is attributed to both direct and indirect medical costs as well as costs associated with production loss. For example, for the vaccination of seasonal influenza, billions of dollars were spent annually in the US²⁷. Also, economic estimates from HPAI outbreaks in Ghana showed the loss of 99.88% of poultry stock valued at US\$ 416 million²⁸. Similar examples of reduced production and losses due to the virus were also reported in Indonesia²⁹ and Australia³⁰.

1.2 Molecular biology of IAV

1.2.1 Morphology of IAV

Influenza viruses are negative-stranded RNA viruses under the family of *Orthomyxoviridae* with 6–8 genomic segments³¹. The structure of Orthomyxoviruses have similarities in that they have a lipid bilayer membrane surrounded by glycoprotein spikes on the

surface, and a matrix (M) protein layer that contains a ribonucleoprotein complex³². The classification of influenza viruses (type A, B and C) was based on the difference in serological reaction of M (matrix) and NP (nucleoprotein) proteins³². However, several other differences exist between the three types of influenza viruses, which includes epidemiological distributions, clinical features, number and length of genomic RNA segments, degree of surface antigenic variability, uncoating mechanisms, and polymerase (PA) proteins³²⁻³⁴. Recently, a new class of influenza virus (Influenza virus D) was identified in cattle and swine. It has 7 genomic segments like influenza C and was found to cause mild respiratory disease^{35,36}.

IAV is a pleomorphic virus with spherical, elliptical, or filamentous structures and size ranging from 80 to 120 nm^{37,38}. There are growing evidences that filamentous virus morphology may be clinically relevant as they are readily produced by human clinical infections. It was suggested that filamentous have increased NA protein activity that could clear respiratory mucus for enhanced infections and spread³⁸. Several studies showed that such morphological variation was associated with the influenza virus matrix (M1) protein and the ion channel protein (M2). The non-filamentous virus particles were able to generate virus filaments after the reassortment of M segments from the filamentous virus, indicating the role of M1 and M2 proteins in the morphogenesis of the virus^{39,40}. M1 protein is the main viral determinant for filamentous morphology. Specific amino acid residues of the M1 protein (30, 207 and 209) were responsible for the formation of filamentous viral particle and mutation of these residues resulted in a change in morphology and virus yield of the virus⁴¹. Disruption of interactions between M2 protein cytoplasmic tail and M1 protein through mutation of the conserved hydrophobic residues of the M2 protein cytoplasmic tail inhibited filamentous virus formation⁴². In addition, the M2 protein cytoplasmic end binds to cholesterol, suggesting that this binding could be essential for the formation and stability of viral filaments⁴³.

The morphology of IAV changed through adaptation to different hosts suggesting the involvement of host factors in the formation of the filamentous virus³⁸. The disruption of actin cytoskeleton reduced budding filamentous virus without the effect on the release of spherical viruses⁴⁴. On the other hand Ras-related protein (*Rab11*) and its association with family interacting protein 3 (FIP3) are important for the formation of the filamentous virus⁴⁵, however the mechanism is not known yet. Taken together, both viral and host factors can influence the

morphology of IAV and further investigations are needed for a detailed understanding of the molecular mechanisms³⁸.

1.2.2 Structural components of IAV

IAV is an enveloped virus surrounded by hemagglutinin (HA), neuraminidase (NA) glycoproteins and the M2 proteins (Fig. 1.2)⁴⁶. The HA and NA proteins are found on the surface of the virion, while the M2 protein crosses the lipid membrane. The M1 protein forms a matrix layer underneath the lipid and that contains the eight rod-like ribonucleoprotein complexes (RNPs) of different size. Each RNP is composed of a viral RNA, a viral polymerase (polymerase basic protein 1(PB1), PB2, polymerase acidic (PA)) and several copies of the viral nucleoprotein (NP) (Fig. 1.2)^{46,47}.

Figure 1.2: Schematic representation of IAV structure (adapted from Nelson et al., 2007; Nature Review Genetics, copy right permission number 4140250559702)⁴⁶.

IAV is an enveloped virus surrounded by two spikes of glycoproteins HA and NA. M1 protein separates the viral membrane from the matrix, while M2 protein is a transmembrane pore which connects the outer membrane and matrix. The matrix consists of eight segments of the viral genome each bound to viral polymerase complex and NP proteins.

1.2.3 IAV genes and their functions

The eight segments of IAV genome encode for seventeen viral proteins (Table 1.1)⁴⁸. Segment 1 (PB2), 4 (HA), 5 (NP), 6 (NA) codes for a single protein, while other segments expressed more than one protein. An alternative splicing of segment 7 and 8 expresses for two (M1 and M2) and three (NS1, NS2 and NS3) viral proteins respectively (Table 1.1)⁴⁸. In addition to the PB1 and PA proteins, the ribosomal frame-shifting and/or leaky ribosomal scanning of the non-spliced mRNA of segment 2 and 3 resulted in additional accessory proteins (Table 1.1)⁴⁹.

Table 1.1: IAV genomic segments, proteins and their functions

Segments (No. of na)	Proteins (No. aa)	Functions	References
1(2316)	PB2(759)	Required for transcription and repletion of vRNA. Associated with pathogenesis and host range restriction. Proposed to have a role in apoptosis and innate immune responses.	50-53
2(2316)	PB1(757)	Core component of polymerase complex. Required for transcription and replication of vRNA	54-56
	PB1-F2 (90)	Has endonuclease cleavage activity (cap-snatching) of cellular mRNA Translocate into mitochondria and induces ROS, apoptosis, <i>NLRP3</i> inflammasome complex and increased inflammations that lead to increased pathogenesis.	57-60
	PB1-N40 (718)	Unknown	
3(2209)	PA(716)	Viral replication and transcription. Has endonuclease cleavage activity.	61
	PA-X(252)	Endonucleolytic activity. Repress cellular gene expressions (host shutoff activity). Decrease the virulence of H5N1 through damping the polymerase activity, viral replication, and host innate immune responses.	62-65
	PA-N115(568)	Unknown	
	PA-N182(535)	Unknown	
4(1752)	HA(566)	Recognition and binding to host cell receptor sialic acid. Fusion between viral and endosomal membrane.	66,67
5(1540)	NP(498)	Activate replication and transcription of vRNA. Promotes nuclear export of vRNP. Mediates vRNA M1 homo-oligomerization.	
6(1433)	NA(454)	Cleaves the glycosidic bond of sialic acid for the release the virus. Have a role in IAV receptor binding activity. Facilitate IAV infection by cleaving sialic acid of mucins.	68-70
7(1002)	M1(252)	Determinant for the morphology of IAV Intracellular trafficking of RNPs and viral assembly	71-73
	M2(97)	Proton selective ion channel, acidification and uncoating of the virus. Has a role in virus assembly and budding. Promotes maturation of HA by regulating the pH of trans-Golgi network.	73-77
8(865)	NS1(230)	Inhibit host anti-viral responses and modulate the replication, protein synthesis, transmissibility and pathogenesis.	78-81
	NS2(121)	Mediate the nuclear export of vRNPs Regulation of transcription and translation of genomic vRNA Recruits cellular ATPase for efficient budding	82-87
	NS3(174)	Suggested to have a role in host adaptations to mouse	88

na=nucleic acid, aa=amino acid

1.2.3.1 The viral nucleoprotein (NP)

The major component of RNPs, NP protein interacts with viral and cellular proteins to activate replication and transcription of the vRNA. It contains domains of a head, a body and a tail loop, and the basic loop and arginine-rich groove, with a projecting component that contains Arg174 and Arg175 which is expected to interact with vRNA⁸⁹ (Fig. 1.3a). Its monomer interacts with approximately 24 nucleotides of the vRNA⁹⁰. The homo-oligomerization of NP protein is

essential for the structural framework and assembly of RNP particles. The tail loop region of NP protein is important for oligomerization, which could be governed by interaction with the insertion groove, conformation and stability of homo-oligomer⁹¹. In addition, the interaction of NP protein with viral polymerases is part of vRNA synthesis. Such functional interaction occurs at residues R204, W207, and R208 which are located at the top of the head domain⁹². Furthermore, the NP protein promotes nuclear export of vRNP interacting with chromosome region maintenance 1 (*CRM1*) mediated by three novel nuclear export signals (NESs) of the NP-NES1, NES2, and NES3⁹³. NP protein also binds to the middle domains M1 viral protein and mediates vRNPs and M1 protein homo-oligomerization⁷¹.

1.2.3.2 Polymerase basic 1 (PB1)

The 250kD IAV RNA-dependent RNA polymerase is a heterotrimeric complex with three subunit proteins (PA, PB1 and PB2) and has a critical role in viral RNA transcription and replication⁹⁴ (Fig. 1.3b). PB1 protein is the core component of the assembly of the polymerase complex and binds with PB2 and PA proteins. The N-terminal of PB2 protein and the C-terminal of PA protein bound to the C and N terminal of PB1 protein respectively⁵⁴. PB1 protein contains highly conserved motifs that function as a catalytic subunit and transcribes the vRNA using a cap-snatching mechanism of cellular mRNA. The N-terminal region of PB1 protein binds with promoter regions of vRNA, cRNA and PA protein. The amino acids residues at the positions 27 and 44 of the N-terminal are responsible nucleotide recognition and viral replication^{55,56}.

The reading frame-shift of PB1 mRNA encodes for PB1-F2 protein, a recently identified protein with high sequence polymorphism⁹⁵. The avian isolates of IAVs usually express full-length PB1-F2 protein, while its truncated variant is frequent in mammalian isolates. However, the HPAI H5N1 had a higher proportion of truncated PB1-F2 protein⁹⁶. Expression of a higher proportion of truncated PB1-F2 protein is associated with virulence and increased pathogenesis. In IAV-infected cells once PB1-F2 is expressed, it is completely translocated and accumulated in mitochondrial inner membrane space via the translocator of the outer mitochondrial membrane (Tom40) channel⁵⁷. The accumulated PB1-F2 protein assembled into amyloid structures or fragmented nanoscale fibrils⁵⁸. These structures have membrane lytic activity and disrupt mitochondrial membrane due to reduced inner membrane potential^{57,58}.

Figure 1.3: The schematic structure of IAV proteins

a) Nucleoprotein⁹⁷, b) Polymerase complex⁹⁸, c) Hemagglutinin⁹⁹, d) Neuraminidase¹⁰⁰, e) M1 matrix protein¹⁰¹, f) M2 ion channel¹⁰², g) NS1 protein¹⁰³.

Alternatively, the PB1-F2 increases the reactive oxygen species (ROS) by suppressing the superoxide anion dismutase (*SOD1*)⁵⁹. The activation of mitochondrial ROS induces apoptosis, *NLRP3* inflammasome complex and increased inflammations that lead to increased pathogenesis^{59,60}.

1.2.3.3 Polymerase basic 2 (PB2)

The PB2 subunit of IAV polymerase is required for the transcription and possesses a tryptophan-rich, cap-binding sequence that is activated by interaction of the viral polymerase to the 5' end vRNA. Such binding subsequently activates the endonuclease cleavage of PB1 subunit⁵⁰. In addition to vRNA, PB2 protein also binds to PB1, PA and NP proteins, and the binding domain of PB2 protein with NP protein acted as transdominant inhibitors of viral gene

expression^{104,105}. The x-ray crystallography analysis of the C-terminal RNA-binding domain of PB2 protein incorporates lysine 627 residues which is related to IAV host adaptation and virulence¹⁰⁶. This domain is also essential for the accumulation of the cRNA replicative intermediate in IAV infected cells⁵¹. The PB2 protein has a functional site called "VRG" sequence and is crucial for RNA polymerase activity, viral replication and acetyl-CoA interaction⁵². PB2 protein is imported into mitochondrial matrix and was proposed to have a role in apoptosis and innate immune responses⁵³.

1.2.3.4 Polymerase acidic (PA)

As part of viral polymerase, PA protein is required for viral replication and transcription through the assembly of the functional complex⁶¹. In addition, the N-terminus of PA protein possesses endonuclease active site that cleaves capped mRNA and has crucial roles in its endonuclease activity preference to PB1 protein^{94,107,108}. The recently identified PA-X protein encoded by segment 3 of IAV genome via ribosomal frame-shifting is comprised of the endonuclease domain of PA protein with its C-terminal encoded by the second open reading frame. The PA-X protein occurs in two forms, the full length and truncated (missing 20 aa at C-terminus) and the truncated PA-X is naturally encoded by several IAV isolates¹⁰⁹.

PA-X protein has endonucleolytic activity similar to the PA protein, and this endonucleolytic activity is associated with increased viral replication and apoptosis, however, the truncated PA-X variant has reduced activity of endonucleolytic activity⁶². This suggests that the C-terminal 20 aa of the PA-X protein has a critical role in virus replication and virulence⁶². It is localized in both the nucleus and cytoplasm and its endonuclease activity targets and degrades cellular mRNA both in the nucleus and cytoplasm, playing a critical role in host shutoff activity via mRNA decay⁶³. In H5N1 virus infection of mice, chicken and duck PA-X proteins diminished the virulence of the virus through damping the polymerase activity, IAV replication, and innate immune responses (ROS, complement components, apoptosis and inflammations)^{64,65}. Further, two proteins PA-N155 and PA-N182 with suggested role in the replication cycle of IAV have been found in the N-terminal truncated form of PA mRNA¹¹⁰.

1.2.3.5 Hemagglutinin (HA)

Hemagglutinin (HA) protein is the most abundant viral surface glycoprotein, mainly involved in binding to host cell receptor sialic acid (SA) and fusion between viral envelope and

endosomal membrane to facilitate the release of the viral genome in to the nucleus of infected cells⁶⁶. It has the major antigenic property that it is recognized by the host immune system for production of neutralizing antibodies. There are at least 18 different HA antigens and are used for subtyping of IAV strains along with the viral surface NA protein¹¹¹. HA protein is a homotrimeric type I membrane glycoprotein with a membrane anchor domain near C-terminus, glycosylated at side chain ectodomain and acylation of cysteine residues in the cytoplasmic tail region (Fig. 1.3c). It has a central α helix coil and three spherical heads that bind to cellular receptor SA. The binding characteristics of HA protein to the cell receptors is one of the contributing factors for cross-species transmission and emergence of pandemic strains. The monomers of HA protein are synthesized as precursors (HAO) and cleaved with cellular proteases into HA1 and HA2 subunits. Ultimately, the fusion peptide (FP) at the N-terminus of the HA2 will be formed which triggers membrane fusion in the acidic environment⁶⁷. The protease proteolytic and membrane fusion property of HA is also a major determinant for IAV pathogenicity. In HPAI H5 and H7 viruses, the junction of HA1 and HA2 contains multi-basic cleavage sites (R-X-R/K-R) for furin and type II transmembrane serine proteases. Such property enhances the systemic spread of the virus throughout the body causing severe illnesses in human and birds. On the other hand, strains with monobasic cleavage of HA protein by only extracellular trypsin-like proteases restricted to respiratory tissues. Several other extracellular proteases including tryptase Clara, plasminogen (PLG), Human Airway Trypsin-like protease were reported in the proteolytic activity of the HA protein (reviewed in^{67,112}).

1.2.3.6 Neuraminidase (NA)

The NA glycoprotein is a tetramer with a box-shaped head attached to a slender stalk (Fig. 1.3d). Each monomer of the head is composed of topologically identical six β -sheets¹¹³. The NA proteins are able to cleave the glycosidic bond of SA on the host cell surface, and respiratory mucus promoting the release and spread of progeny virus^{68,69}. However, it was recently reported that in a background of depleted concentration of HA protein, NA protein was showed to have a role in IAV receptor-binding activity to complement the process of HA receptor binding⁷⁰. Furthermore, this is relevant for helping the virus approach the target cells by cleavage of SAs from respiratory tract mucins, taking part in the fusion of viral and cell membranes, and increasing HA protein activity by removing neuraminic acid residues nearby the SA (HA receptor binding sites)(reviewed in^{114,115}).

1.2.3.7 Matrix (M1) and the ion channel (M2) proteins

The IAV genome RNA segment 7 has two open reading frames that code for M1 and M2 proteins¹¹⁶. Both M1 and M2 proteins are the main determinants for the morphology of IAV. The M1 protein is a multifunctional dimer protein which forms a layer in-between the viral envelope and integral membrane proteins and RNPs⁷¹. It has a crucial role in intracellular trafficking of RNPs and viral assembly, through its binding activity with the lipid membrane, vRNA, NP and M2 proteins and self-oligomerization. The positively charged region of its N-terminal surface is positioned to RNA binding while its opposite hydrophobic surface interacts with the membrane¹¹⁷. The middle domain of M1 protein is mainly involved in binding NP protein to induce vRNP nuclear export and self-oligomerization^{71,72}. The oligomerization of M1 protein is pH dependent in which the N-terminal domain forms multiple order oligomers at neutral pH and completely dissociates at acidic pH, while the C-terminal domain remains dimeric¹¹⁸ (Fig. 1.3e). The interaction of M1 protein with the cytoplasmic tail of M2 protein is also essential for efficient viral assembly⁷³.

The M2 protein is a homotetrameric, highly conserved integral membrane protein of IAV which contains two disulfide-linked dimers tied by noncovalent interactions⁷⁷. It contains an ectodomain (N-terminal), a transmembrane domain, and cytoplasmic tail (C-terminal). The N-terminal end is exposed at the cell surface, while the transmembrane domain serves as membrane anchor and signal sequence during its synthesis¹¹⁹ (Fig. 1.3f). Four identical transmarine domain of the protein forms the M2 protein pore that serves as proton-selective ion channel. The flux of proton ions across the viral membrane occurs through M2 protein pore, leading to acidification and uncoating of the virus for its replication. The IAV drug amantadine binds to the M2 protein pore and prevents acidification and uncoating of the virus⁷⁴. In addition to its uncoating role, the M2 protein cytoplasmic tail has also a role in virus assembly and budding of the infectious virus by interacting with M1 protein and mediating cholesterol-dependent alteration respectively^{73,75,76}. Besides, M2 protein promotes the maturation of HA protein by regulating the pH of trans-Golgi network⁷⁷.

1.2.3.8 Non-structural proteins (NS1 and NS2)

The NS1 of IAV is a protein with several functions. It has two unique domains, an RNA-binding domain (RBD) and an effector domain (ED) joined by a linker region (LR). The LR is flexible governing the NS1 ED domain to interact with specific proteins and RNA molecules

allowing the multifunctional property of the NS1¹²⁰ (Fig. 1.3g). The interaction of NS1 protein with dsRNA prevents the degradation of vRNA by 2'-5' oligoadenylate synthetase (OAS)/RNase. In addition, such interaction inhibits dsRNA-dependent signaling pathways that induce the production of the anti-viral interferon (IFN) response. Several cellular proteins involved in the antiviral response, cellular mRNA processing, host gene expression or gene silencing, signaling pathways, cell cycle and host cell physiology were reported to interact with the NS1 protein¹²¹⁻¹²⁴. Furthermore, it has a role in modulating the replication, protein synthesis, transmissibility and pathogenicity of the virus⁷⁸⁻⁸¹.

The IAV NS2 protein which was renamed as nuclear export protein (NEP) is an adapter that mediates the nuclear export of vRNPs by interacting with the nuclear export complex of the cell and the viral M1 protein⁸². Its N-terminal helix domain contains leucine-rich nuclear export signal (NES2)⁸³ and binds to the nuclear export receptor CRM1 in RanGTP-dependent manner and "GLFG" repeat domain of human nucleoporin 98 (hNup98)⁸⁴. The C-terminal domain which contains antiparallel amphipathic helical hairpin binds to the viral NLS of M1 protein⁸⁵. Recently, it was reported that NEP protein also has a role in transcription and translational regulation of genomic vRNA and recruits a cellular ATPase to the cell membrane for efficient budding of virions^{86,87}.

1.3 IAV life cycle

The life cycle of IAV involves mainly three steps: entry to the host cell, production of viral components (transcription, translation and replication), and assembly and budding of the newly produced virion (Fig. 1.4)⁹⁸.

Nature Reviews | Microbiology

Figure 1.4: Schematic representation IAV lifecycle (adapted from Te Velthuis et al., 2016; Nature Reviews Microbiology, copy right number 4140240593811)⁹⁸.

IAV enter the cell through receptor mediated endocytosis. Once the virus is internalized, it is transported to the nuclear area while maturing in endosomes (lowering pH). Fusion of the endosomal and viral membrane in acidic environment induces the release of vRNPs which in turn transported to nucleus for replication and transcription. The vRNP give rise to mRNA and cRNP. The transcribed mRNA will be transported to cytoplasm for translations, while the cRNP will be used as a template to produce new vRNPs. The new vRNPs and viral protein products will be assembled in the cytoplasm for the release of progeny viruses.

1.3.1 Virus entry

The first step of virus entry process is recognition and attachment of the viral HA with the primary receptor SA. SA is the component of oligosaccharide chains of *N*- and *O*-linked glycoproteins and lipids. Usually, it is bound to galactose by $\alpha(2, 3)$ - or $\alpha(2, 6)$ -linkages. It was well established that the avian and human strains of IAV have a preference to interact with SA receptors with $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA respectively. This receptor specificity of IAV HA and SA interaction was determined by several factors. The primary factor is the

conformation and structure of HA receptor binding site (RBS), in which the RBS of human and pig viruses are “wider” than the RBS in avian isolates¹²⁵. Other factors like modifications on the underlying sugar chains, the topological structure, alternative glycosylations, and the orientation of SA were also reported to have a role in determining the specificity (reviewed in¹²⁶). In addition to SA, other potential receptors of IAV like C-type lectins, Annexin V and 6-sulfo sialyl Lewis X were also proposed (reviewed in¹²⁶).

Once the virus is attached to its receptor, the next step is the internalization of the virus through receptor-mediated endocytosis¹²⁷. Clathrin-mediated endocytosis is the primary internalization pathway of many viruses including IAVs. The binding of the virus to the receptor induces the formation of clathrin-coated pits (CCPs) which involves cargo molecules¹²⁷. Specific internalization signals on cargo receptors are recognized by adaptor proteins. For example, a cargo-specific adaptor Epsin 1 was recruited simultaneously with the assembly of CCPs to the binding site of IAV¹²⁸. The Ras-phosphoinositide 3-kinase (*PI3K*) and the focal adhesion kinase (*FAK*) are also involved the internalization of IAVs¹²⁹. Other endocytic pathways like caveolin-dependent endocytosis¹³⁰, clathrin- and caveolin-independent endocytosis¹³¹ and macropinocytosis¹³² were also used in IAV internalization.

The internalization of the virus forms endocytic vesicles which in turn is followed by formation of early and late endosomes approximately after 10 and 40 min respectively¹³³. The internalized virus particle is transported to the early endosome in an actin-dependent manner. The endocytic compartment that contains the virus leaves the early endosome bearing endocytic carrier vehicle (ECV) which is rapidly transported to the perinuclear region via dynein-directed transport on microtubules. The intermittent activity of the plus- and minus-end directed active transport of the membrane-bound motor protein on microtubule matures the ECV into maturing endosome. The endosome further matures into late endosome by changing the pH to 5, which then activate the fusion of viral envelope with endosomal membrane for the effective release of the virus genome to the perinuclear region¹³⁴.

The acidification of the endosome is an important step for the release of IAV genome. The lower pH (pH < 6) from the endosome transferred to the interior of the virus through the ion channel protein M2 and this condition triggers major and irreversible conformational changes in HA and M1 proteins. The conformational change of M1 protein resulted in its dissociation from RNPs leaving the envelope more flexible for the formation of compatible fusion^{135,136}. The

conformational changes in HA protein lead to the exposure of the fusion peptide that will anchor to the endosomal membrane. Further conformational changes resulted in the subsequent formation of hemifusion, stalk and fusion pore. Then the genomic content that contains vRNPs released into the cytosol through fusion pore (reviewed in¹³⁷). This process also involves aggresome formation which involves histone deacetylase 6 (*HDAC6*), dynein, dynactin and myosin II¹³⁸.

Unlike other RNA viruses, IAV replicates in the nucleus, hence the released vRNPs from the endosome have to be imported to the nucleus for effective replication of the virus. The import of vRNPs involves interaction with cellular nuclear transport receptors particularly importin- β (karyopherin- β) superfamily or importin α . These nuclear receptors bind to the nuclear localization signals (NLSs) of the vRNPs. It was reported that the dissociation of M1 protein from RNPs exposes at least two putative NLSs; at the N-terminal (NLS1) and in the middle (NLSs) on NP protein. The NLS1 is the main mediator of nuclear import of vRNPs^{139,140}. Recently it was demonstrated that the nuclear import of vRNPs is facilitated by heat shock protein 40 (*HSP40*), while suppressed by MxA and Moloney leukaemia virus 10 (*MOV10*) proteins^{141,142}.

1.3.2 Transcription and replication

After entry to the host cell nucleus, the shuttled vRNPs have to be transcribed and replicated to proceed the life cycle of the virus. The polymerase associated with vRNP is responsible for both transcriptions of vRNAs into viral mRNA and replication of the parental vRNA into progeny vRNA and cRNA. Prior to transcription, the PB2 subunit of the polymerase protein binds to the 5'-cap structure of cellular pre-mRNA followed by the generation of 5'capped RNA primers with 3'-hydroxyl group by the endonuclease activity of the PA subunit of the polymerase protein. The addition of guanine (G) residue at the 3' end of the capped primer directed by second last cytosine (C) residue at the 3' end of the vRNA template initiate the transcription^{98,143}. Consequently, the catalytic domain of the polymerase, PB1 subunit extends these primers through condensation of nucleoside triphosphates. The extension proceeds until the recognition of a 5-7 uracil (U) residues sequence located at the 5' end of vRNA template that act as a polyadenylation signal. As the polymerase reaches the site, the steric hindrance caused by 5' end restricts the elongation resulting in poly (A) tail. Then the 5' capped viral mRNA is released from PB2 subunit and bound to nuclear cap binding complex (CBC). This phenomenon activates

the recruitment of cellular factors like TRanscription-EXport (TREX) complex and Nuclear RNA export factor 1(*NXFI/TAP*) making viral mRNPs to be exported to the cytoplasm. Before the viral mRNPs were exported, the association of viral polymerase and cellular RNA polymerase II (Pol II) triggered the takeover of the host splicing machinery by the virus for its alternative splicing of some viral mRNA's (e.g. M, NS mRNAs^{98,143,144}). Further, the virus uses the host cell machinery for translation and viral proteins synthesis in the cytosol¹⁴⁵.

The replication of IAV genome involves mainly two steps, primarily positive sense full-length cRNA is synthesized from the parent vRNA by viral polymerase, and the progeny cRNA is assembled with NP protein encapsidation and the polymerase along with its synthesis to form cRNPs. The cRNP is then used as a template for the synthesis of many progeny vRNPs. The viral transcription is distinct from genome replication in that, transcription involves 5'-cap-snatching mechanism of the host pre-mRNA and 3' polyadenylation⁹⁸. It was also reported that transcription and replication carried out by *cis*-acting and *trans*-acting RNA polymerase respectively^{146,147}. However, recently a new model was proposed that a *trans*-activating polymerase activates vRNA synthesis, while a resident polymerase complex is responsible for replication¹⁴⁸. The newly synthesised vRNPs are exported to the cytoplasm by the nuclear export machinery of the host cell, Crm1 pathway mediated by the viral NEP¹⁴⁹.

1.3.3 Viral assembly and budding

Following the replication of genomic vRNPs and synthesis of viral proteins, efficient assembly and budding are indispensable for the production of infectious virus. IAV assembly and budding occurs in the viral "budozone", a lipid raft domains of the apical cell membrane with characteristics of high cholesterol/sphingolipid concentration. The viral proteins HA, NA, M1 and M2 are membrane-associated proteins which play important role in the process of viral assembly and budding. The forward transport of HA and NA protein from ER to Golgi and to plasma membrane is accompanied by an increased cholesterol content targeted to lipid raft domain membrane formation. The high concentration of HA and NA proteins are important in initiating the budding process through altering the local membrane curvature. The cytoplasmic tails of HA and NA proteins recruit M1 protein and enable its polymerization enhancing further alteration of membrane curvature and budding. However, the ability of M1 protein in recruiting vRNPs and M2 protein at the site of budding impedes the budding process until proper assembly

occurs. At the budding neck, the amphipathic helix domain of M2 protein is inserted into the membrane and exerts positive membrane curvature adjusting the line tension between the membrane lipid phases. Further changes in membrane curvature provide the force that mediates membrane scission and budding of the virus. At this stage, although membrane scission is complete, the virus is still associated with cell membrane due to the interaction of HA proteins and membrane SAs. Subsequently, NA enhances the final release of free progeny virus by cleaving the SA bound to HA (reviewed in^{150,151}).

1.4 IAV host adaptation and virulence

Cross-species transmission and efficient replication of IAV have been documented in different host species (e.g. from avian to mammalian host [human, pigs, ferrets, and mouse, and from human to pigs, ferrets and mouse]) (reviewed in¹⁵²⁻¹⁵⁴). Adaptation to new hosts resulted in the acquisition of phenotypic and genotypic changes. These changes help the virus to cope with the challenges in the new host at different stages of its life cycle (receptor binding specificity, virus enter and uncoating, nuclear transport, replication and release of the virion).

Several genetic changes of the virus (mutations) during host adaptations have been documented in several of the viral proteins (Table 1.2). The host adaptation changes in HA proteins are mainly important for switching the receptor binding specificities. The distribution of IAV receptor is variable within and between different host species. In humans, the upper respiratory tract (URT) is lined with $\alpha(2, 6)$ -linked SA, while the lower respiratory tract (LRT) (alveoli and bronchial epithelium) have higher % of $\alpha(2, 3)$ -linked SA. However, the abundance of $\alpha(2, 3)$ -linked SA in both URT and LRT is higher in children than adults. In addition, other conditions like inflammation increase the expression of $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA through activation of sialyl-transferase (the enzyme that adds SA to galactose). Pigs express both $\alpha(2, 6)$ -linked SA and $\alpha(2, 3)$ -linked SA with a gradual increase in $\alpha(2, 3)$ -linked SA towards LRT. Ferrets have a similar distribution of $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA all through the respiratory tract. The respiratory and intestinal tract of avian species have both $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA, but $\alpha(2, 3)$ -linked SA more abundant (reviewed in¹⁵³).

Table 1.2: IAV viral protein mutations associated with host adaptation and virulence

Mutations				
Viral proteins	Subtypes	Adaptation	Virulence	References
HA	H1N1	E190D, D225G	loss of glycosylation at (71, 142, 144, 147, 172, 286)	154,155
	H2N2	Q226L, G228S		152,153
	H3N2	Q226L, G228S	loss of glycosylation at (165, 246), S145N	152,153,156,157
		Q226L,G228S, N224K, N158D, T160A, T318I, H110Y, S221P, K216E, K153D, S223N/I, G272S, S277N	MBCS, loss of glycosylation at (158, 169), K222Q, S227R, S133A, P235S	153,158-163
	H5N1		L320P, S203F, E273G	164,165
	H5N2		H156N, S263R	166
	H6N6		MBCS, Q450L, E114K, G205E, G218E	167-169
	H7N1		MBCS	170
	H7N3			
		Q226L, S138A, G186V, T221P, A107T, R220I, L226Q, R354K	A107T, R220I, L226Q, R354K	154,171
PB2	H1N1		504V	172
	H3N2	D701N, R702K, S714R, E627K, D701N, R702K, S714R		173
	H5N1		E627K, D701N, R702K, S714R, V25A	168,173-176
	H5N2		Q591K, D701N	165
	H6N6, H7N1		E627K,E627K	166,167
	H7N9		V598T/I	177
PA	Several IAVs	K142E	K142E	172
	H1N1	P224S, 550L	E351K,P224S	172
	H3N2	Q556R	Q556R	172
		K615N, R443K, I353R, T552S	K615N, R443K, I353R, T552S	172,176
	H5N1	T97I	T97I	172
	H5N2		T97I, E133G,	172
	H6N5		I38M	166
	H6N6		S65Y, K615N	172
	H7N7		T97I, K328R, P332T, Q556R	171
	H7N9			
PA-X	H1N1, H5N1		Anti-virulence of PA-X	178
	H9N2		Pro-virulence	178
	H7N7		V63I, A37S/I61T/V63I/V100A(Mfour)	179
NA	H7N1		S350N	167
	H7N9		M26I, N142S, G389D	171
PB1	H1N1		E398G, S524G, I563R	180
PB1-F2	H1N1		S66N	168
	H5N1		S66N	168
NEP/NS2	H5N1	M16I, Y41C, E75G		181
NP	H7N9		A284T, M352I	171
M1	H7N9		M128R	171
NS1			PDZ domain binding motifs (From RSEV to KSEV/ESEV/EPEV0, G45R	182,183
	rH1N1		I64T, D189N, V194I (reduce virulence)	184,185
	H3N2		92E	168
	H5N1		R293K	161
	rH5N3		I106M	186
	H7N9			

The avian viruses usually bind to $\alpha(2, 3)$ -linked SA, while the human isolates often bind to $\alpha(2, 6)$ -linked SA. While the avian viruses adapted to mammalian host the HA protein mutations (Table 1.2) switches the binding specificity from avian $\alpha(2, 3)$ -linked SA to human $\alpha(2, 6)$ -linked SA (H1N1, H2N2, H3N2, H5N1) or dual receptor binding (both $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA) in H7N9 (Table 1.2). The balance between HA and NA proteins is also important for efficient virus attachment, entry and releases¹⁵³. The viral and cellular membrane fusion is an important step in the lifecycle of IAVs and stable fusogenic conformation of HA protein is triggered by either increased temperature or lower pH. Generally, the pH for stable fusogenic conformation of HA protein in human/mammals is lower (33 °C) than the avian IAV isolates (42 °C). Given the pH and thermostability differences, human-adapted avian viruses acquire HA protein stability adaptation through changes in histidine residues (e.g. H110Y in H5N1 virus)¹⁸⁷.

The other adaptive mutations in IAV polymerase proteins were associated with increased polymerase activity, virus replication, increased nuclear import and virulence. The PB2 protein (e.g. E627K, D701N mutations in H5N1 virus) increased the polymerase activity and virus replications in mammalian hosts, while the other mutations (D701N, R702K and S714R) in H5N1 and H3N2 viruses were associated with the enhanced association of vRNPs with importins¹⁷³. Similarly, the mutations in PA proteins were associated with host adaptations and increased virulence¹⁷². In addition to viral polymerase proteins, the NEP (M16I, Y41C, E75G) mutations in the H5N1 virus were associated with increased polymerase activity in mammalian host¹⁸¹. Collectively, HA and polymerase proteins play major role in IAV host adaptation processes.

Increased virulence of IAV infections was associated with several host and viral factors. IAVs that infect the lower respiratory airways are usually associated with increased severity and pneumonia. In addition, viruses that were resistant to IAV neutralizing respiratory secretions e.g. pentraxin 3 (*PTX3*) and surfactant proteins had severe outcomes^{156,188}. Other virulence factors include 1) systemic spread of the virus due to insertion of multibasic cleavage sites (MBCS) in HA (H5 and H7 viruses)^{161,168,170}, and specific mutations in other proteins (several of mutations associated with IAV virulence are described in Table 1.2), 2) removal of HA glycosylation sites (residue 16, 21, 24 of HA protein), 3) increased polymerase activity, virus replication and

inflammatory responses¹⁷², 4) inhibition of antiviral responses (by PB1, PA-X , and NS1 proteins) (reviewed in¹⁸²).

The NS1 protein antagonizes antiviral responses through 1) suppression of IFN- β response by interfering with transcription factors (*IRF3*, *NF-kB*, *AP-1*), *RIG-I* signaling and *CPSF30*, 2) interfere with *OAS*/RNase L to protect viral RNA degradation, 3) activation of *PI3K/Akt* pathway and 4) interfering with pro-apoptotic proteins with its PDZ domain-binding motifs (ESEV-avian, RSKV-human motifs) and 5) hijacking host proteins^{182,189}. The activity of NA protein is also reported to have effect in the virulence of IAVs. The H1N1/WSN and 1918 H1N1 viruses replicated in the absence of trypsin, as NA protein recruits plasminogen that converts to plasmin for HA protein cleavage. The H7N7 virus had increased NA protein activity preventing the formation of virus aggregates and enhanced virus releases. Deletion of the NA protein stalk region (shortened NA stalk) are highly transmissible between avian species and was associated with increased virulence¹⁵².

1.5 Host responses to IAV infections

Innate and adaptive immune responses are the main host responses during IAV infections. The innate immune responses are non-specific first line defenses and involves soluble mediators of the respiratory tract and oral cavity (mucus (mucins), surfactant proteins, pentraxins, ficolins, galectins, mannose-binding lectins (MBL), antimicrobial peptides and complements), chemical mediators (pattern recognition receptors (PPRs) and their downstream signaling, and hematopoietic cells (Macrophages, Natural killer cells, neutrophils, Dendritic cells etc). Adaptive immune responses during IAV infection includes the humoral immune response which is characterized by the production of IAV specific antibodies (immunoglobulin A (IgA), IgM, and IgG) by B-lymphocytes (plasma cells) and cellular responses that involve IAV specific CD4+ and CD8+ T lymphocytes (reviewed in^{190,191}).

1.5.1 Innate immune response

1.5.1.1 Soluble mediators

The soluble mediators of innate immune response are a complex array of IAV inhibitors found in fluids lining the respiratory tract. Most of the inhibitors have SA ligands for trapping and inhibiting binding of the virus particles to epithelial cells, while others interact to activate complement cascades, and modulation of IFN responses (reviewed in¹⁹⁰).

1.5.1.1.1 Mucus (mucins)

Respiratory mucus hydrates the respiratory tract and has a protective barrier role by trapping infectious particles. Mucus is also composed of a higher proportion of mucins which is characterized by high glycosylation and SA content. Mucins inhibit IAV entry by trapping and mimicking the sialylated receptors on the surface of the target cell. In addition, it competitively inhibits the cleavage activity of NA protein. The inhibition of IAV by mucins depends on the activity of viral NA protein, and the virus evades the inhibition by increased NA activity¹⁹²⁻¹⁹⁴.

1.5.1.1.2 Surfactant proteins

Surfactant protein (SP)-D and SP-A are collectins (C-type lectins) that inhibit the HA activity of several IAV subtypes. SP-D is 10 times more potent inhibitor of HA protein activity than SP-A. Porcine SP-D is exclusively sialylated with $\alpha(2, 6)$ -linked SA, while SP-A contains both $\alpha(2, 6)$ -linked and $\alpha(2, 3)$ -linked SA. SP-D has trimetric neck and carbohydrate-recognition domain (NCRD) that binds to the globular domain (HA1) of IAV and inhibits HA protein activity and infectivity. In addition to inhibition of HA activity, both SP-D and SP-A enhance IAV aggregation and opsonization (uptake by phagocytic cells), neutrophil respiratory burst, and modulation of innate immune response¹⁹⁵⁻¹⁹⁷.

Decreased or deficient SP-D was associated with increased susceptibility to avian and seasonal IAV infections, and increased inflammatory cytokines and virulence in human and mice^{190,198,199}. Suppression of IAV by SP-D determined by the existence of mannose-rich glycans on the viral HA protein. Seasonal IAV strains have adequate glycans on HA and are susceptible to SP-D inhibition, while IAV isolates that lack HA protein glycosylation (e.g. pandemic viruses, PR-8 and WSN) are resistant to SP-D inhibition^{200,201}. Recently, Nikolaidis et al., 2014 showed that substitution of D325A + R343V and D325S + R343Y of SP-D inhibited pandemic strains suggesting its value as a potential therapeutic agent²⁰².

1.5.1.1.3 Mannose Binding Lectins

Mannose-binding lectins (MBL) are key components of innate immune response during IAV infections. Lack of MBL is linked to increased IAV infectivity and severity in mice²⁰³. Several mechanisms have been described for the antiviral activity of MBL which includes 1) inhibition of HA and NA protein activities^{203,204}, 2) increasing viral aggregations²⁰⁴, 3) activation of lectin complement pathway and coagulation²⁰⁴ and 4) reduction of inflammation (TNF),

oxidative injury (ROS), and apoptosis in alveolar macrophages²⁰⁵. Similar to SP-D, seasonal IAVs and H9N2 viruses were susceptible to MBL, while pandemic strains and H7N9 viruses were resistant^{206,207}.

1.5.1.1.4 Ficolins

Ficolin is one of the pathogen-associated molecular pattern recognition molecules in animals. It has collagen-like and fibrinogen-like globular domains, for specific binding of carbohydrates specifically N-acetylglucosamine. Three ficolins have been identified in human, L-ficolin, H-ficolin, and M-ficolin. The L-ficolin and H-ficolin have been demonstrated to recognize and bind with carbohydrates of HA and NA proteins of different IAV subtypes. These ficolins lower HA protein titer, lung pathology, and induces production of *IL-4*²⁰⁸. In addition, it forms a complex with mannose-binding-lectin associated serine proteases (*MASPS*) for complement activation. This in turn mediates opsonization, phagocytosis and cytokine productions²⁰⁹. In contrast to SP-D and MBL, H-ficolin inhibits the less HA protein glycosylated IAV strains including PR8 H1N1 and the 2009 pandemic H1N1 viruses²¹⁰.

1.5.1.1.5 Galectins

Galectins are a protein family that specifically bind to β -galactoside sugars (e.g. N-acetylglucosamine) which become bound to proteins in N- or O-linked glycosylation. There are 15 galectins that have been identified in mammals and galectin 1, 3 and 9 were indicated to have anti-IAV activity²¹¹⁻²¹³. These galectins are highly expressed during IAV infections in mouse and human, and galectin 1 and 3 are released to bronchoalveolar space²¹⁴⁻²¹⁶. Galectin 1, 3 and 9 play anti-IAV virus role through inhibition of HA protein activity, reducing viral replication, reduction of inflammatory cytokines and apoptosis^{211,212,214}. On the other hand, deficiency of galectin 9 showed increased virus-specific CD8+ T cells and rapid humoral immune response (IgM, IgG and IgG) followed by rapid IAV clearance²¹³.

1.5.1.1.6 Antimicrobial peptides

Defensins and cathelicidins (LL-37) are the two main classes of antimicrobial peptides (AMPs) present in the human respiratory tract. There are two types of defensins, the α - defensins which are packed in neutrophil granules, and the β -defensins which are primarily produced by epithelial cells lining the respiratory tract (reviewed in²¹⁷). The human defensins have been showed to have a role in neutralizing many IAV strains, however, the mechanism is not known

yet. The human α - defensins induce IAV aggregations by direct binding and formation of multimolecular assemblies of defensins. Retrocyclins are also another class of defensins with similar functions as α -defensins, but retrocyclins are expressed in primates. β -defensins were shown to have an immunomodulatory role during IAV infections. The LL-37 is also another AMP that disrupts IAV membrane inside the cell (reviewed in^{190,217}).

1.5.1.2 Pattern Recognition Receptors and downstream signalling

Once IAV is internalized, the cell utilizes three types of pattern recognition receptors (PRRs) to recognize the virus in the cell to initiate immune signalling. The three PRRs are the retinoic acid-inducible gene I (RIG-I)-like receptors, the Toll-like receptor (TLR) and nucleotide binding oligomerization domain (NOD)-like receptor signalling. RIG-I can be activated by a short 5' ppp-dsRNA structure (panhandle) which is formed by the complementary sequences of the 5' and 3' ends of IAV viral RNA. In addition, RIG-I can also be activated by the U/A-rich ssRNA sequences in the 3' untranslated region (UTR) of IAV genome. Activation of RIG-I induces conformational changes and allows to interaction with mitochondrial-bound adaptor molecule *MAVS*. The process subsequently activates the interferon regulatory factor 3 (*IRF3*) (for IFN and IFN-stimulated genes (ISGs) expression) and NF- κ B (mediates inflammasome). In addition, a direct antiviral activity of RIG-I was demonstrated by its interaction of with 5'ppp dsRNA panhandle and dismantling the viral polymerase complex (reviewed in²¹⁸).

TLRs are expressed in innate immune cells including epithelial cells of the respiratory tract and recognizes a wide range of pathogens including IAVs. There are about 11 TLRs identified in mammalian cells, of which *TLR3*, 7, 8 and 9 are located in endosomal membrane and recognizes nucleic acids, while the others are located on the plasma membrane (reviewed in²¹⁹). In pH1N1 and H3N2 virus infections, *TLR3*, 7, 8 and 9 were up-regulated while *TLR2* and 4 were suppressed²²⁰. During IAV infection, degradation of ribonucleoprotein complex in acidified endosome resulted in the exposure of the double stranded (ds) and ss viral RNA which are recognized by the *TLR3* and *TLR7/8* respectively^{219,221}. Recent studies suggest that viral RNP complex and the glycoproteins (HA and NA) are recognized by *TLR10* and *TLR2/4* respectively^{219,222}. The downstream activation of *IRF3* and *NF κ B* by *TLR3* is mediated by adaptor protein *TRIF/AP1*, while *TLR7* is mediated by *MyD88*²¹⁹.

The NOD-like receptor family, pyrin domain containing 3 (*NLRP3*) have been shown to have an important role in innate and adaptive immune response, however, excessive

inflammasome activation was associated with a severe form of disease. The *NLRP3* inflammasome stimulates the activation of caspase-1, and its subsequent cleavage promotes secretion of pro-inflammatory cytokines (*IL-1 β* and *IL-8*). The mechanism is not known, however recently it was suggested that the viral RNA facilitates inflammasome assembly, and the interaction with other PRRs (RIG-I and TLRs) releases inflammasome-mediated cytokines^{223,224}.

Activation of all the three PRRs results in the production of cytokines (e.g. IFN type I, II and III, *TNF α* , *IL6*, *IL1 β*) and chemokines (e.g. *CCL2*, *CCL4*, *CCL5*, *CXCL10*, *IL8*) for induction of antiviral response, activation of immune cells and adaptive immune responses. Type I IFNs (IFN α/β) and type III IFNs (*IL28A*, *IL28B*, *IL29*) are very important for protective acute IAV infections through activation of an antiviral response in infected and neighbouring cells. Type I IFNs are mainly expressed by alveolar macrophages and dendritic cells (DCs), and their receptors are expressed at the surface of all cells. On the other hand, type III IFNs are primarily expressed by airway epithelial cells and their receptors are restricted to airway and gastrointestinal tracts. Both type I and III IFN induce expression of hundreds of ISGs through *JAK/STAT* signalling pathway (reviewed in^{225,226}). For instance, the human myxovirus resistance protein 1 (MxA) or *Mx1* (in mouse) is one of the ISGs that have antiviral (it blocks the nuclear import of RNPs and export of viral mRNAs) and anti-inflammatory effects during IAV infections. Interferon-induced transmembrane protein 3 (*IFITM3*) is another well-characterized ISG that inhibited early replication of IAVs.

Type II IFNs (IFN γ) are mainly produced by NK cells and T lymphocytes. They promote macrophages and neutrophils for the enhanced killing of virus-infected cells through ROS and lysosomal enzymes. It also activates *JAK/STAT* signalling pathway for expression of IFN γ -related genes (e.g. *IRF1*) that leads to IFN β production. In addition, IFN γ is involved in activation of antigen presentation (macrophages), T helper 1 (Th1) response, and regulation of B cells (reviewed in^{190,225}). The antiviral activity of *TNF α* , *IL6* and *IL1 β* have been described through the production of acute phase proteins, ROS, mediating inflammatory responses, tissue repair, and activation of adaptive immune responses. However, several pieces of evidence showed excess expression of the IFNs and these cytokines resulted in lung tissue damage and severe diseases during IAV infections^{152,211}.

In addition to cytokines, the innate immune response in IAV infections involves chemokines which recruit immune cells to the site of virus infection for rapid clearance. For example, during IAV infections *CCL5* and *CXCL10* recruit macrophages, DCs, NK cells and T-lymphocytes. Basophils, eosinophils, and monocytes were attracted by *CCL4* and *CCL5*, while monocytes, T-lymphocytes and NK cells are also recruited by *CCL2* and *CCL4*. *IL8* highly attracts neutrophil in IAV infected mice²²⁵.

Once these immune cells arrived at the site of IAV infection, they play their own antiviral roles. Macrophages limit virus spread by opsonophagocytosis of IAV particles, phagocytosis of IAV infected cells, and regulation of adaptive immune responses. NK cells lyse IAV infected cells through direct recognition of HA protein and by antibody-dependent cell cytotoxicity (ADCC). Neutrophils also uptake IAV infected cells, produce defensins and LL-37 and promote T-cell response²²⁵.

1.5.2 Adaptive immune response to IAV infections

The humoral and cellular adaptive immune responses are critical host responses against IAV infections and are linked to the innate immune responses by DCs. During IAV infections, DCs migrate to the secondary lymphoid organs to present IAV antigens to B cells by direct interaction. Then the activated B cells will be transformed into antibody-producing plasma cells. In IAV primary infections HA-specific antibodies (IgA, IgG and IgM) were identified in nasal secretions. The IgM antibody is more dominant in primary infections, while IgG and IgA are dominant in subsequent infections. The HA protein-specific antibodies bind to the globular head or stem region of HA protein, neutralize the virus and inhibit the virus attachment to the cell. In addition, the antibodies facilitate phagocytosis and ADCC of virus-infected cells. The NA protein-specific antibodies inhibit its enzymatic activity and facilitate ADCC, whereas the NP protein-specific antibodies trigger complement mediated lysis of virus-infected cells (reviewed in^{191,227,228}).

In addition, IAV infected DCs degrade the viral proteins into small peptides and present them to CD8+T cells by MHC class I molecules. The activated CD8+ T cells differentiate into cytotoxic T lymphocytes (CTLs). The CTLs produce perforins and granzymes for inducing apoptosis of the virus infected cells. In addition, CTLs promote apoptosis via Fas/FasL interaction. On the other hand, DCs phagocytize IAV infected cells or apoptotic cells, which are subsequently degraded and presented to CD4+ T (helper) cells by MHC class II. There are two

major classes of T helper cells, Th1 and Th2. The activated Th1 cells produce *IFN* γ and *IL*-2, and protect IAV infection through the cytolytic mechanism, whereas the Th2 cells promote antibody responses through affinity maturation and generation of memory plasma cells (reviewed in^{191,227,228}).

1.5.3 IAV escape strategies from the immune response

Despite the plethora of immune responses involved against IAV infections, the virus has several strategies to escape from the host responses. The NS1, PB1-F2, PA-X, and M2 viral proteins have been implicated in innate immune response escape, while the change in the HA and NA proteins escape both innate and adaptive immune responses. For instance, the NS1 protein escapes the innate immune response by several mechanisms, which includes binding to viral RNA to mask recognition by PRRs, host protein synthesis shutdown, suppression of type I IFN through up-regulation of *SOCS3*, and by antagonizing the ISGs directly. PB1-F2 inhibit IFN response by interacting with *MAVS* and IPS-1, M2 induces cell apoptosis, and PA-X suppresses expression of cellular immune responses (reviewed in^{191,229}).

Antigenic drift (point mutation on antigenic site, receptor binding or glycosylation site) and antigenic shift (reassortment of the genomic segments between different IAV subtypes) are the major determinants for evasion of both innate and adaptive immune responses. In addition, NP protein mutation (R384G) was reported to evade from CTLs response (reviewed in^{191,229}).

1.6 Genome-wide host gene expressions in IAV infections

The characteristics of IAV and its host response have been extensively studied *in vitro*, *ex vivo* and *in vivo* using different cell line and animal models (reviewed in^{230,231}). *In vitro* studies were applied to understanding different stages of viral life cycle and the corresponding cellular host response. Studies have involved several cell line cultures like the human bronchial adenocarcinoma epithelial continuous cell line (Calu-3), adenocarcinoma human alveolar basal epithelial cells (A549 cells), HeLa cells, and Madin-Darby canine kidney (MDCK) cells, chicken embryo fibroblast (CEF) cells, interferon-deficient Vero cells²³⁰. The *Ex vivo* studies mainly involve human primary epithelial cells derived from respiratory tract (nasal, tracheal, bronchial and alveolar) tissue and were used to address the properties of tissue tropism, infectivity and immune responses (e.g. primary human bronchial epithelial cells (NHBE)). *In vitro* and *ex vivo* models may not reflect the whole system of the *in vivo*²³⁰. The animal model is

important to investigate the pathogenicity and transmission of the different IAV strains and subtypes. Currently, mice, ferrets, guinea pigs, cotton rat, and non-human primates are the most frequently used animals for IAV studies^{230,231}. However, the difference in the host adaptation, the transmissibility, the pathogenicity, challenges in reagent availability, and genome annotation limits representing the human situation^{216,217}. In addition, the cost and ethical issues were the drawbacks to using the closely related non-human primates for experiments^{230,231}. Nevertheless, all the models give a great and deep insight to our knowledge on the replication, pathogenesis, and transmissibility of IAV. Such understanding enables designing and testing of influenza virus therapies and vaccines which could be applicable to human population²³⁰.

1.6.1 Host gene expressions in seasonal (H1N1 and H3N2) IAV infections

The host response to seasonal IAV (PR8/H1N1) has been investigated by global gene expression profile of various cell types. In human alveolar type II cells, it induced up-regulation of pattern recognition receptors (PRRs), IFN, ISGs, *SOSCs*. Among the chemokines response, the *CXCL9*, *CXCL10*, and *CXCL11* were the most induced proteins²³². Further, PR8/H1N1 virus infection activated *PI3K* signalling and inhibition of this signalling reduce production of chemokine response and infectious viruses²³². In mouse airway epithelium, the early induced expression of *NF-κB* binding site motifs, members of the triggering receptor expressed on myeloid cells 1 (*TREMI*) and interleukin 17 (*IL17*) signalling pathways were associated with restricted replication of the virus²³³.

In human macrophages, PR8/H1N1 virus infections increased the expression level of *SOCS1*, *RIG-I* and *IFN-α*²³⁴. Compared to pH1N1 virus it induced a significantly lower level of inflammatory cytokines (*IL6*, *TNF-α*, *IL10*, *CCL3*, *CCL5*, *CCL4* and *CXCL8*) in human macrophages. Similarly, the seasonal H1N1 virus showed lower expression level of pro-inflammatory cytokines including *IFN-α*, *CCL2*, *TNF-α*, and *CCL5* compared to pH1N1 virus infections in NHBE cells²³⁵. In human macrophages, PR8/H1N1 mediate high production of type I IFNs in mouse derived plasmacytoid dendritic cells (pDC)²³⁶. In C57BL/6J mice lung tissue, PR8/H1N1 virus presented with increased expression of IFNs and early innate immune response activating genes, which subsequently increased T and B cell specific gene expression along with infiltration and activation of T and B lymphocytes²³⁷. Lung tissue multigene transcriptional profile of mice infected with the highly pathogenic PR8/H1N1 virus revealed elevated activation of inflammatory signalling and influx of neutrophils and was associated with lethal infection²³⁸.

The host response to the seasonal H3N2 virus has been also explored in a cell line, animals and humans. In human airway Calu-3 cells, H3N2 virus replicates efficiently indicating late antiviral signalling and reduced type I and type III IFN expression compared to the LPAI H1N9 virus. Such differences have been attributed to the difference in their NS1 protein in inhibiting the antiviral response²³⁹. In contrast, in mouse promyelocyte (MPRO) cell line differentiated neutrophils, the highly virulent H3N2 virus infection revealed early and higher expression response of the IFN regulatory factor and IFN signalling pathways followed by increased expression of apoptosis-related factors than the low virulent H3N2 virus²⁴⁰. Similar strong expression of type I IFN, cytokines and apoptosis-related genes were also observed in murine lung infection²⁴¹. Genes involved in proliferation, development and metabolic processes has been down-regulated suggesting to contribute for the lethality and impaired lung repair in murine²⁴¹. The peripheral blood gene expression of human subjects infected with H3N2 virus presented with distinct transcriptional dynamics. Subjects who developed symptoms displayed concurrent multiple PRRs-mediated antiviral and inflammatory responses which are associated with oxidative stress, while individuals with no symptoms exhibit increased expression of genes that involves cell-mediated and antioxidant response²⁴². The swine influenza virus (SIV) H3N2 subtype has been also reported to induce expression of *JAK/STAT* and mitogen-activated protein kinases (*MAPK*) signalling pathways in new-born pig trachea (NPTr) cells. The inhibition of these pathways decreased type I and III IFN responses and viral replication²⁴³. In summary, most of the studies showed that strong inflammatory responses to H3N2 have been associated with the severity of the disease.

1.6.2 Host gene expressions in pandemic IAV infections

The host responses to pH1N1 in NHBE cells rapidly and directly initiated a robust inflammatory gene signature than seasonal H1N1 virus infection²⁴⁴. The transcriptome of pH1N1 and seasonal H1N1 viruses infections in human primary type I-like alveolar epithelial cells showed similar trend and extent of gene expressions. In both viruses, type III IFN has higher expression extents than type I IFN response. Unlike pH1N1 virus infections, seasonal H1N1 virus infection down-regulates genes associated with nucleic acid binding, transcription and translation²⁴⁵.

Different inflammatory cytokines and kinetics of host transcriptional response have been observed in responses to pH1N1 in the lung of mice, macaques and swine²⁴⁶. The expression

profile in macaques described enhanced viral RNA sensing, immune cell signalling and cell cycle arrest. The prolonged immune response has been associated with the enhanced pathology of the diseases²⁴⁷. The IAV gene expression profile of peripheral blood from pneumonia patients infected with pH1N1 was characterized by up-regulation of genes associated with cell-cycle regulation and cell death pathways²⁴⁸. Expression profile of peripheral blood from critically ill and hospitalized patients infected with the 2009 pH1N1 showed highly elevated level of innate cytokine-chemokine in serum, and impaired expression of genes involved in adaptive immune response (antigen presentation, B and T cell development and differentiation), apoptosis and protein ubiquitination²⁴⁹.

The blood RNA transcriptome from children hospitalized with pH1N1 infection showed elevated expression of genes related to inflammatory pathway and down-regulation of genes involved in protein synthesis and adaptive immune pathway²⁵⁰. The pathogenicity of pH1N1 has been further explored in mice that increased pathogenicity was associated with an early and continued inflammatory and interferon response, activation of *NF- κ B*, early and persistence infiltration of immune cells, and deactivation of tripartite motif-containing 24 (*TRIM24*) and lipid metabolism²⁵¹. The determinants of "Spanish Flu" H1N1 IAV pathogenicity have been suggested to be the HA protein and viral RNA polymerase complex. However, HA protein was responsible for increased host gene expression and response in mice²⁵².

1.6.3 Host gene expressions in HPAI infections

Several studies have explored the host response of the HPAI H5N1 virus. The transcriptional response of NHBE after infection with H5N1 virus induces increased up-regulation of pro-inflammatory cytokine and chemokine molecules and antiviral response including type I and type III IFNs²⁵³. Similarly, HPAI H5N1 and H7N7 viruses were characterized by increased induction of *p38MAPK* followed by IFNs and cytokines in endothelial cells of mice and human. Specific inhibitors of *p38MAPK* resulted in reduced expression of type I IFN and other inflammatory cytokines and consequently protects the mice from death²⁵⁴. In addition, the HPAI H5N1 virus induced prominent innate immune response in mice than its LPAI counterpart²⁵⁵. In a murine model, PB1-F2 protein expression of the HPAI H5N1 virus resulted in damage to the epithelial cells by inhibiting early cellular pathways of macrophage activation. Consequently, its expression induces the inhibition of NK cell activity and increased transcriptional expression of chemokines and cytokines²⁵⁶. In contrast, in chickens, it was

suggested that the viral PB1-F2 expression intensely affects the immune response probably by reducing the pathogenicity at low infection dose²⁵⁷.

Pathways from differentially expressed genes of duck trachea infected with HPAI H5N1 virus infection enriched in immune response pathways including *CXCR4* and TLRs signalling pathways and *TRIM33*²⁵⁸. The expression of ferret's lung from HPAI H5N1 virus infection also induced strong expression of IFN response and *CXCL10*. Treatment of the infected ferrets with an antagonist of the *CXCL10*'s receptor *CXCR3* resulted in reduced severity and delayed lethality²⁵⁹. The H5N1 virus infected type II pneumocytes of the cynomolgus macaques induced intense and sustained expression of type I IFN, inflammatory and innate immune genes. In addition, at early infection of H5N1 virus prolonged infiltration of T cells and apoptosis of activated dendritic cells were noted in macaque's lung²⁶⁰. In contrast, the global gene expression of severe diseases from H5N1 virus infection of cynomolgus macaques indicated that delayed or weakened *INF* activated innate immune response, apoptosis and antigen presentation leads to subsequent severe illness²⁶¹. Besides, the transcriptome analysis of human monocyte-derived macrophages infected with HPAI H7N7 and H5N1 viruses revealed weakened inflammatory cytokines, inflammasome and type I IFNs²⁶². Taken together, the host response to HPAI H5N1 virus depends on the host type and either highly induced or delayed inflammatory and innate immune responses may play a role in its pathogenicity. This may be attributed to the viral proteins M2, NS1, PB-F2, and PA proteins of HPAI H5N1^{255,257,262-264}.

The newly emerged LPAI H7N9 virus induced a significantly lower level of INFs and pro-inflammatory cytokines compared to H5N1 in human monocyte-derived dendritic cells (moDCs)²⁶⁵. The transcriptomics response of human lung epithelial cells of the H7N9 virus was close to human H3N2 than the HPAI H5N1 and H7N7 virus infections. Although specific expression of eicosanoid signalling and chromatin modification was observed in H7N9 virus infections. Lack of antigen presenting pathways and low induction of pro-inflammatory cytokines were common among all the HPAI²⁶⁶. In mice models, the responses to H7N9 and H7N7 viruses at early infection were intermediate between the pH1N1 and H5N1 virus infections. When the infection progressed, the H7 subtypes resembled the response of H5N1 virus and were characterized by increased transcription of cytokine response genes and down-regulation of lipid metabolism and coagulation signalling genes²⁶⁷. The gene expression of five subtypes (H1N1, H3N2, H5N1, H5N2 and H7N1 viruses) have been characterized in human

pulmonary epithelium (A549) cells and 16 up-regulated genes common to all of the subtypes which were associated with virus production and immune response (ISGs) were identified²⁶⁸. Similarly, the ISG were induced upon infection of the LPAI H9N2, H5N2, H5N3 and the human strains H1N1/WSN viruses. However, the LPAI exhibited strong stimulation of type I and III IFN expression and continued activation of *STAT1* than pH1N1 and H1N1/WSN strains²⁶⁹.

1.7 Genome-wide siRNA screens for identification of important IAV host factors

Viruses solely depend on cellular machinery for their effective replication and production of progeny viruses. To define important host factors in cellular machinery that support virus replication, single gene knockdown studies have been used. Recently, a high-throughput technique, small RNA interference (siRNA) enable us to investigate the role of the whole genome of a cell simultaneously in a particular condition (e.g. virus infection). The siRNA or RNA interference is a mechanism by which eukaryotic cells regulate their target genes through the generation of double-stranded (ds) RNAs in a cell. The generated dsRNA is cleaved by the enzyme Dicer into small (~20 nucleotides) ds-RNA fragments (siRNA). Small single-stranded RNA molecules which are generated from the unwound siRNA specifically bind mRNAs of target host genes and are then recognized by the RNA-induced silencing complex (RISC) to activate post-transcriptional gene silencing by degrading the target mRNA (reviewed in²⁷⁰). Likewise, transfection of eukaryotic cells with high-throughput synthetic siRNA can specifically silence gene expressions. The cells can be transfected directly by siRNA or indirectly through plasmid/virus-encoded short-hairpin RNA^{271,272}. This technique has been frequently used in functional genomics to investigate phenotypic effects of silencing a gene in various disciplines including virus-host interaction²⁷³. In influenza, siRNA screens are typically set up to study if silencing a human host gene would either increase or decrease the capability of the virus to replicate²⁷⁴⁻²⁷⁷. In the first case, the identified gene would have an antiviral function. In the second case, the identified gene would be a host factor required for the virus which consequently would be a potential therapeutic target.

Several whole genome-wide siRNA or specifically targeted siRNA screens have been conducted to identify IAV pro-viral or antiviral host factors. Although several IAV required host factors (IHF) have been identified in the siRNA screens, very few of them were verified

mechanistically. The first genome-wide siRNA screen attempt was done in *Drosophila*, which identified 176 host factors significant for IAV replication. Upon silencing of these factors, H1N1/WSN IAV replication was significantly inhibited, indicating their role in the life cycle of the virus. Among these host factors, human homologues of *ATP6V0D1*, *COX6A1* and *NXF1* have been verified in HEK 293 cells in H5N1 and H1N1 virus infections. These host factors were suggested to have a role in acidification and fusion, mitochondrial function, and viral RNA nuclear transport respectively²⁷⁸.

In A549 cells, 287 host genes influenced the replication of H1N1/WSN. Further validation revealed that subset of these factors were important for the replication of seasonal H1N1, pH1N1 and HPAI H5N1 viruses. Among which SON DNA binding protein (SON) was found to support the trafficking of IAV virion to late endosomes in early infections. Another host factors CDC-like kinase 1 (*CLK1*) was involved viral M2 protein splicing²⁷⁹. A similar approach with the same type of cells and virus has also identified 295 cellular factors required for early-stage IAV replication. These factors were highly enriched in kinase, phosphatase and ubiquitination signalling. Detailed analysis also identified that 23 factors including vATPase, glycogen synthase kinase 3 (*GSK3*)-B, *COPI*-protein families and fibroblast growth factor receptor (*FGFR*) were important for IAV entry. In addition, nuclear import components, *CAMK2B* and proteases were important for postentry replication of the virus²⁸⁰. The siRNA screen of human protein kinases (*HPKs*) indicated that 22 *HPKs* were important for the replication of H1N1/WSN and/or pH1N1 virus and were regulated by miRNA during infection²⁸¹.

Another study that used pooled shRNA screen identified E3 ubiquitin ligase, Itch, as a crucial host factor for the early life cycle of IAV. Its knockdown impaired the release of a vRNP complex containing endosome into the nucleus²⁸². Similarly, a meta-analysis of the siRNA screens and follow-up validation identified putative ubiquitin ligase (*UBR4*) as important IHF by promoting apical transport of viral proteins through association with viral M2 protein²⁸³. The combination of yeast two-hybrid analysis and siRNA genome-wide screen in HBECs uncovered the host interaction of PR8/H1N1 virus infection. The study identified 616 host factors and mapped the physical interaction networks of viral and host protein interactions²⁸⁴. The study found host factors that affect both IFN responses and viral replications which include

inflammasome-related sensors (*NLRP9*, *NLRP9*, *NLRP10*, *NOD2*), WNT, P53, ER stress, apoptosis and promyelocytic leukaemia protein (PML) pathways²⁸⁴.

Host factors that can resist influenza induced cell death have been also identified in lentiviral-mediated genome-wide screen in A549 cells. Silencing of these genes protects virus-induced cell death through prevention of virus entry to the host cells (e.g.*USP47*) and blocking the late stage (e.g.*TNFSF12-13/TNFSF13*) of IAV replication²⁸⁵. Similarly, 110 host factors which resist H3N2 IAV induced cell upon their knock down have been identified by Random Homozygous Gene Perturbation (RHGP) method. Among these, *MRPL42*, *COX5A*, *TAPT1* or *SLC25A25* were verified to increase the susceptibility to IAV infection²⁸⁶.

Brass et al., 2009, recognized more than 120 host factors that supported the vesicular trafficking, endosomal acidification, mitochondrial metabolism and RNA splicing in the lifecycle of IAV. Conversely, unlike other studies this study focused on validation of the antiviral host factors including *IFITM1*, 2, and 3 for which knockdown increased IAV. These proteins are induced by type I and II IFN and are important inhibitors of IAV at an early stage of the life cycle (endosomal acidification)²⁸⁷. Recently, a targeted siRNA screens on 35 human helicases enzymes (DEAD-box helicases (DDX proteins)) identified *DDX19* as one of an important host factor that promotes viral mRNA nuclear export²⁸⁸. Another siRNA screen which targets 2,732 druggable human genes identified 6 important host factors of which knockdown of Acid phosphatase 2 (*ACP2*) was verified to block virus entry through inhibition of nucleocapsid uncoating and nuclear import of viral RNPs²⁸⁹.

1.8 IAV interactome

Mapping of IAV-host interactions is very important for understanding the process of infection, adaptation, pathogenesis and identification of therapeutic targets. Several studies investigated the interaction of IAV viral proteins with host factors (cellular proteins) using yeast two-hybrid (Y2H), and affinity purification (AP), tandem AP (TAP), and coimmunoprecipitation followed by mass spectrometry (MS) or MALDI-TOF/MS. Similar to the siRNA screen studies discussed above, several IAV interacting host proteins have been identified by these methods, however, very few of them were verified mechanistically.

Many of the IAV polymerase associated proteins were involved in the replication, transcription and nuclear transport of IAV lifecycle^{290,291}. For instance, the polymerase interacting proteins (*NXF1*, *RMB15B*, *DDX19B*, importin β 3, and *PARP-1*) were suggested to

have a role in the nuclear transport of viral mRNA^{292,293}. Other proteins like protein phosphatase 6 (PP6)²⁹⁴, *NUP62*²⁹², and nucleophosmin (NPM)²⁹³ were required for the replication of the virus. DEAD box RNA helicase *DDX17/p72* was also identified as an important host factor for adaptation H5N1 virus polymerase to human cells²⁹⁵.

Host proteins that interact with IAV polymerase subunits have been documented. The cytosolic chaperonin containing TCP-1 (CCT) interact with PB2 protein and guide its folding and incorporation into the polymerase complex²⁹⁶. *TRIM32* and plakophilin 2 (*PKP2*) are antiviral host factors that interact with PB1 protein. *TRIM32* directly ubiquitinate and degrade PB1 protein, while *PKP2* compete with PB2 protein for PB1 protein binding resulting perturbation of IAV polymerase complex^{297,298}. Guanine nucleotide binding protein (G protein) beta polypeptide 2 (Gβ2) and macrophage migration inhibitory factor (*MIF*) interact with PB1-F2 protein, however, the mechanism of their effect on the host response is not known²⁹⁹. The PA protein binds with many mitochondrial proteins, and its interaction with *AIFM1* induces apoptosis³⁰⁰.

Nucleolin interacts with NP and has an antiviral effect through alteration of late viral genes (M1 and HA) expressions³⁰¹. On the other hand, NS1 protein interacts with *PRP19* that supports viral replication³⁰². Similarly, the β subunit of the F-type proton-translocating ATPase (F1Fo-ATPase) interacts with NS2 and it is important for virion formation and budding of the virus³⁰⁰. Despite great efforts in identifying IAV-host interactions, more research is needed to identify the key players in IAV replication and cellular responses.

1.9 Anti-IAVs therapy

Treatment of IAV involves the abolishing of the virus life cycle and/or reduction of the cytokine storm of the host immune responses. Drugs and several inhibitory molecules targeting IAV viral proteins and IHFs are discussed below.

1.9.1 Anti-influenza virus inhibitors targeting the viral proteins

1.9.1.1 HA protein inhibitors

The HA protein inhibitors interfere receptor binding of HA, HA-mediated membrane fusion, proteolytic cleavage of HAO, and maturation of HA protein. Several SA mimicry molecules have been shown to competitively bind the HA protein and significantly inhibit IAV entry. These includes synthetic polymeric sialosides or liposomal carriers (poly (acrylic acid-co-

acrylamides) and dextrans bearing pendant glycy l aminobenzyl sialoside groups)³⁰³, glycopolymers carrying terminal $\alpha(2, 3)$ -linked SA on lactosamine repeats³⁰⁴, saponins with 3-*O*- β -chacotriosyl residue³⁰⁵, a lectin from the green alga *Boodleacoacta* (BCA)³⁰⁶ and peptides³⁰⁷.

Among the HA-mediated membrane fusion inhibitors, Arbidol is the first HA-targeted anti-IAV drug licensed to date. In addition, it is also anti-inflammatory, hence has the benefit in reduction of virus replication and IAV induced cytokine storm (reviewed in³⁰⁸). Nitazoxanide is a small molecule with broad anti-influenza activity that entered phase III clinical trial. Tizoxanide, the active form of nitazoxanide have been demonstrated to inhibit the maturation of HA protein by blocking its terminal glycosylation, the trafficking from ER to Golgi, and assembly (reviewed in³⁰⁹).

1.9.1.2 Neuraminidase protein inhibitors

Neuraminidase inhibitors impair the entry and budding of IAV by interfering with the receptor binding and blocking sialidase activity of NA protein. Approved drugs with NA inhibitors include oseltamivir (Tamiflu), zanamivir (Relenza), laninamivir, peramivir, laninamvir octanoate (reviewed in³¹⁰). These drugs are used for prophylaxis or treatment of seasonal, pandemic and avian strains of IAVs. Zanamivir, oseltamivir and peramivir are administered through inhalation, orally and intravenously respectively. Oseltamivir and Zanamivir are in use for influenza A and B viruses, but they are more effective in IAV compared to influenza B viruses. Zanamivir is more effective than oseltamivir in paediatric patients and exhibits lower emergence of drug resistance³¹¹. Although the benefit of these drugs against IAV infections as therapeutics was controversial, it was suggested that early treatment with NA inhibitors reduced the mortality of hospitalized cases³¹².

Drug-resistance to NA inhibitors (e.g oseltamivir) is relatively low (ranging from 0.3% to 14%), and the NA protein inhibitors are still the choice of IAV therapy³¹³. To improve the drug resistance and effect of NA protein inhibitors, drug analogues to the NA protein inhibitors³¹⁴, modifications with conjugates³¹⁰, and other NA protein inhibitors³¹⁵ including flavonoids³¹⁶ have been investigated and showed promising therapeutic options.

1.9.1.3 Neutralizing monoclonal antibodies

Monoclonal antibodies (mAbs) targeting the HA, M2, NA, and NP protein have been demonstrated to have promising antiviral activity against IAV infection, and some of them

entered phase I and II clinical trials³¹⁷⁻³¹⁹. The anti-HA protein antibodies neutralize the HA protein and block IAV infections, while the anti-M2 protein antibodies (e.g. TCN-032) activate ADCC-mediated killing of IAV infected cells. For example, the RG-7745, VIS-410, CR-6261, CR-8020 are among the mAbs targeting HA protein which entered the clinical trials³¹⁹. The anti-NA protein mAbs targeted the conserved region of NA protein (amino acid 222-230) provided the broad inhibitory activity against wide range of IAV and influenza B subtypes including the drug resistant strains³¹⁸. Further, passive immunization with viral polypeptide specific antisera particularly against NP protein protected birds from a viral infection, suggesting their potential used in poultry industry IAV outbreaks³²⁰.

Although neutralizing mAbs have been approved for the treatment of cancers and immune disorders, its production cost, inadequate pharmacokinetics, accessibility and retention in the targeted tissue, and impaired interactions with the immune system were the major limitations³²¹. In addition, mAbs have limited use in mixed infections, mutant strains and needs detailed pre-diagnosis of the infectious agent³²². For IAV infections, passive immunization using mAbs may be very expensive, however prophylactic treatment could be valuable for vulnerable individuals (immunocompromised, children, elderly, and pregnant women). In addition, mAbs have a potential benefit for the treatment of severe cases of IAV infections (e.g. H5N1 and H7N9 viruses)³¹⁹.

1.9.1.4 M2 ion channel protein inhibitors

The adamantane drugs (amantadine and rimantadine) are the potent inhibitors of the M2 protein ion channel which have been approved by FDA for prophylaxis and therapy of IAV. During the uncoating process of IAV, the M2 protein of the virus allows the transport of H⁺ and acidification of the endosome. In addition, in the HPAI the M2 protein protects the premature changes of HA by maintaining the pH of Golgi in its synthesis⁷⁷. Adamantane drugs bind to the M2 protein ion channel and block the transport of H⁺ and endosomal acidification^{323,324}. Such blockage occurs through the hydrophobic interaction of the N-terminal gate of the M2 protein channel and adamantane group of amantadine or rimantadine. Mutation in the channel pore-like S31N, V27A and L26F disrupts the hydrophobic interaction between the drugs and M2 protein channel pore leading to enlarged pore and transport of H⁺³²⁵.

The high proportion of resistance (90%) of the recently circulating strains of IAV to adamantanes and their low effectiveness in preventing influenza-like illnesses, and

gastrointestinal and central nervous system (more in amantadine) adverse effects contributed for their ban from prophylactic or therapeutic use³²⁶⁻³²⁸. Some improvements were ongoing to target M2 protein ion channels. For instance, newly synthesized pinanamine and camphor derivatives were capable of inhibiting the wild type and partially the drug resistant strains^{329,330}. Recently, it was also modelled that the modification of amantadine by replacing its one hydrogen with an aromatic group or secondary amides would stabilize the blockage of the M2 protein ion channel against the wild and mutant strains³²⁵.

1.9.1.5 IAV Polymerase protein inhibitors

Among the polymerase inhibitors of IAV, favipiravir (T-705; 6-fluoro-3-hydroxy-2-pyrazinecarboxamide) is a broad range antiviral drug under phase III clinical trial. It is effective in inhibiting wide range RNA viruses including human and avian strains of IAV and was approved in Japan for future IAV pandemics. The detailed mechanism is not known however it was believed that favipiravir is recognized by the IAV polymerase mistakenly as a purine nucleotide resulting in reduced viral replication^{331,332}. The guanosine analogue, ribavirin has been also used for the treatment of several RNA viruses including IAVs. However, its use was inconclusive as it has a risk of haemolytic anaemia and teratogenicity³³³. Recently, the PB2 protein cap-snatching blocker VX-787 and the PA protein endonuclease inhibitor ANA-o were newly identified polymerase inhibitors that have an effect on several IAV subtypes^{334,335}.

1.9.1.6 Inhibitors of other viral proteins

Structural based screening of small molecules and single-stranded DNA aptamer that inhibit binding of the viral NS1 protein with cellular mRNA and proteins has been demonstrated in few studies (reviewed^{336,337}). A small-molecule drug, the so-called nucleozin that targeted the NP of IAV has been identified recently and showed promising protective results against H5N1 IAV in cells and mice³³⁸. It has been suggested that nucleozin triggers the non-functional aggregation of NP protein lowering its nuclear accumulation. However, recently it was reported that the primary target of nucleozin is the viral RNP in that; it blocks the cytoplasmic trafficking of RNPs inducing the formation of large perinuclear aggregates together with *Rab11*³³⁹.

1.9.2. Host-targeted antiviral

One of the major problems concerning anti-influenza drugs that target viral proteins is the frequent emergence of drug-resistant viruses. Indeed, most seasonal H3N2, H1N1 and the 2009

pH1N1 influenza viruses are now resistant to amantadine/rimantadine, and oseltamivir-resistant H1N1 viruses have emerged in many countries. Therefore, concern regarding the emergence of drug-resistant isolates is clearly warranted. Hence, there is an urgent need for new antiviral targets, including non-viral targets, to treat influenza virus infection while minimizing the emergence of drug resistance. Host factors have been in the spotlight recently as potential antiviral targets to overcome the problems described above. The emergence of resistance could be less frequent when host factors are targeted compared with directly targeting viral proteins. Recently, several drugs targeting host factors identified to be important for virus replication have been developed²⁷⁴.

The sialidase DAS181 (Fludase), which entered clinical trials, removes the SAs from the epithelial cells and prevent IAV attachment. Several other compounds, including protease inhibitors can block cellular proteases to prevent HA protein cleavage and membrane fusion. The MEK inhibitor U0126, *NF-κB* inhibitors such as acetylsalicylic acid (known as aspirin), and agonists of sphingosine-1-phosphate receptors AAL-R, all of which targeted host cellular functions involved in influenza virus replication, have been studied to explore their potential as new antiviral drugs against influenza. In addition, combination therapy with a protease inhibitor and conventional anti-influenza drugs has been recently tested²⁷⁴.

König et al., 2010 tested a selective inhibitor of calcium/calmodulin-dependent kinase IIβ (*CAMK2B*) called KN93. KN93 regulated viral RNA transcription and reduced virus titers in influenza virus-infected cells. Karlas et al., 2010 showed that treatment with the compound TG003, a small molecule inhibitor of *CLK1* identified as a host candidate involved in influenza virus replication in their screen, led to reduced virus production in human cells probably by affecting the splicing of viral RNAs. Brass et al. found that the *IFITM1*, 2, and 3 hampered the early steps of influenza virus replication influenza virus infection, and also suggested that expression of *IFITM3* defends the host against IAVs. Watanabe et al., 2014 tested 11 compounds that inhibit the cellular functions of several host factors identified in their screen. Of these compounds, they found that ruxolitinib and golgicide A (which are inhibitors of *JAK1* and the Golgi-specific brefeldin A-resistant guanine nucleotide exchange factor 1 (*GBF1*), respectively) markedly decreased virus titers in human cells infected with influenza virus without affecting cell viability²⁷⁴.

In addition to approaches targeting host cellular proteins, Morita et al., 2013 assessed the possible involvement of lipid products in influenza virus replication. By conducting mediator lipidomics and a bioactive lipid screen, they discovered that the DHA-derived protectin D1 isomer (10S, 17S-dihydroxydocosahexaenoic acid) markedly hampers influenza virus replication by inhibiting the nuclear export of viral mRNA. More importantly, they demonstrated that intravenous administration of protectin D1 protected mice from influenza virus infection, suggesting that protectin D1 has potential as a new therapeutic tool in the treatment and prevention of influenza²⁷⁴. Other compounds that target host factors (i.e., KN93, TG003, ruxolitinib, and golgicide A) were also reviewed²⁷⁴.

1.10 Rational of the study

The best way of controlling highly spreading infectious diseases like IAV would be vaccination. However, the efficacy of seasonal IAV vaccines has been called into question. Due to rapid antigenic changes of the virus, these vaccines have to be produced annually before the outbreak, imposing massive economic loss in their preparation and delivery. In addition, the predicted strain may fail to match with the prepared vaccine strain due to potential geographical differences. There could also be population dynamics such as variability in exposure and host immunity.

Anti-IAV therapeutic agents are important alternatives to IAV vaccines. These therapeutics can reduce the replication of the virus, mortality, morbidity and the spread of the virus to healthy individuals. There are few clinically approved anti-IAV drugs for therapeutic and prophylactic use, and the drugs were targeted against viral proteins. However, these drugs have been challenged by the emergence of drug resistance, their toxicity, adverse effects and low efficacy. The ideal broad range anti-IAV drug should be able to target all IAV strains irrespective of their source. This can be achieved by targeting either common cellular targets in response to IAV infections.

To discover effective IAV vaccine and alternative anti-IAV therapeutics, it is important to understand how IAVs adapt to different species host ranges, and vice versa, how different host species respond to different IAV subtypes. It is also crucial to understand IAV-host interactions and identifying critical viral or host factors that could support the replication of the virus or induce pathologic conditions. Recently, high throughput technologies such as mRNA microarray-based gene expression, genome-wide siRNA screening and proteomics analysis are

providing an in-depth insight of host-pathogen interaction of IAVs. A comprehensive analysis of host gene expression changes in different IAV infections in different host species can potentially identify critical host determinants that could explain its host adaptation.

IHF identified by the IAV genome-wide siRNA screening studies could be used as potential anti-IAV targets. However, these studies were not consistent (reproducible) mainly due to false negative results^{277,340}. Nevertheless, protein-protein interactions (PPIs) have become increasingly important for detection of new interactions and protein networks. Thus, computational network growing of the IHFs in PPI databases (tools) will allow extension of networks with additional related IHFs and are important to improve the understanding of the greater functional context. Computational gene network growing is also important for discovering novel functions of genes that could have been missed by these experimental investigations³⁴¹. On the other hand, several of the identified IHFs were not characterised in follow-up studies and identification of the promising candidates from the hundreds of candidate host factors has been challenging. Recently, it has been suggested that computational integration of multiple "OMICS" data set (e.g. transcriptomics, proteomics and siRNA screens) will be more effective in identifying the critical host factor candidates and pathways utilised by the IAVs³⁴².

Developing new drugs is challenging and expensive since the process may take decades. Drug repurposing is a recent strategy that has been proposed as one possible solution to this problem. In this strategy, existing drugs that are in clinical use and with known safety profiles are evaluated for either anti-viral activity or as therapeutics to treat virus infection. System-wide phenotypic data together with bioinformatics methods have been used for *in silico* prediction for repurposing (FDA) approved drugs as alternative therapeutics against infectious diseases and cancer^{343,344}. Therefore, it is reasonable to computationally predict drug-host interactions to identify new drugs that could be suggested as anti-influenza targets.

1.11 Hypothesis

Investigating IAV-host interactions using transcriptome, interactome and siRNA screen data, and gene/protein network analysis may lead to i) identification of therapeutic host targets for IAV; ii) understanding the molecular mechanism of host adaptation iii) and identification of virus-specific characteristics.

1.12 Objectives

1. To identify IAV plausible therapeutic target through integration of transcriptome, siRNA screens and interactome data set.
2. To identify host determinants of IAV host adaptation through comparison of cross-species gene expression signatures (A549, CEF and MDCK) during different IAV infections.
3. To investigate murine macrophage responses during LPAI and human IAV infections.
4. To benchmark selected network growing tools in the context of virus-host interactions.
5. To predict FDA approved drugs that could target IHFs for IAV therapeutics.

Chapter 2: Materials and methods

2.1 Data sources

2.1.1 Microarray data

The microarray data for this study were obtained from Richard J Sugrue's lab from previous experiments. The data were also deposited in National Center for Biotechnology Information (NCBI), Gene Expression Omnibus (GEO) Data set with accession number GSE31524 (A549, CEF and MDCK cells infections) and GSE86446 (macrophage infections).

Table 2.1: The characteristics of microarray data set used in the current study

Cells	IAV strains	Time points (hrs)	Affymetrix platform	References
A549	A/WSN/1933 (H1N1/WSN) (VR-1520)	2, 4, 6, 10	Genechip Human	345
	A/Singapore/478/2009(H1N1) (pH1N1/478)	2, 4, 6, 8, 10	Genome HG U133	
	A/Duck/Malaysia/02/2001 (H9N2)	2, 4, 6, 8, 10	Plus 2.0 Array	
	A/Duck/Malaysia/F118/2004 (H5N2/F118)	2, 4, 6, 8, 10	(GPL570)	
	A/Duck/Malaysia/F189/2004 (H5N2/F189)	10		
	A/Duck/Singapore-Q/F119/1997 (H5N3)	10		
	A/Duck/Malaysia/F59/2004	10		
	H7N9/AICHI	2,10		
CEF	A/WSN/1933 (H1N1/WSN) (VR-1520)	2, 4, 6, 8, 10	GeneChip Chicken	345
	A/Duck/Malaysia/02/2001 (H9N2)	2, 4, 6, 8, 10	Genome Array	
	A/Duck/Malaysia/F118/2004 (H5N2/F118)	2, 4, 6, 8, 10	(GPL3213)	
	A/Duck/Malaysia/F189/2004 (H5N2/F189)	10		
	A/Duck/Singapore-Q/F119/1997 (H5N3)	10		
	A/Duck/Malaysia/F59/2004	10		
MDCK	A/WSN/1933 (H1N1/WSN) (VR-1520)	2, 4, 6, 8, 10	GeneChip Canine	345
	A/Duck/Malaysia/02/2001 (H9N2)	2, 4, 6, 8, 10	Genome 2.0 Array	
	A/Duck/Malaysia/F118/2004 (H5N2/F118)	2, 4, 6, 8, 10	(GPL3738)	
	A/Duck/Malaysia/F189/2004 (H5N2/F189)	10		
	A/Duck/Singapore-Q/F119/1997 (H5N3)	10		
	A/Duck/Malaysia/F59/2004	10		
Macrophages	A/WSN/1933 (H1N1/WSN) (VR-1520)	2, 24	GeneChip Mouse	346
	A/Duck/Malaysia/F118/2004 (H5N2/F118)	2, 24	Genome 430	
	A/Duck/Singapore-Q/F119/1997 (H5N3)	24	2.0Array	

The cell types, the virus strains, the time post infections in hours and the platforms they used were described in Table 2.1. The LPAI isolates (H5N2/F59, H5N2/F118, H5N2/F189, H5N3 and H9N2 viruses) were isolated by the Agri-Food and Veterinary Authority of Singapore³⁴⁷. H1N1/WSN virus was the American Type Culture Collection (ATCC) strain, while the pH1N1/478 was isolated in Singapore during the 2009 pandemic. The sequences and replication of these viruses in different cell types have been characterised previously^{269,347}. The microarray data contains triplicate data (sample) for each virus isolate at each time points and their corresponding controls (mock infections).

2.1.2 Genome-wide siRNA screen data

Till November 2015, there were eleven siRNA screening studies that identified influenza required host factors (IHF) or antivirals (Table 2.2). These studies investigated either larger set genes (e.g. kinases)^{348,349} or whole genomes^{278,279,286,287,350-354}. The genes or IHFs from each siRNA screening studies were curated manually and were refined using a custom Perl script (Appendix A). We noted that a single gene may have multiple Entrez IDs and names and for the sake of consistency we used the “Official gene symbol” to represent a gene. For the studies that used non-human host cells (e.g. DL1²⁷⁸ and MDCK cells²⁸⁶) in their experimental design (Table 2.2), the human homologues of the corresponding genes were used in the analysis. List of IHFs isolated from each siRNA screening study was described in Appendix A.

2.1.3 IAV interactome data

Studies showing influenza virus host-interactome using high throughput screening were identified through online PubMed literature search. Till November 2015, we identified 13^{290,295,296,299,351,354-361} interactome studies that describe a list of host proteins that interact with the IAV proteins (Table 2.2). The viral protein-host protein interactions from each study were manually curated and official gene symbol was used to represent the host proteins. The interaction information from the studies is then summarised by the viral proteins (all the proteins that interact with each viral protein were summarised). The summarised list of viral protein-host protein interactions are described in Appendix A.

2.3 Microarray data analysis

AltAnalyze 2.1.0 and limma (R-package) software were used to analyse the microarray data (Affymetrix .CEL files)³⁶². Proper quality control and normalisation were performed in accordance with the AltAnalyze for Affymetrix.CEL files. AltAnalyze utilises inbuilt robust multichip average (RMA) method for normalisation and limma (R-package) based moderated t-test option (unpaired, assuming equal variance) to calculate an adjusted variance based on the variance of all probes analysed for each comparison. The mean expression for each biological group and mean detection above background probabilities were calculated for each feature (probes). Expression values were calculated if the mean expression and the mean detection above background probabilities for at least one group meet expression thresholds, expression > 70 and p-value < 0.05. Differentially expressed probe sets were selected with statistically significant change (P-value of the student-t-test (raw p-value), ANOVA and the Benjamini-Hochberg False Discovery method (adjusted p-value) < 0.05) and |Log Fold change (FC)| > 1 (changes in virus infection in respect to corresponding mock). A single gene might be represented by more than one probes, hence those genes with an adjusted p-value < 0.05 and their probe mean |log FC| > 1 were considered as differentially expressed genes (DEGs). In case of macrophage data analysis, probe sets with adjusted p-value < 0.05 and |log FC| > 1 were considered as DEGs and probesets with highest |log FC| > 1 were used to construct heat maps. The DEGs in at least one virus at one-time point infection with their corresponding logFC values were used for average linkage hierarchical clustering using Gene Cluster 3.0³⁶³. The heat maps of the clustering were viewed by Java Tree Viewer. In addition, customised Perl scripts (Appendix A) were used to analyse the overlap of DEGs between time points and viruses. Venn diagrams showing the overlap of genes in different groups was performed using Venny 2.1³⁶⁴.

2.4 Functional annotations (GO BP and KEGG pathway analysis)

Functional annotations of the DEGs were analysed using the Gene Annotation and Analysis Resource-Metaspape⁴⁵ and q-value < 0.05 was used to select the top 20 statistically significant pathways. First, both the up- and down-regulated genes at each time point and virus infections were used for Gene Ontology Biological Process (GO BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Then the up- and down-regulated genes were used separately to identify the up- and down-regulated KEGG pathways or GO BPs.

Table 2.2: Summary table for siRNA screening studies

Genome wide Screening type	Cell type	Virus type	Dose/ MOI	Post infection incubation	Number of targeted genes (company)	Reference
siRNA	A549	H1N1/WSN	0.5	12, 24 and 36 hrs	Over 19,000	³⁵⁰
siRNA	A549	H1N1/WSN	0.12	24 hrs	22,843 (Qiagen)	²⁷⁹
shRNA	A549	H1N1/WSN	1	2 weeks	16,368 (National RNAi Core Facility in Taiwan)	²⁸²
siRNA	A549	H1N1/WSN or pH1N1	0.001	48hrs	720 (Dharmacon)	²⁸¹
shRNA	A549	H3N2	7	72 hrs	21,415 (Thermo cientific Open Biosystems, Ottawa, ON, Canada)	²⁸⁵
siRNA	DL1	H1N1/WSN	10	24hrs	13,071 (Ambion)	²⁷⁸
siRNA	HBECs	PR8/H1N1	1	48hrs	1745 (Dharmacon)	²⁸⁴
siRNA	U2OS	PR8/H1N1	0.2–0.3	12 hrs	21,122 (Dharmacon)	²⁸⁷
RHGP	MDCK	H3N2	0.1	48 h	Whole genome	²⁸⁶
siRNA	A549	H1N1/ Brisbane/59/2007	0.005	24, 48 and 72 hpi	30 host kinases	³⁶⁵
siRNA	HEK293	H1N1/WSN	50 pfu/well	48hrs	1,292 IAV interacting host factors (FlexiTubesRNA; QIAGEN)	³⁶⁶

Table 2.3: Summary table for interactome studies

Cell type	Methods	Virus strain (isolate)	Viral protein	Interacting cellular proteins	References
293 human embryonic kidney (HEK)	Pull-Down Assay and Mass Spectrometry	H1N1/A/WSN/33	11 viral proteins	1,292	³⁵⁴
293T cells	Tandem affinity purification followed by MS	H1N1/A/WSN/33	vRNP and viral polymerase complexes	45	³⁵⁹
HeLa cells	Yeast two-hybrid	IAV H1N1 strain A/Sw/NC/44173/00	NP	10	³⁶⁰
293T	Yeast two-hybrid screens and literature mining	A/PR/8/34 (H1N1) A/WSN/33 (H1N1) A/VietNam/1194/200 H5N1 A/Turkey/651242/2006/ H5N1 A/Victoria/3/75 (H3N2)	PA, PB1, PB2 and NP	234	³⁵⁸
Primary HBECs	Yeast two-hybrid	PR8/H1N1	10 viral proteins	87	³⁵¹
293T	Affinity purification followed by MS	H1N1 A/WSN/33	PB2	171	³⁶⁷
293T	Yeast two-hybrid	PR8/H1N1	PB1-F2	8	²⁹⁹
293T	Infectious Protein Complementation Assay	H1N1 A/WSN/33	Polymerase complex	84	³⁵⁷
-	Bioinformatics	H1N1/WSN and H5N1	Polymerase complex	54	²⁹⁵
293T	Tandem affinity purification MALDI /MS.	H1N1 A/WSN/33	Polymerase complex	10	²⁹⁰
293T	Biochemical co-purification followed by MS	H1N1 A/WSN/33	PB2	7	²⁹⁶
293T	Interactome and prediction	H5N1	NS1	69	³⁵⁶
A549 cells	liquid chromatography-tandem mass spectrometry (LC-MS/MS)	H5N1 Vietnam/1203/04	PA, PB1, Polymerase complex	400	³⁵⁵

2.2 Databases and software packages used in this study

Several databases and software packages were used in this study and were described in Table 2.4.

Table 2.4: Databases and software packages used in the current study

Databases	Use in this study	Manufacturers/developers	References
DAVID	Gene Ontology Biological Processes (GO-BP), KEGG pathway analysis and annotations	Leidos Biomedical Research, Inc. USA.	³⁶⁸
DrugBank	Drug-host target interaction	University of Alberta, Edmonton, AB, Canada.	³⁶⁹
Ensembl 83	Source for genomic sequences of canine and chicken, and annotation	European Molecular Biology Laboratory European Bioinformatics Institute (EMBL-EBI), UK.	³⁷⁰
GeneMANIA	Network growing analysis	University of Toronto, Canada.	³⁷¹
GEO Data set	Microarray data source	The National Center for Biotechnology Information (NCBI), NIH, USA.	
HGNC	Source for human official gene symbol lists	EMBL-EBI, UK.	³⁷²
IPA	Network growing analysis	Qiagen bioinformatics, Netherlands.	³⁷³
MetaCore™	Drug-host target interaction	Thomson Reuters, Toronto, Canada.	³⁷⁴
Metascape	GO BP and KEGG pathway analysis	Sanford Burnham Prebys Medical Discovery Institute, CA, USA.	⁴⁵
OMA	Orthologs mapping	Dessimoz Lab at UCL and CBRG at ETH Zurich.	³⁷⁵
STRING	Network growing analysis and PPI interactions	University of Zurich, Switzerland.	³⁷⁶
DR. PRODIS	Drugs killing index and side effects	Center for the Study of Systems Biology, Atlanta, USA.	³⁷⁷
Software/Packages			
AltAnalyze	Microarray data analysis	Cincinnati Children's Hospital Medical Center and the University of Cincinnati, USA.	³⁶²
Customised Perl scripts	Refining gene lists, Pairwise analysis and integration		
Cytoscape 3.4.0	Network topology analysis	Cytoscape consortium, San Diego, USA.	³⁷⁸
Gene Cluster 3	Cluster analysis and heat map	University of Tokyo, Japan.	³⁷⁹
NCBI blast+	Sequence alignment and probe annotation	NCBI, MD, USA	³⁸⁰
NCBI E-utilities (ESearch)	Drug-protein interaction in relation to IAV	NCBI, MD, USA	³⁸¹
OriginPro8	Graphs	OriginLab Corporation, MA, USA.	
R-packages (Rcmdr, ggplot2)	Correlation matrix, PCA analysis and other plots	R Development Core Team and CRAN.	³⁸²
STATA	descriptive statistics, Chi-square and Fishers' exact test)	StataCorp, Texas, USA	
Venny 2.1	Pairwise analysis of genes and pathways, Venn diagram	Juan Carlos Oliveros BioinfoGP, CNB-CSIC, Spain.	³⁸³

For benchmarking the network growing tools (STRING, GeneMANIA and IPA), KEGG pathway and GO BP enrichment analysis were performed using Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8³⁶⁸. First, we did KEGG pathway analysis for small and medium seed sets, as well as the true positive sets (for 153 and 1463 IHFs). Then, the pathway and GO BP analysis were repeated after growing 100 genes with each respective tool. Significantly enriched ($p < 0.05$) KEGG pathways and GO BP were compared against significant pathways and GO BPs from the true positive sets. In addition, KEGG pathway and GO BP for the 100 grown genes by each of the network growing tools were also separately compared with the significant KEGG pathways and GO BPs from the positive sets.

2.5 Integration of microarray, interactome and siRNA screening data set

The genes in each study were manually curated and custom Perl scripts were written to count and perform overlap analysis of the DEGs genes in A549 cells with the collected sets of genes from interactome and siRNA screens. The interactome and siRNA screens were investigated with different cells, virus isolates, methods and settings. To minimise the effect of cell line differences, we restricted the integration analysis in studies that used A549 cells followed by studies used H1N1/WSN and A549 cells (Table 2.2). To identify plausible host factors among the overlapped DE genes with siRNA and interactome; genes that were up-regulated in at least three virus infections at 10hpi were selected.

2.6 Probes annotation and orthologs mapping

The overall overview of probes annotation and orthologs mapping in CEF and MDCK cells gene expressions was described in Figure 2.1. Probe annotation information for Chicken Genome Array, Canine_2 Genome array and Human 133A 2 plus were obtained from Affymetrix NetaffxTM website (document deposition)³¹⁸. The non-annotated probes with no Ensembl ID in chicken and canine chips were isolated and their corresponding sequences were retrieved from chicken probe sequences, FASTA (8MB 08/20/08)³⁸⁴ and Canine_2 probe sequences (FASTA, 9MB)³²⁰ files from Affymetrix. Whole genome sequences of chicken (*Gallus gallus* (17,108 genes) and canine (*Canis familiaris* (24,580 genes) with their corresponding Ensembl IDs were downloaded from Ensembl (Ensemble genes 83)³⁷⁰. The probe sequences with no Ensembl IDs in the chicken and canine Affymetrix chips were aligned against their corresponding whole genome gene sequences from Ensembl, using ncbi-blast+ package³⁸⁰.

A single probe sequence can align with more than one gene; hence, aligned genes with the lowest p-value, higher sequence % identity and higher frequency were identified as the best hits for annotation.

Figure 2.1: Overview of Affymetrix probe annotation and orthologs mapping.

Human orthologs information for chicken and canine were obtained 1) from Affymetrix web document depository 2) by analysing orthologs using “Orthologous Matrix” (OMA) database³⁷⁵. Orthologs between two species (*Gallus gallus* vs *Homo sapiens*) and (*Canis familiaris* vs *Homo sapiens*) were analysed using OMA, and the results were obtained with Ensembl gene IDs. The chicken and canine probes orthologs to Human Genome HG U133 plus 2.0 Array were mapped to Ensembl IDs. The OMA orthologs genes were also mapped to corresponding Affymetrix probes using the Affymetrix and our annotations. The orthologs from the two data set were combined. Chicken vs human and canine vs human and human orthologs of both canine and chicken were used for further analysis.

2.7 Benchmarking network growing tools using IHFs

Three gene network growing tools STRING³⁷⁶, GeneMANIA³⁷¹ and IPA³⁷³ were used for gene network growing analysis. STRING uses experimentally well-described genes, proteins and their interactions and a number of computationally predicted interactions such as gene neighborhood (e.g. prokaryotic operons), paired fusion proteins, gene links via common evolutionary histories (phylogenetic profiles), co-transcription regulators, co-expression patterns, text-mining associations, links inferred from transferring interactions of orthologous proteins interacting in another organism and similarity of the protein structures. STRING (version 10) covers more than 2000 organisms, > 900 million interactions of 9.6 million proteins and is updated regularly³⁷⁶. STRING uses a naive Bayesian algorithm for computing combined scores from different edge types including a correction for the probability of random observation of an interaction adding related genes to grow the query network is based on the closest combined scores³⁸⁵.

GeneMANIA uses association data, including protein and genetic interactions, pathways, co-expression and co-localization similarity information and protein domain similarity data. This tool supports the network analysis for nine organisms (*A.thaliana*, *C. elegans*, *D. rerio*, *D. melanogaster*, *E. coli*, *H. sapiens*, *M. musculus*, *R. norvegicus* and *S. cerevisiae*) and uses aggregated interaction links from hundreds of data set (experimentally validated and/or computationally predicted)³⁷¹. GeneMANIA utilises two algorithms: 1) a linear regression algorithm to calculate composite functional association network (based on GO BP, molecular function and cellular compartment) from multiple networks obtained from different data sources and 2) a Gaussian field label propagation algorithm for predicting gene function from the

composite network. For a longer list of query genes (> 5) the weights are chosen automatically using linear regression maximising interaction between seed genes while minimising interactions of seed genes to other genes not on the query list^{371,386}.

The main attribute of the IPA knowledge database is the high-quality manual curation of texts from peer-reviewed journals and both public and private biomedical databases. The retrieved knowledge is structured into ontologies and the ontology made available as knowledge base for various applications including functional network analysis. Given user seed genes and the knowledge-base of the interaction of thousands of genes with each other, IPA uses a multi-stage, heuristic six step algorithm to produce networks. Briefly, IPA sorts seed genes based on their interconnectivity to construct multiple small networks. The small networks are then merged by growing with genes from the knowledge-base that can connect the small into bigger networks. Then a p-score is calculated as the probability of finding f more seed genes in a set of n genes randomly selected from the knowledge-base¹⁰⁹.

Before growing networks, we first mapped the IHFs (Appendix A) to the network growing tools. The IHFs that were mapped to the three tools were used for growing networks (Appendix A). For the network analysis we first defined two seed sets with different size (number of genes); 1) 28 genes that were shared in more than 2 studies, 2) 153 genes shared by at least two studies (Appendix A). Using these two sets of genes as a seed we automatically grew 10, 20, 50 and 100 genes with each of the network growing tools. These numbers (10, 20, 50 and 100 genes) were selected because GeneMANIA uses these fixed numbers for growing networks and the other tools allowed defining any numbers.

The second step of the network analysis was utilising randomised gene sets as seed base to confirm whether the detection rate was by chance or not. The 153 genes that were shared by two studies were randomised into five groups of 30 genes (small sets (30A to E)), and the rest 1,463 genes into 10 groups of 146 genes (medium sets (146A to J)) (Appendix A). Each set (either 30 or 146) was used as a seed to grow 100 genes in each of the networking tools. In parallel, we randomly selected an equal number of small and medium set genes from 19,006 human genes obtained from HUGO Gene Nomenclature Committee (HGNC)³⁷². The grown genes were separated from the seed genes and the newly grown genes from each network growing tool were compared against the list of known IHFs.

The third step of analysis was combining or interchanging grown genes between the network growing tools (STRING and GeneMANIA: as these tools are freely available online) to test if this improves the detection rate of known IHFs. For combination, the first 50 genes from two network growing tools were combined and compared against the list of known IHFs. For interchange, smaller sets of genes were used, and the first 50 grown genes in one tool were used the seeds to grow additional 50 genes in the other tool. Like the previous analysis, the grown genes from each network analysis were separated from the corresponding seeds and compared against the list of known IHFs.

Topological parameters (average shortest path length, betweenness centrality, closeness centrality, clustering coefficient, degree distribution, eccentricity, neighbourhood connectivity, radiality and stress) of the networks from STRING and GeneMANIA were analysed using Network Analyzer in Cytoscape v 3.4.0³⁷⁸.

2.8 Drug-host target interactions

Drugs that interact with the known IHFs (1,445 genes) and the new candidate IHFs (1,538 genes) (excluding anti-viral genes) from the network growing analysis were identified using (MetaCoreTM)³⁷⁴. Therapeutic drug-target interactions and secondary drug interactions were obtained via drug interaction analysis and those drugs with an effect on the known or new host targets were extracted if they were FDA approved and only have inhibitory effects on the IHFs. Previously *in silico* predicted or experimentally tested anti-IAV drugs or small molecules were also manually curated from previous reviews and reports (Appendix B)^{266,268,276,354,387}, and were compared using their DrugBank IDs with the current predictions.

2.9 Prioritising host targeted FDA approved drugs

In drug-host interactions, we found hundreds of drugs that could target the known and new candidate IHFs. To prioritise these drugs for further experimental screening we considered certain parameters which include: killing index, the number of side effects, the cost of the drugs; the number of the interactions between IHFs, the number of IHFs targeted by a drug; and the co-occurrence likelihood ratio of drug-target and target-influenza in Pub Med abstracts.

In the selection of drugs for IAV, considering the killing index and side effects are very important that the side effects of the drugs shouldn't exceed or be equivalent to the influenza illness. The side effects and killing index of each drug were obtained from DR.PRODIS

database³⁷⁷. In addition, drugs that target proteins that interact more with other host proteins tend to have higher toxicity (Biruhalem Taye, Sebastian Maurer-Stroh; unpublished observations). The interaction of the IHFs targeted by the FDA approved drugs was analysed using STRING and the number of interactions of each host target was used for further analysis. The costs of the drugs were obtained from DrugBank database³⁶⁹.

Text mining of the published articles in PubMed or other databases have been used to extract drug-gene relationship and the automatic text mining was equal or better than manually curated knowledge³⁸⁸. This text mining approach assumes that the higher the co-occurrence of X and Y within the literature databases, the more significant the association is likely to be³⁸⁹. Hence, we hypothesise that the higher the number of publications that the drugs and targets co-occur, the higher the activity of the drugs on the targets. We did online PubMed "esearch" and count the number of publications related to the drugs and their corresponding targets. "Esearch" executes Entrez search using terms (genes, drugs and influenza virus) in indexed fields (in the abstract of the PubMed publications). In addition, the higher the number of publications that reported target and Influenza co-occurrence would indicate the higher confidence of the target associated with influenza infections. To reduce the effect of drugs, targets or influenza that would have higher publications we calculated the likelihood ratio of publications. The "eserach" include drugs to target and targets to influenza.

$$LHR = 2 \frac{ABi}{Ai + Bi}$$

Where, LHR = Likelihood ratio, ABi = the number of PubMed abstracts that both Ai and Bi co-occur; Ai= the number of publication in Ai (e.g. drug); the number of publications in Bi (e.g. target).

Then values of each parameters were standardised by calculating Z-score.

$$zi = \frac{xi - \mu}{\sigma}$$

Where, Zi= Z-score, Xi = population mean, μ =sample mean, σ = standard deviation.

Based on our assumption, we used higher Z-score values of LHRs and number of IHFs targeted by the drugs and lower Z-score values of killing index, side effects, the number of IHFs interactions and cost of the drugs for prioritising the drugs. Finally, we used Z-score total for ranking the drugs.

$$Z_{total} = (\sum Z_i \text{ high}) - (\sum Z_i \text{ low})$$

Where:

$\sum Z_i \text{ high}$ = the summation of higher Z-score (LHRs and number of IHFs targeted by the drugs),

$\sum Z_i \text{ low}$ = the summation of lower Z-score (killing index, side effects, the number of IHFs interactions and cost of the drugs).

2.10 Statistical analysis

Descriptive statistics (percentage (%), mean of percentages) and hypothesis testing (Student-t-test: two-tail assuming unequal variance, one-way ANOVA), chi-square and fisher's exact tests were performed by STATA version 11. Cross-correlation matrix and principal component analysis were analysed using R program R Commander package³⁸² to observe patterns of DEG expressions and to discriminate the key determinant DEGs respectively. Customised Perl scripts were used to refine and count gene lists, for randomization, overlap (pairwise) analysis, and filtering the grown genes from the input seed genes. Percentage of detection for the network growing analysis was calculated as the overlap of grown genes with the positive sets of genes (either 153 or 1463 IHFs): (equation 1 for genes, 2 and 3 for pathways and GO BPs).

$$\text{Percentage} = \frac{\text{Number of grown genes in each networking tool overlapping with the known IHFs}}{\text{Total number of grown genes in growing networks (usually 100)}} \dots\dots\dots 1$$

$$\text{Percentage} = \frac{\text{Number of pathways or BPs before or after growing networks overlapping with positive set pathways or BPs}}{\text{Total number of pathways or GO BP from the whole hostfactors}} \dots\dots\dots 2$$

$$\text{Percentage} = \frac{\text{Number of pathways or BPs from the 100 recruited genes overlapping with positive set pathways or BPs}}{\text{Total number of pathways or BP from the whole hostfactors}} \dots\dots\dots 3$$

Chapter 3: Identification of IAV-therapeutic plausible host targets through integration of genome-wide transcriptome, siRNA screens and Interactome data set

3.1 Introduction

Due to the emergence of drug-resistant IAV strains for existing anti-influenza drugs, there is an urgent need for new alternative antiviral drugs which are less likely to be affected by drug-resistance mutations, antigenic shift and antigenic drift of the virus. We seek to capitalize on the recent emergence of novel antiviral drug discovery strategies through targeting cellular host factors. As described above (see Introduction section 1.7 and 1.8), there were several IHFs and IAV-interacting proteins that may serve as potential targets for anti-IAV drug development. Although exciting findings have been described by characterising only a few of these host targets, several of them were not characterised in follow-up studies. Identification of the promising candidates from the hundreds of these host factors has been often difficult. Shapira et al., 2009²⁸⁴, Watanabe et al., 2014³⁶⁶ and Tripathi et al., 2015²⁸³ integrated the siRNA and interactome data set to highlight high confidence host targets. Recently, it has been suggested that computational integration of multiple "OMICS" data set (e.g. transcriptomics, proteomics and siRNA screen) will be more effective in identifying the critical host factor candidates and pathways utilised by the virus³⁴².

Some of the studies validated the role of few host factors first by transfection of the siRNA to suppress the host factor, followed by infection with the virus^{279,350,390}. This is the best approach to dissect the role of the host factors in the life cycle of the virus even though it does not ensure their potential as therapeutic targets. This is because people usually get treated with drugs after they got infections unless for prophylaxis. To consider the host factors as a therapeutic target, one should ask their abundance during infections, which can be partially addressed by integration with gene expression data set. In this study, we first investigated the genome-wide transcriptomic responses during eight isolates of IAV infections in A549 cells. Subsequently, these DEGs were integrated with the interactome and siRNA screens to identify plausible host targets for anti-IAV therapy.

3.2 Results

3.2.1 Host gene expression changes in A549 cells infected with different IAVs

Microarray based gene expression profiles of A549 cells upon infection of eight different IAV isolates were analyzed. The H1N1/WSN, pH1N1, H5N2/F118 and H9N2 virus infections had temporal expression data, while the rest (H5N2/F59, H5N2/F189, H5N3 and H7N9 viruses) were obtained at 10hpi. Overall, we identified a total of 5,563 DEGs in at least one of the observations ($\text{adj.p-value} < 0.05$ and $|\log^{\text{FC}}| > 1$).

The kinetics of up-and down-regulated genes from the four isolates at 5 time points post infection were described in (Fig. 3.1a and b). Generally, the numbers of up- and down-regulated genes were increased over time. However, in pH1N1 the numbers of both up- and down-regulated genes were decreased from 2 to 6 hpi and were increased from 6 to 10 hpi. The kinetics of up-regulated genes were in line with the replication of the viruses for two isolates (H5N2/F118 and H1N1/WSN viruses), while the trend was different in pH1N1 and H9N2, where they had slightly increased and constant viral titer unlike their corresponding up-regulated expression trend respectively (Fig.3.1a)³⁴⁵.

The trends of down-regulated genes were matching the replication behavior of the respective viruses. These down-regulated genes were relatively higher in H1N1/WSN and H5N2/F118 virus than in H9N2 and pH1N1 virus (Fig.3.1b). This result was similar with Geiss et al., 2001, where changes in cellular mRNA levels particularly of down-regulated genes were dependent on the replication of IAV³⁹¹. Pairwise analysis of the temporal DEGs indicated that most of the DEGs at early time point infections are also differentially expressed at 10hpi except in pH1N1 where most of the 2 and 4hpi DEGs were distinct from the 8 and 10hpi (Fig. 3.1c-f). Among the 5,563 DEGs, 5,495 (98.7%) of the genes were differentially expressed at 10hpi. As described in Figure 3.2a, the numbers of DEGs at 10hpi were higher in H7N9 and H1N1/WSN virus infections followed by H5N2 virus isolates. In these viruses, the proportions of down-regulated genes were higher than up-regulated genes. In contrast, pH1N1, H5N3 and H9N2 viruses showed lower number of DEGs and up-regulated genes are proportionally higher than the down-regulated genes in the respective viruses. Pairwise analysis of the DEGs at 10hpi in each virus infections indicated that the number of up-regulated genes shared between the avian viruses were higher when compared with avian-human shared DEGs (Fig. 3.2b). However, H1N1/WSN and H7N9 virus shared higher number of DEGs than other viruses (Fig. 3.2b).

Figure 3.1: Kinetics of up- and down-regulated genes in A549 cells infected with four IAVs compared to mock infections.

a) the kinetics of up-regulated genes at 2, 4, 6, 8 and 10hpi; b) the kinetics of down-regulated genes at 2, 4, 6 and 8 hpi of A549 cells infected with the four IAV isolates; c-f) overlap of DEGs at different time points in the four virus infections.

Collectively, 44 DEGs were identified to be common to all of the virus infections at 10hpi and all of them were up-regulated. The extent of their expression across the virus infections is shown in Fig. 3.3. Most of these common genes were the product of innate immunity in response to virus infections.

Figure 3.2: DEGs in A549 cells infections at 10hpi compared to mock-infections.

a) the number of up and down-regulated genes in A549 cells infected with eight IAVs, b) Pairwise analysis of the up-regulated genes (purple) and down-regulated genes (yellow) between the eight IAVs at 10hpi.

3.2.2 Pathways and GO BP of the DEGs

Pathway and GO BP analysis of the DEGs was analyzed using Metascape 1) from the four temporal infections and 2) at 10hpi of all the viruses. First we performed the pathway and GO BP analysis for both up-and down-regulated genes together for each virus infection. Then we did separate pathway and GO BP analysis for the up- and down-regulated genes. The top 20 significantly enriched pathways and GO BPs with q-value < 0.05 were selected for cross-comparison across the virus and time point infections. As expected, the temporal pathway enrichment analysis showed immune responses (defense in response to IAV infections) after 6hpi, except in H5N2/F118 virus which induced the response from 4hpi onwards (Fig. 3.4a). The immune response pathways were enriched mainly by the up-regulated genes. In contrast, genes involved in cellular transcription and translation, cell cycle and metabolic process were enriched

by the down-regulated genes. These pathways were prominent at 10hpi in H1N1/WSN and H5N2/F118 virus infections (Fig. 3.4a).

Figure 3.3: The common DEGs across all IAV infections and their extent of expressions in A549 cells at 10hpi compared to mock-infections.

The top 20 pathways and GO BPs at 10hpi indicated that the avian viruses have similar pathway enrichment patterns except in H7N9 virus infection (Fig. 3.4a). Based on the enrichment analysis, the avian virus H7N9 and H1N1/WSN viruses induced closely related pathways in A549 cells (Fig. 3.4a). Overall, the innate immune response, particularly response to type I IFN was prominent in all IAV infections except in H1N1/WSN which has relatively lower extent of enrichment. Pathways related to apoptosis and cell death are relatively common between the viruses except in pH1N1 virus infection (Fig. 3.4b). Whereas, pathways related to cell cycle, and cellular response to stress and DNA damage are prominent in H1N1/WSN and H7N9 virus infections (Fig. 3.4b). Overall, pathways related to immune response were up-regulated in IAV infected A549 cells. On the other hand, pathways related to metabolic process, cell cycle were down-regulated (Fig. 3.4a and b).

Figure 3.4: The top 20 Pathway and GO BPs of the DEGs in IAV infected A549 cells compared to mock-infections.

a) pathways and GO BPs of the DEGs at different time point infections, b) pathways and GO BPs of the DEGs at 10hpi. The $-\log_{10}(q\text{-value})$ was indicated in the chart where the highly significantly enriched pathways had higher values of $-\log_{10}(q\text{-value})$ (dark brown). Lower $-\log_{10}(q\text{-value})$ (light yellow) indicate less significant, while gray indicates no enrichment.

3.2.3 Integration of the transcriptome, siRNA screens and interactome data set

To identify important host factors for anti-IAV targets and to relate the expression changes into the context of the life cycle of the virus, we integrated the DEGs with host factors from siRNA screens and interactome (Table 2.2, Table 2.3, and Appendix A).

Figure 3.5: Integration of all DEGs at 10hpi in A549 cell infections with interactome and siRNA screens.

a) Venn diagram showing overlap of all the DEGs at any virus and any time point infections in A549 cells with the two other data set b) Cross-comparison the up-regulated pathways in A549 cell virus infections with siRNA and Interactome pathways. c) Cross-comparison the down-regulated pathways in A549 cell virus infections with siRNA and Interactome pathways.

Comprehensive curation of genes from the 11 siRNA screening studies, we identified 1,713 host factors of which 1,561 and 152 were pro-viral (IHF) and antiviral host factors respectively (Appendix A). Similarly, from the 13 IAV interactome studies 1780 IAV interacting host factors were identified (Appendix A). First we integrated the genes from siRNA screens and interactome with the 5,563 DEGs from the eight IAVs (H1N1/WSN, pH1N1, H5N2/F118, H5N2/F59, H5N2/F189, H5N3, H7N9 and H9N2) in A549 cells at different time point infections (Table 2.1, Fig. 3.5a). The result showed that 205 genes were shared between the three data set, while 398 and 541 genes of the DEGs were exclusively shared between the siRNA screens and interactome data set respectively.

To investigate the functional context of these genes from the three data set, we performed KEGG pathway analysis using the up- and down-regulated genes in each virus infections along the KEGG pathways of the siRNA and Interactome (Fig. 3.5b). As expected, all virus infections of A549 cells, the siRNA screens and interactome genes were enriched for influenza A pathways (Fig. 3.5b). Genes associated with TNF signaling and cytokine-cytokine receptor interactions were up-regulated in all the data set particularly prominent in H5N3 virus infection but not in interactome and pH1N1 virus infection (Fig. 3.5b). This suggests that the immune responses upon IAV infections in A549 cells is not only as antiviral response but also to support the virus replication (pro-viral) response. Detailed investigation of up-regulated genes in immune responses indicated that *IRF2*, *PIK3R3*, *JAK2*, *SAMHD1*, *TRIM21* and *TRIM26* were among the highly expressed in pro-viral siRNA screens. Apoptosis and MAPK signaling pathways were also enriched in pro- and anti-viral siRNA screens, however had low enrichment which varies between the virus infections (Fig. 3.5b). On the other hand, spliceosome, RNA transport, ribosome, and protein processing in ER were pathways enriched in pro-viral siRNA screens and interactome genes. These pathways were not or very slightly induced by the virus infections (Fig. 3.5b).

Cross-comparison of the down-regulated pathways (Fig. 3.5c) indicated that H7N9, H5N2/F59, H5N2/F118, and H1N1/WSN virus infections down-regulates pro-viral pathways (e.g. endocytosis, cell cycle, protein processing in ER). There were very minimal changes in other virus infections (Fig. 3.5c).

Figure 3.6: Integration of DEGs in at least 3 IAV infections in A549 cells with host factors identified by interactome and siRNA screens.

a) Pairwise analysis showing the shared genes between the three data set, b) Expression of the selected host factors (up-regulated) shared between the three data set.

Overall, the data demonstrated that induction of the pro-and anti-viral influenza A pathways or innate immune response is common to the IAV isolates examined. Induction or down-regulation of other pro- and anti-viral pathways were IAV virus specific. This suggests that part of the immune responses during IAV infections are in favor of virus replication. The data also suggested that cell cycles, endocytosis, PIK-Akt signaling, and protein processing in ER pathways were also important for efficient virus replication, hence down-regulation of these pathways could be the host-antiviral mechanisms.

To identify conserved DEGs among the viruses, we considered those genes which were differentially expressed in at least 3 viruses. Thus from our analysis we found 711 DEGs expressed in at least 3 viruses. The integration analysis showed that 37 genes were shared between the three data set (Fig. 3.6a), while 99 and 116 DEGs were shared with the siRNA screen and interactome data set respectively (Fig. 3.6a). Among the 37 shared genes, only 1 (*STAT1*) was an antiviral host factor and all the rest were pro-viral genes. Of the 37 shared genes, 12 were up-regulated and 25 were down-regulated in at least 3 virus infections (Fig. 3.6b). Targeting host factors that support the replication of IAV and/or those interacting with the viral proteins has been one of the alternative drug development approaches for IAV. We hypothesise

that among the currently identified host factors, those up-regulated during virus infections could be the plausible host targets. Because considering those down-regulated as therapeutic targets could have a limitation in efficiency as they are already down-regulated in actual infections.

Thus, we focused on the up-regulated genes for further discussions. The 12 up-regulated genes and their extent of gene expression across the eight viruses in A549 cells at 10hpi were described in Fig. 3.6b. These host factors are involved in regulation of innate immune responses (*IRF2*, *SP100*, *TRIM21*, *STAT1*, *TRIM26*, and *SAMHD1*), protein translation (*EIF2AK2*, *RPS15*, *RPS15A*, *FXR2*, and *HEXIM1*) and mitochondrial DNA stability (*TIMM17A*).

SAMHD1, *RPS15*, *RPS15A* and *FXR2* interacts with many of the IAV proteins, while proteins that interact with single viral protein include PB2 (*TRIM21* and *TRIM26*), NS1 (*IRF2* and *SP100*), NP (*HEXIM1* and *EIF2AK2*). The *STAT1* and *TIMM17A* interact with two viral proteins (Fig. 3.7).

Figure 3.7: Interactions of IAV viral proteins with selected host factors.

Green nodes= the up-regulated pro-viral host factors that shared between the three datasets, red nodes = antiviral, light blue nodes = viral proteins, red edges= interaction between the viral and host proteins, purple edge= the interaction among the host proteins.

3.2.4 Integration of the three datasets that used the same cell type (A549 cells)

As experiments in different cell line may vary, we extended our analysis with siRNA screens and interactome studies that were performed in A549 cells. With such criteria, 763 and 407 genes were identified from the 6 siRNA screening studies and 1 interactome study respectively (Table 2.2 Table 2.3, and Appendix A). The integration of these genes with the 711 DEGs showed that only 3 genes (*EGFR*, *COPA* and *DOCKS*) were shared between the three datasets (Fig. 3.8a), and all of the three genes were down-regulated in A549 cell infections. Among the 28 genes exclusively shared between the expression and siRNA screen, 18 pro-viral genes were up-regulated in at least 3 viruses at 10hpi (Fig. 3.8b).

Figure 3.8: Integration analysis of the three data set that were performed using A549 cells.
a) Pairwise analysis of the three datasets from A549 cells, b) The expression profile of up-regulated pro-viral host factors c) The expression of up-regulated host factors that interact with IAV polymerase proteins.

Of the 18 pro-viral genes, four of them (*TRIM21*, *EIF2AK2*, *SP100* and *IRF2*) were also identified in the previous analysis. These four genes were also functionally related with the *FBXW10*, *MX2*, *TRIM14*, *PLALA*, *JAK2*, and *BTC*. Alternatively, among the 42 genes shared between the expression and interactome datasets, 5 were identified to be up-regulated (Fig. 3.8c). These five genes have been demonstrated to have an antiviral effect in influenza and other virus infections.

3.2.5 Integration of the data set that used the same virus strain (H1N1/WSN virus and cell type (A549 cells)

Further, we narrowed down our analysis to the data set that were performed using similar virus strain (H1N1/WSN) and cell type (A549 cells). With this criteria, we identified 671 genes from four siRNA screens. For the transcriptome data, we considered all DEGs of A549 cells during H1N1/WSN virus infections (DEGs in at least one of the four-time point infections (2, 4, 6 and 10hpi)). There was no interactome study that fulfilled this criterion, hence we were restricted to integrate only the expression and the siRNA screen data set. As described in Fig. 3.9a and b, a total of 81 (4 up- and 77 down-regulated) genes were shared between the expression and siRNA screen data set. The expression profile of the 4 up-regulated genes was described in Fig. 3.9b, of which the highly expressed two genes (*FBXW10* and *TRIM21*) were also found in the previous analysis (Fig. 3.8b, Fig. 3.9b).

Figure 3.9: Integration of the three data set performed using the same virus strain (H1N1/WSN) and cell type (A549 cells).

a) Pairwise analysis showing the overlap of the three data set, b) the temporal expression profile of the four host factors shared between expression and siRNA screens.

3.2.6 Drugs targeting the plausible host factors

Finally, we performed drug-host gene interaction analysis using MetaCore for the identified up-regulated IHFs (plausible host factors) and identified 7 FDA approved drugs that had inhibitory activity on the three IHFs. The drugs and their targets are Dabrafenib (*EIF2AK2* and *JAK2*), Ruxolitinib (*JAK2*) and Tofacitinib (*JAK2*); antibiotics (Gentamicin and Tobramycin (*PLA1A*)); Antihypertensive (Propranolol (*PLA1A*)); and antiviral (Acetylcysteine (*JAK2*)). Collectively, the integration of the three data set identified important host factors for IAV drug targets. These identified pro-viral host factors are involved in many cellular processes which include regulation of innate immune responses, ubiquitination, endocytosis, protein translation and metabolism and mitochondrial maintenance. In addition, few of these plausible host factors can also be targeted by FDA approved drugs.

3.3 Discussions

The global host gene expression analysis indicated several isolate specific and common host factors and pathways that could be involved in the lifecycle of IAVs. Generally, the LPAI and human isolates of IAVs up-regulated genes related to immune response and down-regulated genes related to cellular processes and metabolic pathways. The expression and pathway results were generally consistent with previous reports in that higher proportion of DEGs were down-regulated and were related to metabolic pathways and cellular processes while, innate and pro-inflammatory cytokines were induced^{268,392-396}. In addition, isolate/strain dependent host gene expression similarities and differences have also been indicated in previous reports^{268,393}. The expression patterns in HPAI (H5N1 and H7N7 viruses) have been indicated to have increased level of pro-inflammatory cytokines compared to LPAI *in vitro* and *in vivo* studies²⁵³⁻²⁶⁰. Our results showed that LPAI and human isolate shared up-regulation of immune responses and inflammatory cytokines. The result is in line with those of LPAI H7N2, H7N3³⁹³, H5N2²⁶⁸, and H7N9 virus infections that also showed some common and some isolate specific expression changes for human and avian isolates in IAV infections.

The up-regulation of innate and inflammatory immune response has been shown to down-regulate metabolic pathways for example the sterol pathway³⁹⁵ and cellular processes such as gap junction intracellular communication (GJIC)³⁹⁷. This may suggest that cellular processes

and metabolic pathways could be down-regulated in response to the innate and pro-inflammatory responses activated upon IAV infections.

As indicated in the result, several of the host factors either from siRNA or interactome were down-regulated in IAV infections. Down-regulating the majority of the IAV required host factors could be a host cell strategy as a result of an antiviral response³⁹⁸. In addition, general host gene expression inhibition through IAV induced degradation of host polymerase³⁹⁹ and degradation of cellular mRNA by the cap-snatching mechanism of the viral polymerase could also count for the higher proportion of down-regulated genes⁴⁰⁰. The ultimate goal of understanding IAV host interaction is to identify plausible host therapeutic targets. Considering down-regulated IHFs as therapeutic targets could have a limitation in efficiency as they are already down-regulated in actual infections. Hence, inhibition of the up-regulated IHFs would be the ideal strategy.

Our investigation highlighted the up-regulated IHFs that have a role in ubiquitination, immune response, protein translation, endocytosis and phospholipid metabolism. *TRIM21*, *TRIM14*, *TRIM26* and *FBXW10* are involved in ubiquitination processes. Ubiquitination of either viral or host cell proteins could support or suppress IAV replication. For instance, *TRIM32* directly targets and ubiquitinates PB1 protein which in turn leads to proteasomal degradation and reduction of IAV viral replication²⁹⁷. Similarly, mutation of a lysine residue targeted for ubiquitination in PB1-F2 protein increased its expression, the replication activity and antagonises the *IFN* β response in H5N1 and H1N1 viruses⁴⁰¹. In addition, the modification of M1 protein ubiquitination to SUMOylation by aminoacyl-tRNA synthetase complex-interacting multifunctional protein 2 (*AIMP2*) stabilises IAV M1 protein and promotes nuclear export to increase viral replication⁴⁰² suggesting the indirect antiviral role of ubiquitination. Furthermore, *TRIM25* and *RNF135*/Riplet ubiquitin E3 ligases ubiquitinate RIG-I for activation of optimal downstream signalling and interferon response⁴⁰³. While the poly-ubiquitination of *TRIM14* recruits *NFkB* and modulates *MAVS* signalosome for activation of *IRF3* and *NFkB* mediated antiviral response⁴⁰⁴.

Conversely, ubiquitination can induce IAV replication and suppresses the antiviral immune responses. For example, *TRIM21* is an E3 ubiquitin ligase localised predominantly in the cytoplasm and translocated to the nucleus during pro-inflammatory conditions. It negatively regulates interferon responses through ubiquitin-mediated proteasomal degradation of interferon

regulatory factors (*IRF3*, *IRF5*, *IRF7*) and mono-ubiquitination of *IKK β* that resulted in the suppression of *NF κ B* signalling. Similarly, *TRIM21* negatively regulates interferon responses and increased viral titer in Sendai and Japanese encephalitis virus (JEV)^{405,406}. Overexpression of the deubiquitinating enzyme *USP11* repressed IAV genomic RNA replication probably through deubiquitinating the mono-ubiquitinated IAV NP protein⁴⁰⁷. Several studies also showed increased IAV replication through ubiquitination and proteasomal degradation of the antiviral host responses including *p53*⁴⁰⁸, *IFITM 3*⁴⁰⁹⁻⁴¹¹ and host RNA polymerase II (Pol II)⁴¹². We suggest that reduced IAV titer in down-regulation of *TRIM21* could be due to its negative regulatory role in IRFs and IFN response. When *TRIM26* is translocated in to nucleus, it blocks *IFN- β* induction through K48-linked ubiquitination and degradation of *IRF3*⁴¹³. In contrast, the pro-viral *TRIM14* was shown to activate the IAV antiviral response⁴⁰⁴ while the role of *FBXW10* in IAV infection is not clear. Taken together, we suggest that *TRIM21* and *TRIM26* may support IAV replication by negatively regulating IRFs and IFN responses.

Sterile α motif and histidine-aspartate domain-containing protein 1 (*SAMHD1*) is also induced in virus infection and plays a role in the regulation of innate immune response⁴¹⁴. *SAMHD1* exhibits phosphohydrolase and nuclease activity and inhibit HIV-1, Hepatitis B virus, and herpes simplex virus type 1 (HSV1) infections in dendritic cells, monocytes and macrophages by depletion of dNTPs and degradation of HIV-1 RNA template⁴¹⁵⁻⁴¹⁹. However, nuclease activity of *SAMHD1* varies between the non-cycling (dendritic cells, myeloid cells and resting CD4⁺ T cells) and cycling cells (HEK293T)⁴²⁰. In addition, it was suggested that the degradation products of HIV RNA template can be recognised by intercellular PRRs and induce *NF- κ B* signalling. Conversely, *SAMHD1* deficiency leads to accumulation of the DNA species which is internally recognised by cyclic GMP–AMP (cGAMP) synthase (cGAS) that induces IFN induction leading to an immunological disorder called Aicardi–Goutières syndrome (AGS)^{421,422} suggesting IFN response regulatory role of *SAMHD1*. Recently, *SAMHD1* was also found to be a non-classical ISG and was regulated by miR-181a and miR-30a. In primary human monocytes, upon stimulation of IFN type I and II, miR-181a and miR-30a were down-regulated with the inverse up-regulation of *SAMHD1*⁴²³. In a similar manner as *TRIM21* and *TRIM26*, *SAMHD1* might also support replication of IAV through regulation of IFN responses.

The *EIF2AK2*, *SP110* and *MX2* are interferon-inducible genes in response to viral infections. *EIF2AK2* codes for double-stranded (ds) RNA-dependent protein kinase (PKR) a key

antiviral agent through *eIF2α* mediated blockage of viral protein translation (reviewed in^{424,425}). In IAV infection, the interaction of NP viral protein and *Hsp40* results in dissociation and activation of *P58IPK*, consequently, *P58IPK* impairs the antiviral activity of *PKR*^{426,427}. In contrast, *P58IPK* knockout mice in IAV infection resulted in increased mortality and *PKR* activity⁴²⁸, suggesting the role of *PKR* in pathology and inflammation.

SP110 is a member of a multi-complex nuclear body with a role in activation of gene transcription⁴²⁹ and is also induced by IFNs⁴³⁰. Recently, it was reported that one of the isoforms of *SP100* (*SP100A*), a paralogue of *SP110*, amplifies adenoviruses gene expression⁴³¹. Both adenoviruses and IAVs mRNA transcription occur in the nucleus of the host, suggesting that *SP100* may induce expression of IAV mRNA like in adenoviruses. Overexpression of *SP100* is not protective against vesicular stomatitis virus (VSV) and IAV⁴³², in agreement with a pro-viral function. *MX2* is a cytoplasmic GTPase and has an antiviral activity for several viruses but not against IAV⁴³³⁻⁴³⁶. However, the nuclear-localized *MX2* is indicated to have anti-IAV activity⁴³⁷.

IRF2 and *JAK2* may play a role in modulating the immune response in IAV infection. *IRF2* regulates transcription of type I IFN and ISGs in cells through antagonistic role against *IRF1*⁴³⁸ and attenuation of *STAT1* function⁴³⁹. In line with our investigation, *IRF2* was up-regulated in IAV infection⁴³⁹ and its pro-viral activity could be due to its suppression effect of the *IRF1* and *STAT1* induced antiviral responses. *JAK2* activates signal transduction through phosphorylation of tyrosine residues⁴⁴⁰. In IAV infection, phosphorylation of the viral proteins M1 and NP controls the nuclear import of both proteins^{441,442} and *JAK2* inhibitor decreased the tyrosine phosphorylation of M1 and inhibits its nuclear import⁴⁴¹. *JAK2* is also associated with activation of the inflammatory response in infections and its inhibitors were suggested to have a pharmacological advantage for treatment of sepsis⁴⁴³. Hence, we suggest that targeting *JAK2* in IAV infection could produce benefits from inhibition of both viral replication and the cytokine storm.

TIMM17 is a core component of the *TIMM23* complex, a channel for mitochondrial protein import. It is important for regulation of pore structure and voltage gating of the channel. There are two copies of genes that encode for *TIMM17*, the *TIMM17A* and *TIMM17B* and both of them are expressed in human tissues⁴⁴⁴. Overexpression of *TIMM17A* prevents mitochondrial DNA (mtDNA) loss in human and yeast cells. The study suggested that such stabilisation of mtDNA may be due to selective pressure to maintain a functional mitochondrial genome⁴⁴⁵.

TIMM17A overexpression was also associated with adverse pathological and clinical outcomes in breast cancer⁴⁴⁶. In contrast, decreased the level of *TIMM17* resulted in impaired biogenesis of mitochondrial proteins and activates stress-induced signaling pathways⁴⁴⁴. Hence, overexpression of *TIMM17A* could induce IAV replication through stabilization of mitochondrial biogenesis and prevention of apoptosis. *TIMM17A* interacts with M1 and M2 proteins of IAV proteins (Fig. 3.7). Although the M2 viral protein have been demonstrated to have toxicity effect on mammalian cells through mitochondrial proton channel activity, the relation between the M1 and M2 proteins and *TIMM17A* is not known⁴⁴⁷.

Hexamethylene bisacetamide (HMBA) inducible protein 1 (*HEXIM1*) is a regulator of the positive transcription elongation factor b (*p-TEFb*), which is the master modulator of RNA polymerase II during transcriptional elongation⁴⁴⁸. The activity of HMBA which is a potent inducer of HIV production was reported to be through phosphorylation of *HEXIM1*. The phosphorylated *HEXIM1* subsequently releases active (*p-TEFb*). Then, HIV promoter is activated by p-TEFb for transcription elongation and viral production⁴⁴⁹. Betacellulin (*BTC*) is one of epidermal growth factor receptor (*EGFR*) ligands, which are synthesised primarily as a transmembrane precursor and then mature through proteolytic cleavage (reviewed in⁴⁵⁰). It activates *EGFR*⁴⁵¹ and activation of *EGFR* promote IAV uptake⁴⁵², however, the role of betacellulin in facilitating the uptake of IAV is not known.

Phospholipase A1 (*PLA1A*) is an enzyme that hydrolyzes phospholipids into fatty acids and lysophospholipid⁴⁵³. The corn-shaped lysophospholipids hydrolyzed by *PLA2* alter plasma membrane structure and play a role in IAV membrane raft budding⁴⁵⁴. The fatty acids are important for palmitoylation of many cellular proteins and could also help in palmitoylation (sialylated) at HA and M2 proteins of IAV⁴⁵⁵. IAV HA protein can be palmitoylated on cysteine residues at the cytoplasmic domain (211, 218 and 21) and the transmembrane region (TMR) (C540 and C544). Removing the palmitoylated fatty acid or mutation of the cysteine residues of IAV HA protein particularly at the cytoplasmic tail severely affected the membrane fusion, replication, number of NP and M1 protein internal and budding of IAV⁴⁵⁶⁻⁴⁵⁹. Palmitoylation depends on availability and concentrations of the fatty acids (reviewed in⁴⁶⁰), and whether *PLA1A* is involved in the regulation of HA protein acylation is not clear.

Among the up-regulated IAV interacting host factors; *ZC3HAV1*, *STAT1* and *MX1* are involved in the antiviral immune response. *ZC3HAV1* also called *ZAP* is induced by 5'

triphosphate modified RNA (3pRNA) and stimulates the interaction of *RIGI* with *MAVS* for induction of antiviral interferon responses through activation of *IRF3* and *NFkB* transcription factors⁴⁶¹. The interferon responses induced by IAV infection activates *STAT1*³⁴⁵ and subsequently increases the expression of ISGs^{345,462}. It plays role in both innate and adaptive immune response in IAV infections (reviewed in⁴⁶³). *MX1* is one of the antiviral ISGs induced by type I and type III interferons⁴⁶⁴. In A549 cells, *ZC3HAV1* interacts with PA and PA-PB1 viral proteins, *STAT1* with PA and *MX1* binds to the polymerase complex³⁵⁵. The binding of *ZC3HAV1* with PA-PB1 (or polymerase complex) has been suggested to counteract the antiviral immune response through inhibition of *RIGI* signalling³⁵⁵. Increased expression of *ZC3HAV1* in IAV infected chicken lung has been correlated with survivability⁴⁶⁵.

PARP14 and *RNF213* interact with PA of IAV in A549 cells³⁵⁵. *PARP14* is involved in cell death pathways and it has been shown to be activated by IAV infection in human Alveolar Type II (ATII) cells²³². *RNF213* is a probable E3 ubiquitin-protein ligase which is involved in angiogenesis and ATPase activity⁴⁶⁶ and its role in IAV infection are not clear. On the other hand, *MX1* blocks virus transcription and replication through inhibition of PB2-NP proteins interaction and interfering with virus assembly⁴⁶⁴. Overall, the up-regulated IAV interacting host factors have antiviral activity, and IAV uses such interaction as an evasion of host antiviral responses. Hence, the activity of these host factors can be maintained by blocking the interaction of viral polymerase proteins with these host factors.

Other known antiviral host factors such as interferon induced transmembrane protein (*IFITM*) family^{287,467} and 2'-5'-oligoadenylate synthetase 1(*OAS1*)⁴⁶⁸ were up-regulated in the present study. *IFITM1* and *OAS1* were up-regulated in all of the isolates at 10hpi and after and/or at 6hpi in the four isolates. *IFITM2* was only up-regulated at 10hpi in H5N2/F59 virus while *IFITM3* was up-regulated in pH1N1, H5N2/F118, and H5N2/F59 viruses at 10hpi. None of these host factors were indicated in interactome studies. However, *IFITM3* and *IFITM2* were reported as antiviral host factors in siRNA screens^{287,351}. *IFITM3* and *OAS1* restrict IAV infections through inhibition of fusion pore formation at a post-hemifusion stage in late endosomes⁴⁶⁹ and activation of viral RNA degradation⁴⁶⁸ respectively.

Some of the up-regulated IHFs may be targeted through repurposing FDA approved drugs although the high price and side effects of some of them would not make them suitable for treating influenza. The antineoplastic agents are kinase inhibitors and recently it has been

indicated that the kinase inhibitors like dinaciclib and flavopiridol, showed potent antiviral activity against IAV⁴⁷⁰. Similarly, the antihypertensive Propranolol⁴⁷¹ and the antiviral acetylcysteine showed therapeutic and anti-IAV activity⁴⁷². These support the reliability of our integration approach for identification of plausible therapeutic targets.

3.4 Conclusions

The integration of the three data set highlighted plausible IHFs that could be targeted against IAVs. The up-regulated IHFs (*TRIM21*, *TRIM26*, *IRF2*, and *SAMHD1*) might support the replication of IAV through suppression of the antiviral responses. The other up-regulated IHFs could also enhance the replication of IAV at different stages of the virus life cycle: endocytosis (*BTC*), prevention of apoptosis (*TIMM17A*), nuclear import (*JAK2*), and translation elongation of viral proteins (*HEXIMI*). However, the detailed mechanisms for several of these up-regulated IHFs need to be investigated further. This can be achieved by knocking in/out of the identified plausible host factors in A549 cells, primary cell lines and also in animal models (e.g. mice) and measuring the level of the viruses through different virus assays. In addition, new inhibitory molecules against up-regulated IHFs, and drugs that interfere with the interaction of viral proteins with the up-regulated antiviral host factors would provide new alternative anti-IAV therapies.

Chapter 4. Comparison of gene expression signatures during IAV infections of different species cell lines

4.1 Introduction

Influenza A virus (IAV) is a zoonotic virus that is circulating in various animals (e.g. avian, pigs)⁴⁷³. These animal viruses play an important role in the evolution of influenza viruses that circulate with the human population. However, when viruses transmit from avian to a human host several barriers must be overcome for efficient infections. Furthermore, once the virus has transmitted to the new host the maintenance of the virus in the new host requires the ability to overcome several other obstacles⁴⁷³. In essence, the virus must adapt to overcome these barriers to infection. While the process is well understood, there have been extensive studies to understand the bases for the host adaptation that allow maintenance of the avian IAVs in mammalian hosts⁴⁷⁴. High mutation rates and reassortment of IAV gene segments are, for example, some of the major drivers of host adaptation⁴⁷⁵.

Several IAV proteins have been implicated in host adaptation. Avian virus (H5N1) polymerase proteins (PB1, PB2 and PA) demonstrated to have poor activity in mammalian hosts (e.g human). However, upon acquisition of specific mutations in these viral polymerase proteins considerably increased the polymerase activity and mammalian host adaptation^{181,476}. The adaptation and tissue tropism of IAV also depends on host cell receptor binding of viral HA protein. Human influenza viruses primarily infect ciliated cells that express $\alpha(2, 6)$ -linked SA receptor, while the avian viruses mainly infect non-ciliated cells with $\alpha(2, 3)$ -linked SA receptor expression⁴⁷⁷. The HA protein mutations can allow avian viruses to preferentially recognize human SA receptors¹⁸⁷.

Although several studies have examined the genetic determinants of IAV adaptation, the host response elicited by different viruses is not well understood. In particular, the specific effects that virus infection has on general cell physiology would be expected to be an important factor in enabling a virus to replicate in a new host. Although animal models of infection can be used to examine pathogen host interactions, the viruses used need to be adapted to animal models⁴⁷⁸. This process of adaptation can change fundamental properties of the virus isolates,

and in general these changes are multi-factorial and poorly defined. In addition, if a virus is not able to replicate efficiently in the animal model this may also impact on other biological properties of the virus that requires efficient virus replication such as host response. In this context cell types are normally permissive to influenza virus replication represent a convenient model system to examine influenza virus replication, since virus infection can be accurately controlled and the virus-specific molecular and cellular changes that occur early in infection can be analyzed. Furthermore, these cells can retain complete signaling networks that are related to the innate host response to infection [e.g. IFN], and it is also expected that these signaling networks retain properties of the species they are derived from.

Our group previously described several LPAI viruses that were isolated from duck imported to Singapore. These viruses included H5N2, H5N3 and H9N2 virus subtypes³⁴⁷. The replication characteristics of these viruses in A549, CEF and MDCK cells were examined in detail²⁶⁹. Interestingly, the avian viruses exhibited a cell-specific capacity to replicate in cell types. We hypothesised that the capacity of specific virus isolates to increase or repress the expression of specific host genes may determine the capacity of viruses to replicate in mammalian and avian cell types, and by extrapolation, different animal hosts. Our group had previously examined the host response to infection with these viruses²⁶⁹, but, due to limitations in the annotation of chicken and canine Affymetrix probes it was not previously possible to perform a detailed side-by-side comparison of the host responses in CEF and MDCK cells with the human A549 cells. In this current study we extended our previous analysis by applying bioinformatics tools to annotate and use ortholog mapping between the different species. In this way we have been able to compare the expression signatures in different cell types of different species. This approach is able to highlight the key expression signature differences between different species during IAV infections and gives insight for the host-adaptation mechanism of IAVs.

4.2 Results

4.2.1 Annotation and orthologs mapping

In this analysis we examined six IAVs (H1N1/WSN, H5N2/F118, H5N2/F59, H5N2/F189, H5N3 and H9N2 viruses) infections in A549, MDCK and CEF cells which were derived from human, canine and avian sources respectively. The rational for using these cell types was that these are permissive cell types which are extensively used to examine influenza

virus infection. In addition, MDCK cells are commonly used to propagate clinical strains of influenza virus. In this context the GeneChip Canine Genome 2.0 Array (Affymetrix, USA), the GeneChip Chicken Genome Array (Affymetrix, USA) and Genechip Human Genome HG U133 Plus 2.0 Array (Affymetrix, USA) were used in the microarray analysis. However, the interpretation of microarray data into biological context requires accurate mapping of the probe sets on the array to the genes that they are targeting⁴⁷⁹. Affymetrix Gene Chips are one of the leading array research platforms for functional genomics⁴⁸⁰. However, the annotations of probe sets to genes are not complete, and 50% and 40% of the chicken and canine probesets were not annotated⁴⁸⁰. Incomplete annotation of microarray probes sets can restrict and limit the interpretation of expression results as important differentially expressed probes can be missed in the process of functional annotation or pathway analysis. In this study we mapped the non-annotated probe sets to Ensembl reference gene sets of chicken and canine genomes as basis for probe annotation.

Our investigation observed that 41.6% (16,055/38,535) and 30.6% (13,195/43,035) of the chicken genome and canine genome Affymetrix arrays were not annotated with ensemble IDs (Table 1). Using ncbi-blast+ package we were able to annotate 99.52% (15,979/16,055) and 99.7% (13,162/13,195) of the non-annotated probes with ensemble IDs in chicken and canine genome arrays respectively (Table 1). The annotated probes represent 7412 chicken and 6869 canine genes. Although 84% (6,229/7,412) and 85.9% (5,899/6,869) of the annotated chicken and canine genes respectively were represented by other Affymetrix annotated probes, still nearly 52.9% (7,412/14,007) chicken and 38.5% (6,869/17,813) canine genes would have been missed if the probes were not annotated. In addition, we used annotations from Ensembl genes 83, NCBI and DAVID for annotation of Ensembl IDs/probes to official gene symbols. Furthermore, with the support of orthologs mapping (see below), we increased the annotation of chicken and canine genome array from 58.4% to 99.8 % (38,474/38,535) and 69.4% to 99.9% (43,006/43,035), respectively (Table 4.1 and Appendix C).

Combining gene expression results from multiple species of a specific condition (e.g. disease) can lead to additional findings related to cross-species conservation or host-specificity that cannot be seen in a single species result alone⁴⁸¹. However, cross-species comparison of microarray data is complicated because several factors including platform variations, probe

quality, laboratory effects, dynamic environment, genetic background and organism annotation status⁴⁸¹. Similar tissues have significant conserved expression patterns across species⁴⁸².

Table 4.1: Annotation of affymetrix probe sets with no ensembl IDs

Annotation	Chicken Genome	Canine Genome 2.0
	Array	Array
Total number of affymetrix probes	38535	43035
Total number of genes (ensembl IDs)	12639	16602
Probes without ensembl ID annotation	16055	13195
Probes annotated in this study (new)	15979	13162
Genes corresponding to the new annotation	7412	6869
New annotated genes represented by other affymetrix probes	6229	5899
New annotated genes unique to this study	1183	970
Total number of fully annotated probes	38474	43006
Total number of fully annotated genes	14007	17813

N.B. The total number genes (from ensembl genes 83) for Chicken and Canine are 17108 and 24580 respectively.

Comparison of host-gene expression signatures between species is important to evaluate the host-specific changes during IAV infections. Side-by-side comparison of single genes expression in different species could be possible by mapping orthologous genes of the species. This approach could also evaluate whether the experiment in one species is representative of the others. To study these effects, we performed a comprehensive human ortholog mapping of chicken and canine genes to compare microarray results from infections in the three cell types under identical experimental condition and with the same influenza virus isolates. As indicated in Table 4.2, only 35-37% of the chicken and canine genome Affymetrix arrays were mapped as human orthologs (human genome HG U133 Plus 2.0 array).

To maximize the coverage, we analyzed human orthologs of chicken and canine genes using OMA. Overall, 11,674 and 16,002 chicken and canine genes respectively were orthologs to human genes (Table 4.2, Appendix C). From OMA orthologs mapping in chicken, 11,220 of the 11,674 genes were common to the fully annotated probes-ensembl genes and represent 80% of

the array genes. Similarly, 15,284 of the 16,002 human orthologs of the canine genes were in the fully annotated sets representing 85.8% of the canine genome array. After mapping probes to ensembl IDs and vice versa, the combination of orthologs from OMA and Affymetrix covered 89.1% and 91.6% of the chicken and canine genome array probes, respectively (Table 4.2, Appendix C).

4.2.2 Comparison of DEGs in A549 with MDCK and CEF cells

The microarray data was analyzed using GEO2R (limma, R-package) and the comparison of differentially expressed genes (DEGs) in chicken (CEF) and canine (MDCK) relative to their human orthologs (A549) was described in Tables 4.3 and 4.4. Generally, the expression changes were mainly observed after 6hpi. The number of DEGs were significantly higher in MDCK than A549 cells, particularly at 10hpi. This is similar to the number of down-regulated genes except in H5N3 virus infections; however the reverse is true for the up-regulated genes (Table 4.3). Unlike A549 and MDCK cells, the numbers of both up and down-regulated genes were significantly higher in A549 cells compared to CEF with exception of H5N2/F118 virus infections (Table 4.4). Taken together, the data suggested the expression changes in IAV infection were in the order of MDCK > A549 > CEF with exceptions of up-regulated genes in A549 vs MDCK and the H5N2/F118 virus infection in A549 vs CEF cells.

Table 4.2: Human ortholog mapping of *Canis familiaris* (Canine) and *Gallus gallus* (Chicken) genes/probes

Orthologous mapping	Genes/probes	Chicken orthologs	Canine orthologs
Affymetrix orthologs	Probes	14415 (37.04)	15053 (34.97)
	Genes	8654 (61.78)	9109 (51.13)
OMA orthologs	Genes	11674 (83.34)	16002 (89.83)
OMA orthologs common to the full annotation (ensemble ID)	Genes	11220 (80)	15284 (85.8)
Mapping OMA orthologs to Affymetrix probes	Genes	11175 (79.78)	15552 (87.30)
	Probes	32244 (83.67)	38755 (90.05)
Total Orthologs (OMA + Affymetrix)	Genes	11788 (84.15)	15577 (87.44)
	Probes	34335 (89.10)	39436 (91.63)

Table 4.3: Comparison of ortholog DEGs in IAV infected A549 and MDCK cells

Virus isolates	Post infection	A549	MDCK	Common	p-value
All DEGs					
In all isolates	All time points	5,684	13,737	5,301	<0.001
H1N1/WSN	6h	2	135	0	1*
	10h	4,173	10,683	3,619	<0.001
H9N2	6h	0	2	0	
	8h	19	1	0	1*
	10h	70	11	2	0.001*
H5N2/F118	4h	1	0	0	
	6h	33	0	0	
	8h	83	17	1	0.087*
	10h	93	6,106	40	0.45
H5N2/F59	10h	2,269	2,664	729	<0.001
H5N2/F189	10h	974	2,194	281	<0.001
H5N3	10h	281	9,691	192	0.033
Up-regulated genes					
In all isolates	All time points	1,282	9,765	871	<0.001
H1N1/WSN	6h	2	93	0	1*
	10h	145	79	11	<0.001
H9N2	6h	0	2	0	
	8h	18	1	0	1*
	10h	61	8	2	<0.001
H5N2/F118	4h	1	0	0	
	6h	30	0	0	
	8h	54	3	1	0.01
	10h	40	131	2	0.044*
H5N2/F59	10h	957	83	35	<0.001
H5N2/F189	10h	459	31	6	<0.001
H5N3	10h	189	9,664	136	<0.001
Down-regulated genes					
In all isolates	All time points	4,675	11,762	4,220	<0.001
H1N1/WSN	6h	0	42	0	
	10h	4,028	10,604	3,532	<0.001
H9N2	6h	0	0	0	
	8h	2	0	0	
	10h	9	3	0	1*
H5N2/F118	4h	0	0	0	
	6h	3	0	0	
	8h	29	14	0	1*
	10h	53	5,975	28	0.03
H5N2/F59	10h	1,312	2,581	587	<0.001
H5N2/F189	10h	515	2,163	213	<0.001
H5N3	10h	92	27	5	<0.001

* Fisher's exact test

Table 4.4: Comparison of ortholog DEGs in IAV infected A549 and CEF cells

Virus isolates	Post infection	A549	CEF	Common	p-value
All DEGs					
In all isolates	All time points	4,743	3,716	2,099	<0.001
H1N1/WSN	4h	0	1	0	
	6h	1	14	0	1*
	10h	3,541	45	13	0.866
H9N2	8h	11	0	0	
	10h	42	0	0	
H5N2/F118	6h	18	0	0	
	8h	51	3,645	18	0.498
	10h	65	84	0	1*
H5N2/F59	10h	1,894	19	7	0.014
H5N2/F189	10h	786	5	1	0.292*
H5N3	10h	219	18	1	0.287*
Up-regulated					
In all isolates	All time points	980	159	32	<0.001
H1N1/WSN	6h	1	8	0	
	10h	91	42	7	<0.001
H9N2	8h	12	0	0	
	10h	31	0	0	
H5N2/F118	6h	18	0	0	
	8h	27	103	2	0.023*
	10h	24	11	0	1*
H5N2/F59	10h	756	13	5	0.001*
H5N2/F189	10h	328	5	1	0.132*
H5N3	10h	136	18	1	0.189*
Down-regulated					
In all isolates	All time points	3,983	3,573	1,870	<0.001
H1N1/WSN	4h	0	1	0	
	6h	0	6	0	
	10h	3,450	3	1	1*
H9N2	8h	1	0	0	
	10h	11	0	0	
H5N2/F118	6h	0	0	0	
	8h	24	3,542	9	0.425
	10h	41	73	0	1*
H5N2/F59	10h	1,138	6	2	0.107*
H5N2/F189	10h	458	0	0	
H5N3	10h	83	0	0	

* Fisher's exact test

4.2.3 Functional annotation of DEGs at 10hpi

To compare the expression changes of the three cells in side-by-side, we mapped the chicken and canine genes/probes together using human gene/probe orthologs as common basis (Appendix C). As many of the expression changes occurred at 10hpi, the DEGs at this time point of infections were used for further functional annotation and pathway enrichment analysis. A total of 10845 genes were differentially expressed in at least one of the events at this time point. The functional annotations of the DEGs in each cell type and virus infections were described in Fig 4.1. In the figure, the suffix labelling A, C and M indicates the cell type A549, CEF and MDCK respectively, (e.g. A_10h H1N1/WSN, A549 cells infection with H1N1/WSN virus at 10hpi).

Generally, the functional annotations of the DEGs were related to cell cycle, cellular organization, metabolic processes and immune response pathways (Fig. 4.1a). Detail pathway enrichment analysis of the up-and down-regulated genes were separately given in Fig. 4.1b and Fig. 4.1c respectively. The up-regulated genes were associated with immune response (interferon and inflammatory) responses, intracellular signaling and cellular transport (Fig. 4.1ab). The interferon and inflammatory responses were generally up-regulated in A549 cells infected with all viruses and also for H1N1/WSN virus in CEF cells and H5N2/F59 virus in MDCK cells (Fig. 4.1b).

On the other hand, IAV infection down-regulates genes associated with cell cycle and nucleic acid metabolism with strongest effect in MDCK cells, followed by A549 cells (with the H1N1/WSN, H5N2/F189, H5N2/F59 viruses) (Fig. 4.1c). No such characteristics was observed in all infections of CEF cells, A549 (the H5N3, H5N2/F118, H9N2 viruses) and MDCK (the H5N3 virus). Overall, the data demonstrated that IAV infections in A549 cells mainly induce interferon and inflammatory pathways, while virus infections also induce cell cycle pathway arrest mainly in MDCK cells and variably in A549 cells in a strain dependent manner. Conversely, relatively very low interferon and no inflammatory responses were observed in CEF infections,, pathways related to cell cycle were not arrested in CEF cells.

Figure 4.1: Functional annotation of the DEGs in A549, CEF and MDCK cell lines during IAV infections at 10hpi compared to mock infections.

a) Functional annotation of both up-and down-regulated genes together, b) functional annotation of up-regulated genes, c) functional annotation of down-regulated genes.

4.2.4 Correlation and clustering of DEGs at 10hpi

Correlation analysis of the DEGs at 10hpi (Fig. 4.2a) indicated that the expression correlation of different viruses in one cell type is higher compared to the expression correlation of a single virus in different cell types. For instance, the expression of DEGs in MDCK cells following H1N1/WSN virus infection at 10hpi (M_H1N1.WSN) was correlated better with the other viruses infecting the same cell type (M_H5N2.F118, M_H5N2.F59, M_H5N2.F189 and M_H5N3 viruses) compared to the same virus infecting A549 cells (A_H1N.WSN) and CEF cells (C_H1N1.WSN) (Fig. 4.2a). This phenomenon is also similar for other viruses (Fig. 4.2a). In addition, the principal component analysis (PCA) clustered the expression patterns at 10hpi into three groups (Fig. 4.2b) that clustered mainly towards the cell type rather than IAV strains.

Figure 4.2: Correlation and clustering of the DEGs.

a) Correlation matrix of the DEGs in the three cell lines across the IAV infections at 10hpi. b) the principal component analysis of the DEGs to identify the determinant genes in each cluster (species).

DEGs associated with expression in CEF cells (PCA1) include *OSBPL1A*, *MRPS27*, *RPRD1B*, *ARHGAP21* and *COPS2*, while *CCL5* expression was associated with MDCK cells (Fig. 4.2b). Interferon stimulatory genes (ISGs) expressions are mainly associated with A549 (e.g. *RSAD2*, *OAS1*, *USP18*, *XAF1*, *MAX1*, *SAMD9L*) followed by CEF cells. The hierarchical clustering of the top 56 highly expressed genes (using Gene cluster 3.0) is in accordance with the PCA analysis that expression patterns of the same cells were clustered together except in M_H5N3 virus (in MDCK) and A_H1N1/WSN virus (in A549 cells) at 10hpi (Fig. 4.3a).

Figure 4.3: The hierarchical clustering of the top 56 DEGs across the cell lines and virus infections compared to mock infections.

a) the hierarchical clustering of the top 56 DEGs at 10hpi b) hierarchical clustering of the top 56 DEGs during temporal infections in three viruses (H1N1/WSN, H5N2/F118 and H9N2).

The temporal expression of these genes in the three of the viruses (H1N1/WSN, H5N2/F118 and H9N2) in the three cell types indicated few changes before 10hpi. Some of the ISGs were up-regulated at 6 and 8hpi during the avian virus infection in A549 cells. In MDCK cells, H1N1/WSN virus induces *RSAD2* and *OAS1*, while H9N2 virus induces *CCL5* from 6hpi onwards (Fig. 4.3b). Taken together, this data suggested different species cellular responses in IAV infections.

4.2.5 Integration of the DEGs with interactome and siRNA screens

For the better understanding of the DEGs in the context of IAV replication, we integrated the orthologs DEGs at 10hpi with siRNA and interactome data set (Fig. 4.5a). The result revealed that 271 pro-viral and 16 anti-viral siRNA screen genes were shared with DEGs and interactome data set. Cross-comparison of the KEGG pathways in siRNA screens and interactome across the up-regulated pathways in the virus infections at 10hpi showed that influenza A pathway was mainly enriched in A549 cell infections particularly in H5N3, but minor changes in MDCK and CEF cells infections. In addition, in MDCK cells the ribosome pathway was highly induced during H5N3 virus infection. In contrast, several of the pro-viral pathways were down-regulated in MDCK (H1N1/WSN and H5N2 virus infections) and A549 (H1N1/WSN and H5N2/F59 virus infections) cells. Overall, the data demonstrated that activation or suppression of pro-or anti-viral pathways were cell type and virus strain specific.

Among the top 56 highly expressed ortholog genes (Fig. 4.3a and b), five (*MRPS27*, *MX2*, *PLAU*, *MTHFD2*, *DAAMI*) and three (*TMEM140*, *KLHL21*, *MYH9*) were pro-viral and anti-viral genes respectively. Among which *MRPS27* specifically up-regulated in CEF cells, while *MX2* and *TMEM140* were up-regulated mainly in A549 cell infections.

4.2.6 DEGs associated with metabolic pathways

Detailed investigation of the DEGs involved in metabolic pathways identified 665 genes differentially expressed at least once across the three cell types and the virus infections at 10hpi. The hierarchical clustering and heat map of the 665 DEGs was described in Figure 4.5a. As indicated in the figure, most of the metabolic pathways were up-regulated during H5N3 virus infections in MDCK cells. Conversely, many of the genes were down-regulated in MDCK cells (H1N1/WSN and H5N2 virus infections), and A549 cells (H1N1/WSN and H5N2/F59 virus infections).

Figure 4.4: Integration of the DEGs in the three cell types during the IAV infections at 10hpi with interactome and siRNA screens data set.

a) Venn diagram showing the pro-and anti-viral genes overlapping with the DEGs and interactome. b) Top 20 shared and specific pathways of the siRNA screens and interactome along with the up-regulated DEGs. c) Top 20 shared and specific pathways of the siRNA screens and interactome along with the down-regulated DEGs.

Almost no effect on metabolic genes were observed in CEF cells infected with all of the tested viruses. Similarly, H9N2 virus was silent in changing the metabolic state of the three viruses.

For pathway enrichment analysis, we identified the up- and down-regulated genes in each cell and virus infections. The KEGG pathway analysis of the up-regulated metabolic genes indicated that energy (oxidative phosphorylation), lipid, amino acid, nucleic acid, glycan and vitamin metabolic pathways were enriched in H5N3 virus infections in MDCK cells at 10hpi (Fig. 4.5b). No induction of the metabolic pathways were found in MDCK cells in other virus infections. To the reverse, nucleic acid, amino acid, lipid, carbohydrate and glycan metabolic pathways were down-regulated in MDCK cells (in H1N1/WSN, and H5N2 infections) and A549 cells (H1N1/WSN and H5N2/F59 virus infections) (Fig. 4.5c).

Figure 4.5: DEGs and pathways involved in metabolic pathways.

a) Hierarchical clustering and heat map of the DEGs involved in metabolic pathways b) top 20 up-regulated metabolic pathways c) top 20 down-regulated metabolic pathways.

To evaluate whether these metabolic pathways were in support or against IAV replications, we identified 194 genes that were involved in metabolic pathways from the pro-viral siRNA screen genes.

Figure 4.6: Comparison of metabolic pathways between the pro-viral siRNA screens and DEGs.

a) Top 20 pro-viral metabolic pathways (from siRNA screens) across the up-regulated metabolic pathways across the cells and the virus infections. b) Top 20 pro-viral metabolic pathways (from the siRNA screens) across the down-regulated metabolic pathways across the cells and the virus infections.

Then we compared the metabolic pathways from the pro-viral siRNA screen with the up- and down-regulated pathways in the virus infections. The result in fact showed that oxidative phosphorylation and other metabolic pathways were important for the replication of IAVs. Parallel comparison of the replication important metabolic pathways (from pro-viral siRNA screens), with the up-regulated genes indicated that H5N3 induced several of the IAV replication important metabolic pathways in MDCK cells. In addition, purine metabolism was also induced by H5N2/F59, H5N2/F118 and H5N3 virus infections in A549 cells (Fig 4.6a). To the opposite, these IAV replication important pathways were down-regulated in MDCK cells (infected with H1N1/WSN and H5N2 viruses) and A549 cells (infected with H1N1/WSN and H5N2/F59 viruses) (Fig. 4.6b). Very few down-regulated metabolic pathways were observed in A549 cells infected with H5N2/F189 virus. Overall, the data demonstrated that metabolic pathways are important for the replication of IAV replications. While H5N3 virus induces many metabolic pathways in MDCK cells, other viruses down-regulate these pathways. The induction or down-regulation of the metabolic pathways in A549 cells were virus specific, while in CEF cells metabolic pathways were not affected by IAV infections. This suggested that induction or maintaining the metabolic pathways of the cells during IAV infections could be important for replication of IAVs.

4.3 Discussion

The CEF, MDCK and A549 cells represent the most common cell types that are used to examine influenza virus infection. Although these cell types would be expected to exhibit transition properties that are distinct from the tissue they are derived, it would be expected that signaling networks that are present within these cells types would retain species-specific properties (e.g. how they are regulated). This therefore allows us to examine the host response to AIV infection and forms the basis of the study we have outlined in this manuscript. It is expected that avian viruses would favor cells that contain an avian background, and as a consequence different pathways are regulated compared to that in human cells. There are several factors that can influence the capacity of avian and mammalian virus to replicate in different cell types. One current paradigm is that species adaptation involves the interactions of multiple virus proteins with specific cellular proteins. These interactions induce changes in the expression and activity of a myriad of other cellular pathways that play a diverse role from the host response to

infection, to virus-induced changes in cell metabolism. At the very basic level these virus-induced changes could determine the capacity of the virus to productively replicate in these cell types. If such changes in gene expression occur in tissues, then these would be expected to occur in a species-specific manner and would be expected to have a major impact on virus transmission and pathogenicity. However, although these host cell proteins perform similar activities in these cells it is likely that the sequences of the same proteins in the avian or human background will differ. This would be expected to influence the interaction of the virus proteins, and in turn these differences may lead to fundamental changes in the host expression profile.

A major limitation of comparing how avian influenza viruses affect host gene expression is the functional annotation of the genes in chicken and canine genome which are less well annotated than those from human. This therefore makes a direct comparison of changes in gene expression difficult. In this study we were able to extend the annotation in the Affymetrix probe sets in chicken and canine to give a more complete understanding of the changes in gene expression that result from LPAI virus infection. This therefore allowed us to compare the expression profile of homologous proteins in these different cell types, through mapping of their genes to human orthologs.

Generally, the number of orthologous genes whose expression changed when infected with IAV was in the order of MDCK cells > A549 cells > CEF cells. Comparison of human and canine orthologous gene expression showed similarity in normal conditions⁴⁸³. Similarly, global transcription patterns of most orthologous genes were conserved between mouse and human immune cells though hundreds of genes showed divergent expressions across the cell lineages⁴⁸⁴. The ideal way of investigation would be using similar experimental setup (viruses, platform, dose of infection, laboratory setup as we used in this study) and the same cell type (tissue) of different species. Although we were comparing different cell types of different species (as a limitation of this study), recently it has been shown that expression patterns of many genes in human and mouse are more similar in different tissues of the same organism compared to the same tissues of different species suggesting the underlying differences between the two species⁴⁸⁵. Hence, more or less the cells used in this study would represent the individual species.

In A549 cells, type I IFN (*IFNB1*) was up-regulated in all of the virus infections at 10hpi except in H5N3 and pH1N1 virus infections. Type III IFN (*IL28A* and *IL29*) were highly expressed except in H5N3 virus infections as *IFNB1*, but with higher extent (Appendix D). In

CEF cells only H1N1/WSN and H5N2/F59 viruses were able to induce *IFNB1*. In contrast to A549 and CEF cells infections, H5N3 virus induced expression of *IFNA* and *IFNB1* in MDCK cells at 10hpi. Collectively, the avian viruses induces higher to moderate level of *IFN* responses and TNF signaling (inflammatory response) in A549 cells. The avian viruses were unable to induce the antiviral immune responses in CEF and MDCK cell infections, but H5N2/F59 and H5N3 viruses showed very low antiviral response in CEF and MDCK cells respectively. These observations showed that antiviral responses (e.g. *IFN* γ and *IFN* β responses) were prominent when a virus adapted to one species infected other species. This condition is not consistent in MDCK cells, however, it has been demonstrated that MDCK cells possess both $\alpha(2, 3)$ -linked SA and $\alpha(2, 6)$ -linked SA on their surface making them suitable for susceptible and high yield of the avian and human IAV strain⁴⁸⁶. In both human and avian IAV strains, MDCK cells show low levels of antivirus responses. Our data give an additional explanation that this phenomenon is likely to be an important aspect that allows recovery and propagation of viruses in these cells.

In line with our study, both LPAI H5N2 and HPAI H5N1 viruses induce significantly very low level of innate immune responses in chicken lung and CEF cells compared to duck. It was suggested that this difference was associated with the absence of RIG-I genes in chicken or CEF cell lines. The CEF cells that express RIG-I gene after transfection induced increased level of innate immune responses in H5N1 virus infections⁴⁸⁷. However, recently chicken lung tissue microarray showed increased level of type I IFN, cytokines and ISGs in HPAI H5N1 virus infections, but not in LPAI H9N2 virus⁴⁸⁸. This suggested that there might be alternative ways independent of RIG-I (e.g. Stimulator of IFN genes (STING)) which could induce excessive innate immune response in H5N1 virus infected chicken⁴⁸⁹.

In contrast to our investigation, A549 cells infected with low LPAI H5N2 exhibited very low expression changes²⁶⁸. Such difference could be due to very low level of infections (MOI=1), and time point infections (24hpi) used in their study. The same study demonstrated that HPAI H5N1 and H7N1 virus infections induce up-regulation of several ISGs in A549 cells²⁶⁸.

In this investigation, we identified highly expressed gene signatures that could partially explain the role of host factors in host adaptation of IAVs. Generally, regardless of the IAV subtype infection, the key difference between mammalian and avian (chicken) cell line expression signatures come from vesicular transport, apoptosis and interferon responses. We

have selected the highly differentially expressed genes to explain this phenomenon. *OSBPL1A*, *MRPS27*, *COPS2*, *ARHGAP21* were highly expressed in CEF cells (Fig. 4a). *OSBPL1A* and *ARHGAP21* are involved in vesicular trafficking while *MRPS27*, *COPS2* have an anti-apoptotic role. Conversely, *CCL5* which is highly expressed in MDCK and A549 cells may induce apoptosis and together with *CCL5*, ISGs and pre-apoptotic signals (*XAF1*) would induce apoptosis and hinder the replication of IAV in A549 cells.

Oxysterol binding proteins (OSBPs) are involved in cell signaling, cytoskeletal organization, lipid homeostasis, sphingolipid (SM) metabolism and *LXR/ABCA1* regulations⁴⁹⁰. OSBP-related protein (ORP) 1 is one of the 12 proteins in the family. It has two variants ORP1S (*OSBPL1A*) and ORP1L (*OSBPL1B*) mainly localized in the cytosol and late endosome respectively⁴⁹⁰. Cholesterol and oxysterols binds ORP1S with high affinity and this observation suggested that binding of ORP1S with sterol promotes transport from the cytosol to the nucleoplasm⁴⁹¹. In HeLa cells, the sterol transport from plasma membrane (PM) to the endoplasmic reticulum (ER) and lipid droplets (LDs) was enhanced by overexpression of ORP1S⁴⁹², which can be used for new membrane synthesis. It has been observed that depletion of membrane cholesterol from the IAV envelope significantly reduced the IAV infectivity in Madin-Darby bovine kidney (MDBK)⁴⁹³ and MDCK cells⁴⁹⁴. Conversely, virus infectivity was maintained upon addition of exogenous cholesterol^{493,494}. It was also suggested that IAV virus genome segments are transported to the plasma membrane by cholesterol-enriched recycling endosomes⁴⁹⁵.

ORP1L the variant of *OSBPL1A* co-localizes and interact with GTP-bound *RAB7*. ORP1L modifies the functional cycle of *RAB7* and interferes with late endosome (LE)/lysosome endocytic membrane trafficking⁴⁹⁰. Cholesterol levels in LE are sensed by ORP1L and depending on the concentration of cholesterol ORP1L changes its conformation and guides the transport of LE to the minus and plus end of microtubules accordingly⁴⁹⁶. ORP1L forms RILP-*RAB7*-ORP1L complex to recruit dynein motors which initiate translocation of LE to the microtubule minus end⁴⁹⁷. Overexpression of ORP1L has been shown to induce protein clusters in endosomes in the perinuclear area of the cell⁴⁹⁸. This suggests that it might also have a role in the transport of IAV containing LE to the nucleoplasm.

It has been demonstrated that, *RAB5* and *RAB7* are required for IAV (H1N1/WSN and H3N2) infection¹³³ and *RAB7* was mainly involved in the intermittent and confined movement of

the virus in the perinuclear region⁴⁹⁹. *RAB7* is LE/lysosomal small GTPase that is required for the fusion of many PH-dependent viruses, including IAV. The known antiviral *IFITM3* partially co-localizes with *RAB7* and LAMP1, and overexpression of *IFITM3* by IFN treatment expands these components⁵⁰⁰. Recently, it was demonstrated that *IFITM3* disrupt intracellular cholesterol homeostasis to block viral entry⁵⁰¹. *IFITM3* inhibit viral entry through disruption of intracellular cholesterol homeostasis and this was accomplished by preventing the interaction of VAPA and OSBP⁵⁰¹. This in turn induces accumulation of cholesterol in multi-vesicular bodies and late endosomes which inhibits the fusion of endosomal membranes with viron-containing vesicles⁵⁰¹. Further, the expression patterns of *IFITM* genes showed little response in chicken compared to higher expressions in duck⁵⁰². Duck *IFITM3* also mediates restriction of H1N1, H6N2 and H11N9 virus infections⁵⁰³. In our investigation, *IFITM3* was not differentially regulated in CEF cells, while it was up-regulated in MDCK (H5N2/F59, H5N3 virus infections) and A549 (H5N2/F59 and H5N2/F189 virus infections) cells. Taken together, increased expression of *OSBPL1A* in CEF cells could contribute for increased rate of virus replication compared to MDCK and A549 cells²⁶⁹.

The highly expressed *ARHGAP21* in CEF cell infections at 10hpi, was also reported to be involved in trafficking through the control of *CDC42* activity⁵⁰⁴. The overexpression of *ARHGAP21* negatively regulates the transport of IAV neuraminidase (NA) to the cell surface and inhibits IAV (H1N1/WSN) virus replication in A549 cells¹¹⁸.

Mammalian mitochondrial ribosomal protein of small subunit 27 (*MRPS27*) is one of the pentatricopeptide repeat domain (PPR) proteins vital for mitochondrial protein translation. Its knockdown is associated with low level of respiratory complexes and cytochrome c oxidase activity (*COXI*, *COXII*)⁵⁰⁵, and disruption of mitochondrial membrane potential (MMP)⁵⁰⁶. Viruses control various cellular organelles for efficient replication and to escape from host defense and immune responses. In addition to its bioenergetics activity, mitochondria is implicated in antiviral response and apoptosis⁵⁰⁷. Apoptosis or cell death is associated with decreased mitochondrial inner MP, increased cytochrome c release, a reduction in cytochrome c oxidase activity, caspase activation and DNA damage⁵⁰⁸. Several virus induced pathological cell deaths or other cellular toxicity are also relies on MMP as a critical event⁵⁰⁹.

Increasing survival of IAV infections in macrophages is associated with translocation of *RIG-I/MAVS* components to the mitochondria to regulate antiviral or apoptotic signals⁵¹⁰. The

PB1-F2 protein of IAV has been implicated to reduce MMP, alters mitochondrial morphology and release of pro-apoptotic proteins and cell death through formation of protein channel pores⁵¹¹. The reduced inner membrane potential induced by PB1-F2 protein also suppresses *RIG-I* signaling pathway and activates *NLRP3* inflammasomes⁵⁷. PB1-F2 protein is not expressed in all IAV strains, however it has been indicated in most of the HPAI H5N1, the pandemic H1N1 (except the 2009 pandemic), H2N2 and H3N2⁵¹² viruses. From sequence prediction, among the viruses used in this study H5N2/F59 and H5N3 couldn't express PB1-F2 protein. In addition to PB1-F2 protein, the M2 protein of the IAV reduces MMP suggesting new pathological mechanism of IAV pathogenesis⁴⁴⁷. In our study, the increased expression of *MRPS27* in CEF cells and MDCK cells (in H5N3 virus infections) could reduced apoptosis and cell death resulting intact cytoskeleton. *MRPS27* was also one of the pro-viral host factors identified by siRNA screens³⁶⁶. Hence, the pathologic effect of IAV either by PB1-F2 or M2 cytotoxicity on cellular mitochondria might compensated by increased expression of *MRPS27* that reduce apoptosis and cell death while increasing IAV replication.

COPS2 is one of the *COPS9* signalosome proteins localized in both cytoplasm and nucleus^{513,514}. It is implicated in the pluripotent maintenance of human and mouse embryonic stem cells (ESCs), somatic cell reprogramming, and regulation of ubiquitin-dependent proteasomal degradation^{514,515}. Knock down of *COPS2* resulted in decreased proliferation rate and G2M arrest in mouse ESCs, while its overexpression increased somatic cell reprogramming even in virus (retrovirus) expressing cells⁵¹⁴. In our investigation, we have observed down-regulation of genes associated with cell proliferation including G2M checkpoint (Fig. 4.1c) in MDCK and A549 cell infections, but not in CEF cells. This suggests that the overexpression of *COPS2* could have a role in maintaining CEF cell proliferation or reprogramming IAV infected cells. However, the overexpression of *COPS2* in CEF cells is not consistent in infection with all IAV subtypes.

CCL5 expression is induced through activation PRRs at early stage of virus infection and by *TNF α* and *IFN γ* at the late stage. It protects against certain infectious diseases; however, it is a key pro-inflammatory cytokine that can also lead to immunopathology⁵¹⁶. Production of *CCL5* has been observed in several respiratory virus infections including RSV and influenza^{516,517}. *TNF α* and *CCL5* were the predominant produced cytokines following H1N1/PR8 virus infection of A549 cells. *TNF α* regulates *CCL5* production with subsequent increased expression of *IL8*,

CCL2 and IFNs in H1N1/PR8 virus infections⁵¹⁸. *CCL5* is not differentially expressed in CEF cells, and this could be due to the lack of the *RIG-I* receptor in chicken⁵⁰².

Elevated level of *CCL5* expression have been also shown in H1N1/PR8 and pH1N1 virus infection in human macrophages^{234,519}, and in the lungs of H5N1 virus infected patients⁵²⁰. Micromolar concentration of *CCL5* have been shown to induce apoptosis in T cells and such *CCL5* induced apoptosis was accompanied by release of cytochrome c with activation of apoptotic caspases (caspase 9 and 3)⁵²¹. In contrast, *CCL5-CCR5* interaction has been shown to inhibit virus induced apoptosis in macrophages and A549 cells, suggesting increased macrophage survival through the control of virus infection⁵²². However, higher level of *CCL5* in macrophages activate ERK1/2 pathways in a CCR5-independent manner⁵²². Activation of ERK1/2 generally promotes cell survival; but under certain conditions, ERK1/2 can also have pro-apoptotic functions⁵²³. How ERK1/2 activation promotes apoptosis is not well understood, however it has been linked to increased mesenchymal apoptosis and decreased epithelial proliferation in fetal lung explants⁵²⁴. Taken together, increased expression of *CCL5* (MDCK) or/and ISGs in A549 cells could play a role in reduced replication of IAV and induction of apoptosis in mammalian cell lines.

Integration and cross-comparisons of the DEGs with host factors from siRNA screens and interactome data sets showed that activation or suppression of the pro-viral cellular pathways in IAV infections was strain and cell dependent. While H5N3 virus induces pro-viral pathways in A549 and MDCK cell infections, the other viruses down-regulate these pathways in strain dependent manner in the same cell lines. H5N3 was the highly replicating virus in all of the cell types, thus we suggest that this virus induces pro-viral pathways for its highest replication in mammalian cell lines³⁴⁵.

In this study, we have also separately investigated DEGs involved in metabolic pathways in the virus infections. We noted that H5N3 virus induced oxidative phosphorylation, nucleic acid, amino acid, lipid and glycan metabolisms in MDCK cells infection. Supported with the siRNA screens data, these pathways were found to be important for the replication of IAVs. In line with this study, previous reports showed induction of glycolysis, lipid and cholesterol biosynthesis in H1N1/PR8 infected MDCK cells^{525,526}. Similarly, a proteomic study in pH1N1 infected NHBE cells showed increased level of proteins associated with oxidative phosphorylation, fatty acid metabolism, glycolysis, TCA-cycle, pentose phosphate pathways. In

addition, the induction of the metabolic pathways were associated with survival of the virus infected cells⁵²⁷. It was suggested that activation of cellular metabolism in IAV infections was a compensatory mechanism of the cell for resources (energy, lipid, amino acids and nucleic acids) which were depleted during virion production⁵²⁷. Recently, Meyer et al., 2017 demonstrated increased expression of mitochondrial genes associated with oxidative phosphorylation in LPAI H6N2 and H7N1 viruses in chicken lung epithelial cell line (CLEC213)⁵²⁸. The same study showed that inhibition of oxidative phosphorylation using oligomycin reduced virus replication without cell toxicity⁵²⁸. Overall, the present data suggested that induction or maintaining of the metabolic pathways during IAV infections could be important for cell survival and IAV replication. To the reverse, down-regulation of metabolic pathways and induction of apoptosis may reduce IAV replication but subsequently may activate inflammatory processes.

4.4 Conclusion

In conclusion, re-annotation of the chicken and canine Affymetrix chips through human ortholog mapping allows for identification of more differentially expressed genes that can be mapped to functional pathways in our setup of infection with different low pathogenic avian influenza viruses. The expression signatures of different IAV isolates in a single cell type are more closely related to each other than the expression signatures of a single isolate in different species cell types. The number of differentially expressed genes in influenza infected MDCK is higher than in A549 cells and lowest in CEF cells. The additional up-regulated genes in MDCK and A549 are linked to increased antiviral, pro-apoptotic and inflammatory activity in strain dependent manner. Conversely, up-regulated genes in CEF cells induce virus infectivity, reduce apoptosis and increase cell proliferation. This is consistent with the fact that there is low impact of LPAI on chicken that can harbor and shed these viruses without developing disease.

Chapter 5: Systems-based approach to examine the cytokine responses in primary mouse lung macrophages infected with low pathogenic avian Influenza virus circulating in South East Asia.

5.1 Introduction

The murine model of influenza virus infection is commonly used to study the transmission, host adaptation, immune responses and pathogenesis of IAVs (reviewed in⁴⁷⁸). However, many human and avian IAVs are not intrinsically pathogenic in mice, hence the virus has to be adapted to the murine host to allow efficient virus replication⁴⁷⁸. Although the mode of species adaption is not fully understood, the available evidence suggests that it involves several factors (e.g. optimization of receptor binding specificity of the hemagglutinin (HA) protein and adaptation of the virus polymerase complex) to allow replication in the cells lining the upper and lower airway^{181,529,530}). Therefore, the adaptation of LPAI viruses to a new host involves the introduction of new genetic changes that may mask inherent biological properties of these viruses¹⁵². The initial sites for IAV infection are the epithelial cells that line the upper respiratory tract (URT), but disease severity is associated with replication in the lower respiratory tract (LRT), and involves other cell types including the lung resident macrophages⁵³¹. Therefore, the capacity of influenza viruses to cause disease in the animal models depends on their capacity to replicate in the airway cells lining the URT prior to LRT. This biological property is therefore distinct from the capacity of viruses to induce pro-inflammatory cytokines. The analysis of LPAI infection in immune cells isolated from the lung tissue allows us to examine the biological properties of the viruses without the prerequisite for productive infection in the upper airway of the mouse.

Macrophages play an important role in disease progression following influenza virus infection, and the host response in macrophage cells following virus infection are important determinants of disease severity. The biological properties of different macrophages are dependent on their tissue location, and macrophages that are resident in the lung exhibit subtle differences in their biological properties when compared to the macrophages resident in other

tissues. At least three categories of lung-resident macrophage cells have been described, and are referred to as alveolar, interstitial and exudate macrophages. Additionally, macrophages can be changed from quiescent to activated states in response to external cues, which include cytokines (e.g. *IFN* γ , *IL-4*, *IL-13*, *TGF* β), microbes and their products (e.g. LPS), growth factors (e.g. *GSF-1* and *GM-CSF*) and other modulators (e.g. phagocytosis following bacterial and viral infections)⁵³¹.

Depending on host homeostasis signals and external environment challenges, macrophages present heterogeneous functions such as the release of reactive toxic species, production of pro/anti-inflammatory cytokines and lipid mediators, tissue remodeling clearance of apoptotic cells and oxidized lipoproteins, wound healing and regeneration⁵³¹⁻⁵³³. The interaction between different IAVs and macrophages varies and the host response to infection is largely virus isolate-specific⁵³⁴⁻⁵³⁶. In this context there are marked differences in the ability of different virus strains to induce macrophage-related inflammatory mediators, but the mechanisms responsible for this variation are largely unknown. Understanding the molecular mechanisms by which different IAV can induce pro-inflammatory cytokines in virus-infected macrophages may allow the development of therapeutic strategies that can mitigate IAV infection. In addition, this may lead to a better understanding of how viruses that are circulating in the natural environment attain these biological properties, and if they can transfer these properties to other circulating viruses (e.g. by gene reassortment).

Our group have previously described the biological properties of several LPAI viruses that were isolated from imported live poultry during routine surveillance, which included H5N2, H5N3 and H9N2 viruses^{269,347}. It was concluded that these viruses were representative of LPAI viruses that were circulating in ducks within the SE Asia region. In this current study, we have extended the investigation and have applied a system-based approach to investigate the host response to these viruses in pulmonary murine lung macrophages (PM Φ).

5.2 Results

5.2.1 Global changes in host gene expressions at 24 hours post-infection.

In the Affymetrix microarray system (GeneChip Mouse Genome 430 2.0 Array) there are 28,974 annotated genes, and analysis of the microarray data indicated that 89.8% (25,886), 89.8% (25,886), 89.7% (25,850) and 84.6% (24,368) of the annotated genes were expressed in H1N1/WSN and mock-infected cells, H5N2 and H5N3 virus-infected cells respectively. The

global gene expression in PMΦ infected with each of the three viruses was examined at 24hpi. Overall, 1830 probe sets (1103 genes) were differentially expressed in at least one of the virus infections at 24hpi. The numbers of DEGs at 24hpi in H1N1/WSN, H5N2, and H5N3 virus-infected cells were 193, 21 and 1071 respectively (Fig. 5.1a). The numbers of up- and down-regulated genes are 3.9 and 10.4 times higher in H5N3 virus-infected cells than H1N1/WSN virus-infected cells. There were no significantly up-regulated genes in cells infected with H5N2 virus, and the down-regulated genes in cells infected with the H5N3 virus were 23.8 times higher than H5N2 virus infections. The number of DEGs with $|\log^{\text{FC}}|$ value > 3 were also higher following the H5N3 virus infection (Fig. 5.1b (i and ii)).

Overlap analysis of the DEGs at 24hpi is shown in Fig. 5.1c. One hundred thirty-two (91.6%) (132/144) of the genes showing up-regulated expression in cells infected with H1N1/WSN virus infection were shared with the H5N3 virus-infected cells. The up-regulated shared genes include *APOL9A*, *CCL12*, *CMCK2*, *DHX58*, *H28*, *IFI44*, *IFIT1*, *IFIT2*, *IFIT3*, *ISG15*, *MX1*, *OAS1A*, *OASL2*, *RSAD2*, *RTP4*, *SLFN3*, *USP18*, *XAF1*, and *ZBP1*. These genes were associated with innate immune response to viral infection and were up-regulated in different IAV strains (H1N1, H3N2, H5N1, H5N2, and H7N1) in A549 cells²⁶⁸, human blood, PBMC, and primary epithelial cells⁵³⁷. These genes (e.g. *IFI44*, *IFIT1*, *OAS1A*, *OASL2*, and *XAF1*) were also implicated in response to other respiratory viruses (RSV and HRV) in human PBMC and lung epithelial cells²⁶⁸. Similarly, most (76.2% of the H5N2 virus and 60.4% of the H1N1/WSN virus) down-regulated genes were shared with the H5N3 virus down-regulated genes.

Fourteen down-regulated genes were shared between the three viruses among which *AGER*, *SFTPD*, *EGFL6*, *NDNF*, *DSP* and *TTN* were involved in regulation of cell adhesion and wound healing (Fig. 5.1c). Approximately, 76.8% (439/571) up-regulated and 93.8% (469/500) down-regulated genes were unique to H5N3 virus infection (Fig. 5.1c). The hierarchical clustering of DEGs/probe sets at 24hpi for the three viruses was illustrated in Fig. 5.1d. In H5N3 virus-infected cells, there were unique up-regulated and down-regulated gene cluster patterns which was consistent with the higher proportion of unique up-and down-regulated genes shown in Fig. 5.1c. Collectively, the data showed that LPAI viruses induce few common gene expression signatures compared to other IAV isolates and other respiratory viruses, suggesting that most of the DEGs were strain-specific expression signatures.

Figure 5.1: Global host gene expression in mouse PM Φ infected with IAVs at 24 hours-post infection compared to mock infections.

a) Number of up-and down-regulated genes b) the extent of up-and down-regulated genes with log^{FC} values i) up-regulated genes ii) down-regulated genes c) Overlap analysis of DEGs in three IAV infections d) unsupervised hierarchical clustering of DEGs

5.2.2 Functional annotation of DEGs

Functional annotations of the DEGs i.e. GO biological process and KEGG pathway was analyzed using Metascape (<http://metascape.org/gp/index.html#/main/step1>)³⁵⁴. We compare significantly enriched pathways and their enrichment p-values (q-value) for all the viruses at 24hpi. The top 20 statistically significant pathways (q-value < 0.05) were described in Fig. 5.2a. Generally, the pathways are related to immune responses (cytokines, interferons and inflammatory responses), morphogenesis and cell migration (Fig. 5.2A). Separate pathway enrichment analysis for the up-regulated genes (Fig. 5.2b) and down-regulated genes (Fig. 5.2c) indicated that the cytokine, interferon and inflammatory responses were enriched in the up-regulated genes, while the enrichment of down-regulated genes were associated with cell migration, morphogenesis and metabolic pathways. It was expected that infection of macrophage cells with IAVs induces immunological responses⁵³⁸. The activation of type I and II interferon pathways (Fig. 5.2a and Fig. 5.2b) alleviate IAV infections⁵³⁹, while increased activation of cytokine responses could result in the induction of a cytokine storm (inflammatory cytokines) which are associated with immune-pathological conditions^{538,539}. The H1N1/WSN and H5N3 viruses significantly induced interferon pathways (Fig. 5.2a and 5.2b) and while these viruses induced inflammatory responses, these were more prominent in H5N3 virus-infected cells at 24hpi. Increased induction of inflammatory cytokines is a characteristic of infections caused by HPAI viruses⁵⁴⁰, and the induction of inflammatory pathways in H5N3 virus-infected cells could indicate its potential to induce similar pathologic-related immune conditions. Inflammatory cytokines could also disturb the normal cellular structure, morphology and migration contributing to infiltration of fluids and immune cells at the site of inflammation⁵⁴¹. H5N2 virus infections showed no induction of inflammatory pathways, deregulation of pathways related to cellular growth and pathways related to morphogenesis compared with the mouse-adapted H1N1/WSN virus at 24hpi (Fig. 5.2b, 5.2c). Taken together, these data indicate that the H5N2 virus is less able to induce pathways related to inflammatory cytokines compared to the H1N1/WSN virus, while the H5N3 virus induced inflammatory pathways at higher levels.

Figure 5.2: Top 20 statistically significant ($q < 0.05$) functionally enriched pathways or/and GO biological process (GBP) of the DEGs.

a) Pathways/GBP from both up- and down-regulated genes. b) Pathways/GBP from up-regulated genes c) Pathways/GBP from down-regulated genes. Increased intensity of the colour shows increased significance of the pathway/GBP corresponding to each virus infection conditions. More related pathways/GBP were clustered together (left) and the viruses inducing similar pathways/GBP are clustered together (top).

5.2.3 Host response to the IAVs infection

Based on the functional annotation and the extent of gene expression changes (highest $\log^{FC}$ values), we focused on genes related to cytokines, interferon and cell death and macrophage activation pathways in more detail since these have been implicated in increased virus pathogenicity. Selected genes from virus-induced expression changes were also validated by biochemical cytokine assay.

Figure 5.3: Selected host responses of PMΦ in IAV infections compared to mock infections.

The expression levels of PMΦ genes associated with a) cytokines and chemokines b) ISGs c) cell death d) macrophage activation e) Boxplot and beeswarm plots indicating the expression distribution of the DEGs involved in macrophage activation.

5.2.4 IAV induced cytokines

Infections with the H1N1/WSN and H5N3 viruses induce expression of several cytokine genes in PMΦ (Fig. 5.3a). As indicated in Fig. 5.3a, the extent of cytokine gene expression following H5N3 virus infection is higher than in cells infected with the H1N1/WSN and H5N2 viruses. In cells infected with the H5N3 virus, the levels of cytokine gene expression was significantly higher than in cells infected with either the H1N1/WSN or H5N2 viruses. This has been experimentally validated by measuring the cytokines and showed consistent with the expression result that there were increased level of inflammatory mediators (*IL-1α*, *IL-1β*, *IL-6*, *MIP-1α* (*CCL3*), *MIP-1β* (*CCL4*), *RANTES* (*CCL5*), *MCP-1* (*CCL2*)) following H5N3 virus infection at 4, 8 and 20 hpi compared to cells infected with the H1N1/WSN, H5N2 and H9N2 viruses³⁴⁶. In contrast, the expression of cytokine that are important for the survival of macrophages was down-regulated in H1N1/WSN and H5N3 virus-infected cells at 24hpi (Fig. 5.3a)⁵⁴². Taken together, inflammatory chemokines and cytokines are highly induced in H5N3 virus-infected cells, compared with cells infected with H1N1/WSN and H5N2 viruses.

5.2.5 Interferon stimulated genes (ISGs) and cell death

Virus infections are first recognized by the pattern recognition receptors of the host cell and subsequently induce activation of IRFs and production of IFN. This can be followed by transcription of several antiviral ISGs⁵⁴³. The ISGs have several cellular effects which include anti-viral, immuno-modulation, anti-angiogenic, cell cycle arrest and anti-apoptotic functions⁵⁴⁴. As expected, several of the ISGs were induced in PMΦ upon infection with the H1N1/WSN and H5N3 viruses (Fig. 5.3b). Unlike the higher expression of cytokine responses in H5N3 virus infection, generally the extent of ISGs expression were similar in H1N1/WSN and H5N3 virus infections (Fig. 5.3b). However, certain ISGs (e.g. guanylate binding protein 1, 2, 3 and 6 (*GBP1*, 2, 3 and 6)) have higher expression levels following infection with the H5N3 virus than the other two isolates (Fig. 5.3b). *GBP1*, *GBP2* and *GBP3* have antiviral activity against IAV⁵⁴⁵ and can also be activated by inflammatory cytokines⁵⁴⁶.

IFIT1 is significantly expressed in both H1N1/WSN and H5N3 virus-infected cells. It is an antiviral protein for viruses that bear 5'-triphosphate RNA⁵⁴⁷, however recently it was reported to have no effect on negative sense RNA viruses including influenza⁵⁴⁸. Raw264.7 cells pre-treated with *TNFα* presented with an elevated level of *IFIT1* induced cell death⁵⁴⁹. In line with this, in H1N1/WSN and H5N3 virus infections, an elevated level of *IFIT1* was associated

with increased expression of anti-apoptotic genes XIAP associated factor 1 (*XAF1*) in both viruses and TNF receptor-associated factor 1 (*TRAF1*) in H5N3 virus (Fig. 5.3c) in response to early cell death⁵⁵⁰.

5.2.6 Macrophage activation

There is growing evidence that macrophages (MΦ) are functionally subtyped into M1MΦ (activated by bacterial lipopolysaccharides (LPS) or IFNγ) and M2MΦ (activated by IL-4, IL-13 or TGFβ). M1MΦ favour glycolytic metabolism, production of reactive oxygen species (ROS) and pro-inflammatory reactions, and are associated with a Th1 response which is essential to combat intracellular infections. The M2MΦ utilize fatty acid oxidation and mitochondrial respiration and induce anti-inflammatory mediators and are linked to Th2 response, which is important for extracellular parasitic responses, tissue remodeling, and repair and wound healing. M1MΦ and M2MΦ have a different metabolic profile^{531-533,551-555}. M1MΦ cells are characterized by overexpression of *NOS2*, *MTOR*, *PFKFB3*, *COX2*, *PTGES*, *HIF1A*, *SLC2A1*, *G6PD*, *GCK* and down-regulation of *DHCR24*, *COX1*, *LTA4H*, *TBXAS1*, *ALOX5*, and *LCAT*. On the other hand, M2MΦ cells showed increased expression of *ARG1*, *PRKAB1*, *PFKFB1*, *STAT6*, *COX1*, *ATF3*, *PPARA*, *PPARG*, *SDH*, *NRF1*, *ESRRA*, *TNFAIP812*, *EPAS1* and low level of *PTGES*.

We examined if the expression changes in IAV infected PMΦ lead to the formation of either of these MΦ phenotypes. The expression changes of M1MΦ/M2MΦ associated metabolic genes indicated that H5N3 virus induces classical activation of PMΦ presenting with M1MΦ like phenotype with up-regulation of *SLC2A1* and down-regulation of *PPARG*, *ATF3* and *TNFAIP812* (Fig. 5.3d, 5.3e).

Similarly, genes associated with wound healing and tissue remodeling were down-regulated in cells following H5N3 infection (Fig. 5.2d). In contrast, H1N1/WSN and H5N2 virus infection induced a M2MΦ like phenotype which is characterized by increased expression of arginase 1 (*ARG1*) at 2hpi (Fig. 5.3d, 5.3e) and had a similar metabolic profile as unpolarized (inactivated) MΦ (mock infected) (Fig. 5.2c, 5.3d, 5.3e)⁵⁵⁶. Taken together, these data suggest that while H1N1/WSN and H5N2 viruses induce M2MΦ like activation, the H5N3 virus induces M1MΦ phenotype MΦs.

Figure 5.4: DEGs associated with metabolic pathways.

a) up-regulated DEGs associated with metabolic pathways b) down-regulated DEGs associated with metabolic pathways c) the top 20 up- and down-regulated KEGG pathways of the metabolic genes.

5.2.7 DEGs associated with metabolic pathways

Among the DEGs in at least one of the virus infections in PM Φ , 110 genes were found to be involved in metabolic pathways. Fifty four of the 110 genes were up-regulated mainly in H5N3 virus infections (Fig 5.4a), while 56 of the genes were down-regulated (Fig 5.4b). KEGG pathway analysis of these up-and down-regulated genes indicated that metabolism of cofactors and vitamins, amino acid metabolism, and carbon metabolisms were up-regulated (Fig 5.4c). In contrast, N-glycan biosynthesis and lipid metabolism were down-regulated. Three pathways were enriched by either up- or down-regulated genes, of which purine metabolism was prominent in down-regulated genes. Collectively, in PM Φ cells infections, H5N3 virus induces higher extent of amino acid, vitamin and carbon metabolism compared to H1N1/WSN and H5N2/F118 viruses. In contrast, lipid, glycan and purine metabolisms were down-regulated.

5.3 Discussion

Influenza virus infection of alveolar epithelial cells type II (AECII) and murine lung tissue showed increased expression of cytokines (*IL-6*, *CXCL10*, *CXCL5*, and *CXCL9*) and ISGs (*RSAD2*, *IFIT1*, *IFI44*, *IFIT2*, *MX1*, *OASL2*, *GBP3*, *IIGP1*, *SLNF4*). Although the levels of expression differed in these studies, several of the inflammatory cytokines and interferon response were detected in both the lung and AECII transcriptome analysis⁵⁵⁷. Similarly, several other cell types of the lung e.g. epithelial cells type I, macrophages, dendritic cells, and mast cells contribute to inflammatory cytokines production during IAV infections⁵⁵⁸⁻⁵⁶⁰. We observed similar increases in expression of inflammatory cytokines in the H5N3 virus-infected PM Φ cells.

The H5N3 virus induced higher levels of pro-inflammatory cytokine in mouse PM Φ compared with the H1N1/WSN and H5N2 viruses. Increased inflammatory cytokine expression in whole mouse lung tissue infected with different mouse-adapted pathogenic/virulent IAV subtypes have been described previously using a transcriptomic approach^{246,557,561-566}. The 2009 pH1N1 virus induced increased expression of *CXCL10*, *OASL*, *OAS2*, *IFIH1*, *USP18* and *XAFI*²⁴⁶. Similarly, up-regulated expression of inflammatory cytokines (*CXCL10*, *CXCL11*, *CXCL9*, *CXCL13*, *CCL9*, *CCL5*, *IL-6*, *IL1*, *IL1A*, *IL1B*, *TNF*, and *TNF- α*) and ISGs (*GBP2*, *GBP4*, *GTP1*, *ISG15*, *DdX58*, *IFIH1*) were shown in virulent H3N2/HK31 infections⁵⁶¹⁻⁵⁶³. In addition, the lung transcriptome of HPAI H5N1 virus infection in susceptible mouse revealed increased expression of inflammatory cytokines (*CCL2*, *CXCL2*, and *TNF- α* , *CSF3*) and interferon responses (*IFN- α* , *IFN- β* , *IL28B*, *IFNA4*, *IFNA5*) which were associated with

increased viral load and disease severity^{564,565}. The H7N9, H7N7 and H5N1 viruses infected mouse lung transcriptome also showed higher levels of *CXCL13*, *IL-6*, *IFNG*, *IFNB1*, *IRF9*, *IRF1*, and *TNF* compared with the 2009 pH1N1 infections. Furthermore, increased expressions of inflammatory cytokines and decreased transcription of genes associated with lipid metabolism correlated with increased disease severity²⁴⁶.

Lung-resident MΦ have been implicated in disease severity following IAV infection^{534,567}. The pathogenicity of IAVs correlated with their ability to induce high levels of pro-inflammatory cytokines in MΦ⁵⁶⁸, which in turn is proposed to cause the increased migration of granulocytes into the lung that leads to localized tissue damage²³⁸. Although the association of high levels of infectious virus production and higher levels of inflammatory responses is suggested, other factors may also play a role in this process. The PR8/H1N1 virus strain is highly virulent and was characterized by higher level of virus production and lung inflammation than HKx31/H3N2 and BJx109/H3N2 viruses⁵⁶⁹. Characterization of inflammatory cell types in mice lung showed that overexpression of *CCL2* and *CCR2*+ monocyte-derived cells (monocyte-derived DCs and exudate MΦ) were the predominant cause of immune pathology during PR8/H1N1 virus infection⁵⁷⁰. In addition, TRAIL-expressing MΦ (exudate MΦ) were implicated in influenza-induced pneumonia through increased expression/ release of *TRAIL* and alveolar epithelial cell (AEC) apoptosis⁵⁷¹. The HPAI H5N1 virus is a potent inducer of cytokine storms (very high level of inflammatory mediators (e.g. *IL1*, *IL-6*, *TNFα*, *IL-8*, *MCP-1*, *IP-10*, *IFNγ*) in patients, *in vivo* mouse model experiments and *in vitro* macrophage infections^{568,572-578}.

Although all the three viruses infected PMΦ with similar efficiency, surprisingly the H5N3 virus induced highest inflammatory responses at 24hpi compared to H1N1/WSN and H5N2 viruses. However, we obtained no evidence that the infection of murine MΦ with either of the viruses induced a productive infection (Sutejo, Tan and Sugrue, unpublished observations). In this context mouse MΦ infected with either of H1N1/WSN, A/HK/483/97 H5N1 (HPAI) or A/X-31/H3N2 viruses showed induced expression of inflammatory cytokines through activation of either *NF-KB* or *MAPK* signaling pathways in strain dependent manner⁵⁷⁹⁻⁵⁸³. However, it was reported that a productive infection in murine-derived MΦ was characteristic of the HPAI H5N1 virus⁵⁷⁹. Although these studies have indicated stain-specific differences in pro-inflammatory cytokine induction in murine MΦ, it is not clear if these observations can be extrapolated to human MΦ. However, HPAI H5N1 infection lead to the expression of higher levels of

inflammatory mediators than seasonal H1N1 and H3N2 viruses in human AM^{519,573}. Moreover, the avian viruses H5N1 and H7N9 induced higher inflammatory responses than the seasonal or 2009 H1N1 pandemic viruses^{234,568,574,584-588}.

Interestingly, H5N3 virus-infected PMΦ cells induce expression patterns and pathways (including cytokines) that were similar to the response that we have observed in respiratory syncytial virus (RSV)-infected PMΦ cells (Appendix E)⁵⁸⁹. RSV is another important virus that can cause lung tissue damage via pro-inflammatory cytokine induction in the lower airway. Therefore comparing the cellular response to these viruses may give insights into the interaction that these viruses with MΦ, and possibly the induction of common pathways that may be activated during virus infection that leads to inflammation. We compared the DEGs in this study with DEGs of PMΦ in response to RSV infection from our previously published work⁵⁸⁹. Overlap analysis of the up-regulated genes in IAVs infection with RSV indicated 465 (277 up-regulated and 188 down-regulated), 102 (100 up-regulated and 2 down-regulated) genes shared in H5N3 and H1N1/WSN virus-infected cells respectively (Appendix E), indicating, higher numbers of DEGs shared between RSV and the H5N3 virus than with the H1N1/WSN and H5N2 viruses. The severe form of diseases due to RSV infection is associated with increased level of inflammatory cytokines and chemokines (*TNFα*, *IL-1αβ*, *CXCL8*, *CXCL10*, *CCL5*, *IL6*, *CCL2* and *CCL3*) production by epithelial cells and MΦ (reviewed in^{590,591}). Similarly, RSV-infected PMΦ showed an elevated level of these inflammatory mediators⁵⁸⁹. The levels of either gene expression or cytokine proteins (*IL-1α*, *TNFα*, *IL-6*, *CXCL10*, *CCL5*, *CCL2*, *CCL3*, *CCL4*, and *CCL5*) in RSV-infected PMΦ were comparable to the H5N3 virus infection. This suggests that the H5N3 LPAI virus and RSV could share common features in the way that they interact with the MΦ.

Since the murine macrophage cells were equally susceptible to infection with each of the three viruses these differences cannot be explained by reduced infections rates. The H5N3 virus belongs to a H5N1 clade of viruses and internal genes have been shown to share common ancestry with different subtypes of avian viruses circulating in different geographical locations and environment³⁴⁷. This previous analysis revealed a high level of sequence conservation between the H5N3 and H5N2 viruses. However, we can hypothesize that small changes in gene sequences may be sufficient to account for the significant differences in the host responses to infection. Although the H1N1/WSN virus was able to replicate efficiently in BALB/c mice, the

H5N2 and H5N3 viruses replicated poorly in the mouse infection model (Yeo, Tan and Sugrue, unpublished observations). This is consistent with the expected low efficiency of infection of these avian viruses in mammalian hosts, including humans, and highlights a limitation in the mouse model when examining the interaction of host response in murine MΦ. The capacity of the H5N3 virus to induce pro-inflammatory cytokines in MΦ is therefore a property of the virus that is distinct from its capacity to replicate in non-avian cell types. We have no evidence that this biological property is unique to the H5N3 virus examined in this study, and it is possible that the capacity to induce high levels of cytokines may also be a property of other circulating LPAI viruses. The expected low replication rates in non-avian hosts may normally mitigate this intrinsic biological property of such viruses. However, we can speculate that in the natural environment such viruses (e.g. the H5N3 virus) could transfer their property of inducing high levels of pro-inflammatory cytokines to another virus that replicates better in mammals. Assuming that such a reassorted virus would be viable in the natural environment, this could result in inherent biological changes with medically important implications. It is therefore, necessary to understand this process more completely and identify sequence motifs that may correlate with this biological property (e.g. using the FluSurver platform: <http://flusurver.bii.a-star.edu.sg>). This information would be of use in surveillance studies which would allow the threat of circulating viruses to be better assessed.

The complete mechanism of avian IAV adaptation and pathogenicity to mammalian hosts is not well understood yet. However, it is likely that it involves HA protein receptor binding specificity and polymerase complex activity⁵⁹². In terms of key host specificity markers like HA 226 (Q avian, L mammalian, H3-based numbering) and PB2 627 (E avian, K mammalian), the H5N3 virus strain studied here still has the avian-preferred residues in both cases. Interestingly, the HA does have one S159N mutation that has been implicated in increased binding to mammalian-type $\alpha(2, 6)$ -linked SA receptors without loss of binding to avian-like $\alpha(2, 3)$ -linked SA⁵⁹³. However, the same mutation is also found in the H5N2 virus strain used in this study so it cannot be held responsible for the observed differences. On the other hand, we have found a K615R mutation in the PA protein of the H5N3 virus isolate, but not in the H5N2 and H9N2 viruses used in our study. Combination of mutations in internal genes including PA 615R has been suggested to contribute to the virulence of avian viruses in mammalian hosts⁴⁷⁶.

Under normal circumstance, a cell has its own inherent metabolic activity that is responsible for a variety of different physiological processes. Increasing evidence suggests that metabolism under the influence of specific stimuli (e.g. virus infection) is an important determinant for functional phenotype differences in MΦ⁵⁵². There are two mechanistic options to connect MΦ and metabolism, i.e. either the macrophage cell can regulate metabolic function extrinsically, or intrinsic metabolic functions dictates macrophage activation states⁵⁵². The current study demonstrated increased expression of *ARG1* in H1N1 and H5N2 virus-infected PMΦ at 2hpi. In activated M2MΦ there is increased expression of *ARG1* catalyzes polyamine production which is essential for the synthesis of collagen, tissue remodeling and cell proliferation⁵⁵³. In line with this, following H1N1/WSN and H5N2 virus infections there was a remarkable change in the maintaining of the down-regulated pathways that were associated with cell proliferation and cellular organization from 2hpi to 24hpi (Fig. 5.2C, Appendix E). In contrast, M1MΦ are characterized by increased expression of NOS2 and NO production. The subsequent induction of *NF-κB* mediated inflammatory mediators (*TNF-α*, *IL-6*, *IL-12*, NO and others) leads to tissue inflammation and cell death. In addition, the M1MΦ reprogram glucose metabolism to the glycolytic pathway and the pentose phosphate pathway (PPP), which results in production of reactive oxygen species (ROS) that mediate tissue disruption⁵⁵³. Similarly, the current study revealed that the H5N3 virus activated the PMΦ into a M1MΦ phenotype which is characterized by increased expression of *NF-κB* mediated inflammatory cytokines and down-regulation expression of pathways related to proliferation and myogenesis. Although the effect of different IAV strains in activation of different MΦ subtypes is not well understood, recently Zhao et al., 2014 suggested that activation of MΦ into M1MΦ/M2MΦ phenotype is time and macrophage type dependent⁵⁹⁴. Continuous activation of MΦ into the M1MΦ phenotype is associated with severe disease while the M2MΦ phenotype is an indication of recovery from disease/ infection⁵⁹⁴.

This analysis demonstrates that LPAI viruses interact with immune-related cells in a virus-specific manner, leading to virus stain-specific changes in pro-inflammatory cytokine expression. The capacity of the avian H5N3 virus to infect mammalian host cells and result in increased expression of inflammatory mediators suggests that this is a unique property that is distinct from other factors, e.g. virus replication in mammals. Since in the natural environment the H5N3 virus can potentially reassort with other circulating influenza viruses, and if such virus

is viable this may be able to transfer unique properties that are associated with increased cytokine induction. The identification of sequence motifs in the H5N3 virus that is associated with the increased cytokine induction will be the subject of future studies. This information will allow the identification of sequence motifs that can be used in surveillance studies to assess LPAI viruses which could impose threats to human health in the future.

5.4 Conclusion

The H5N2 and H5N3 LPAI viruses can infect PM Φ similarly to the mouse-adapted H1N1/WSN virus. The global and temporal host gene expression indicated that while there were comparable expression levels of ISG in H1N1/WSN and H5N3 virus-infected cells, increased pro-inflammatory responses were associated with H5N3 infection. This systems-based approach provides evidence that increased cytokine induction during influenza virus infection is not a property that is unique to HPAI viruses, but it is also an inherent property of some LPAI viruses. A better understanding of the signals in LPAI viruses that induce host responses to infection should provide insights into the molecular mechanism of pathogenesis. It will also allow a better understanding of the contribution that these viruses play in the emergence of novel influenza strains in epidemics and pandemics.

Chapter 6: Benchmarking selected computational gene network growing tools in the context of virus-host interactions

6.1 Introduction

Gene and/or protein networks are used as a cardinal representation of various types of biological processes and help in the prediction of molecular and cellular function⁵⁹⁵. It has been indicated that interacting genes/proteins may be part of the same pathway or biological process and, in larger scale, work together in similar cellular functions^{596,597}. Reliable prediction and precise treatment of complex cellular functions require knowing the network of as many as possible genes and their products (e.g., proteins) connected in functional pathways. Pairwise interactions at the gene and protein levels have been analyzed extensively, often from low or high-throughput experimental screens, with the links (annotated as functional interactions) available in online databases (reviewed in⁵⁹⁸). However, these methods do not fully investigate the highly complex interaction patterns in cellular systems and variability in terms of their accuracy and reproducibility⁵⁹⁹ is encountered. Additionally, interactions seen *in vitro*, especially in large-scale screens, may not occur *in vivo* due to spatial and temporal constraints. At the same time, the exact number of protein-protein interaction (PPI) in human is not known, and the available data is estimated to represent approximately 10% of the total PPIs in human⁶⁰⁰. Hence, computational PPI predictions have become increasingly important for detection of new interactions and protein networks. There are several computational gene and/or protein network databases that combine different experimental and computationally predicted interactions through the integration of PPI information that are obtained from different sources (reviewed in^{371,376,601}).

Network growing functions allow extension of networks with additional related genes and are important to improve the understanding of the greater functional context and organization of genes and/or proteins, as well as for discovering novel functions of genes that could have been missed by experimental investigations³⁴¹. This could also be achieved by integrating gene expression or “omics” data from specifically investigated conditions with PPI information into a complete large network followed by extraction of active sub-networks related

to the respective conditions⁶⁰². The non-expressed genes which could be identified in the network are expected to have similar function with the expressed genes that could be missed by the experiments⁶⁰². Such sub-network extraction tools, given user input of gene expression data, include DEGAS (Dysregulated Gene set Analysis via Subnetworks)⁶⁰³, KeyPathwayMiner⁶⁰², and JActiveModules⁶⁰⁴. While other methods (e.g. Search Tool for the Retrieval of Interacting Genes/Proteins (STRING)³⁷⁶, GeneMANIA³⁷¹ and Ingenuity Pathway Analysis (IPA) (<http://www.ingenuity.com/products/ipa>)³⁷³ support the construction of networks for query gene lists without explicit user-provided expression data and automatic network growing of specific numbers of additional genes for discovering functionally related genes⁶⁰⁵⁻⁶⁰⁷.

The motivation of this study was to benchmark the selected gene/protein network growing services STRING³⁷⁶, GeneMANIA³⁷¹ (both academically freely available) and IPA³⁷³ (commercialized) in the context of virus-host interaction (e.g. Influenza A virus (IAV)) in a user-relevant setup. These three services were selected for their ability to take user gene lists as sufficient input and execute the network growing with user-defined numbers of nodes automatically added to the query genes (e.g 10, 20, 50 and 100 nodes).

It would be theoretically interesting to compare the tools' basic algorithms in a well-defined equal search space. However, it would not reflect the reality that a characteristic difference of the tools is the use of different underlying interaction sources and information, which will directly influence the performance a typical user will experience. With the benefit of the user community and intrinsic inseparability of algorithms and used databases in mind, our aim was to benchmark the available tools in their full implementation as tool/web service.

We chose influenza virus required host factors (IHF) interactions for this benchmark because 1) several large-scale screens have been performed recently with data available, 2) experimentally validated host factors do not correspond to a single pathway but are functionally loosely linked covering a broad range of cellular functions and 3) network growing services can be used to find new candidate host factors as potential drug targets against influenza in future.

6.2 Results

6.2.1 The siRNA screening studies show low reproducibility.

So far there were 11 siRNA screening studies that have identified host factors required for influenza (IHF). Overall, 1,580 human IHFs were identified from the 11 siRNA studies of

which 28 and 158 IHFs were shared by at least three and two studies respectively (Fig. 6.1a, Appendix A).

Figure 6.1: Meta-analysis of IHFs from siRNA screening studies.
a) Number of IHFs shared between the siRNA screening studies, b) Number of IHFs shared between each siRNA screening study: the higher the number of overlapping genes, the higher the intensity of the color (from 0 overlapping genes = yellow, to 30 = green).

The number of overlapping IHFs among the individual siRNA screens ranges from 0 to 30 genes, with a maximum overlap of ~10% (between Karlas et al., 2010 and König et al., 2010)

(Fig. 6.1b). Previous reviews from 4 to 6 siRNA screening studies showed similarly low level of overlapping IHFs²⁷⁵⁻²⁷⁷. The low overlap between the siRNA screens was explained due to methodological differences (e.g. differences in cell type, virus strains, infection dose, time of post infection and siRNA libraries (Table 2.2)). However, recently it was reported that this discordance is mainly due to false negative results^{277,340}. Therefore, a merged list of all the siRNA screening studies in the context of IAV would give the most comprehensive positive set of IHFs.

Mapping of these IHFs to all three network growing tools for fair comparison was successful for most of IHFs (Fig. 6.1a). Next, we used the 28 (small set) and 153 (medium set) IHFs as query genes (seed sets) to grow with 10, 20, 50 and 100 additional genes in each of the network growing tools. The intersection of grown genes from small set seeds were evaluated against the 153 and 1463 IHFs, while the grown genes from medium set seeds were evaluated against the whole positive sets (1463) IHFs.

6.2.2 Network growing algorithms partially recover IHFs from siRNA screens.

The results showed that as the number of genes building the network increases from 10 to 100, the number of genes intersecting with known IHFs also increased (Fig. 6.2a, b and c). However, the rate of IHFs detections (percentage) shows variable trends with slightly decreasing performance after growing 20 and/or 50 (Fig. 6.2d, e and f), suggesting early recruitment of known host factors. Among 100 genes grown from the small set seeds, 60%, 24% and 27% of the automatically recruited genes in STRING, GeneMANIA and IPA respectively, overlapped with the known IHFs (Fig. 6.2e). Similarly, the detection rates were 42%, 44% and 21% in STRING, GeneMANIA and IPA, respectively with the medium set seeds (Fig. 6.2f). Therefore, while using the small set as seed genes, STRING was the sole strongest performer (Fig. 6.2e), both GeneMANIA and STRING seem to perform equally well with the medium set seed genes (Fig. 6.2f).

Among the 100 genes recruited by growing networks using the 28 small set seeds, we also examined the distribution of correctly identified IHFs among the 125 genes with support by 2 siRNA studies vs 1,310 genes with support by only one siRNA screen (Fig. 6.1a; Fig. 6.2a and d). For a gene to be detected in the 125 shared lists, proportionally we expect ~10 genes to be detected in the 1,310 non-shared genes.

Figure 6.2: The trend of network growing tools in detecting known IGFs.

a) Number of detected known IGFs upon growing networks with the small set (28 IGF) seeds and 153 known IGFs as positive set. b) Number of detected known IGFs upon growing networks with the small set (28 IGF) seeds and 1463 known IGFs as positive set. c) Number of detected known IGFs upon growing networks with medium set (153 IGFs) seeds and 1463 known IGFs as positive set. d) Percentage of detected known IGFs upon growing networks with the small set (28 IGF) seeds and 153 known IGFs as positive set. e) Percentage of detected known IGFs upon growing networks with the small set (28 IGF) seeds and 1463 known IGFs as positive set. f) Percentage of detected known IGFs upon growing networks with medium set (153 IGFs) seeds and 1463 known IGFs as positive set.

Interestingly, the proportions of shared genes detected in STRING, GeneMANIA and IPA were 3.4, 3.1 and 2.4 times higher than the non-shared genes respectively. This means that the network growing tools recruit shared IHFs in higher proportions, suggesting the reliability of the methods for identification of highly relevant IHFs candidates.

6.2.3 The network growing tools recruit other functionally related IHFs.

To verify that the performance of the network growing tools is neither by chance nor biased by generally highly connected nodes in human interaction networks, we randomized the IHFs into two groups:- 1) the 153 IHFs were randomly distributed into 5 non-redundant groups with ~30 genes each (small set IHFs), and 2) the 1463 IHFs were randomized into 10 non-redundant groups each containing ~146 genes (medium sets) (Appendix A). Similarly, we randomly selected equal numbers of small set and medium set genes from the whole pool of human genes (19,004 genes) obtained from HUGO Gene nomenclature committee (HGNC) database³⁷² (Appendix A). The result showed that in both small (Fig. 6.3a) and medium (Fig. 6.3b) set genes the percentage mean detection of known IHFs was significantly higher with IHF seeds compared to the seeds from random human genes (Fig. 6.3a and 6.3b).

Figure 6.3: Detection rate of network growing tools after randomization of IHFs and random human proteins.

a) Comparison of the detection performance of the network growing tools after growing 100 genes using either small set (30 genes) IHFs or random human proteins as seeds and 1463 IHFs as positive sets. P-values are the result of a paired Student-t test analysis. b) Comparison of the detection performance of the network growing tools after growing 100 genes using either medium set (146 genes) IHFs or random human proteins as seeds and 1463 IHFs as positive sets. P-values are the result of a paired Student-t test analysis.

This suggested that given IHFs as seeds, the detection of other IHFs by using these network growing tools was not likely to be by chance but may indeed reflect relevant functional connections.

6.2.4 Performance of IHFs detection varies between the network growing tools due to different data sources used.

Using the randomized subsets, we could also benchmark statistical significance of the different detection rates of IHFs by the network growing tools confirming highest performance by STRING followed by GeneMANIA and IPA (Fig. 6.4a). The advantage of STRING was most dominant for both the smaller and medium sets, while GeneMANIA had relatively small improvement of performance when the medium set was used (Fig. 6.1b, 6.1c and 6.4a).

To gain more insight on the differences in the performance, we investigated the grown gene differences, and the effect of edges (data sources) and network topological parameters. Pairwise analysis of the grown genes from the small and medium set IHFs seeds from each network growing tool showed that among the 1966 grown genes (from all network growing tools (Appendix A)), only 3.7% were shared by all three tools and high proportions of distinct genes were being grown by IPA (40.3%) followed by GeneMANIA (29.6%) and STRING (7.9%) (Fig. 6.4b). These differences could be due to several reasons, including data sources as well as the growing algorithms used by the three tools. It was difficult to directly compare the tools over the same search space restricted to the same underlying database because their success and performance perceived by the user (the aim of our benchmark) depends critically on the combination of their unique and different sets of edge sources.

Nevertheless, we tried to evaluate the performance of the three tools in a setup of closest possible search space where the original source databases are largely overlapping, which is the case for protein-protein interaction data covered by edge info labels “experiment and database” for STRING, “physical interaction” for GeneMANIA and “experiment” for IPA (Appendix F and Fig. 6.4c). However, also only using this overlapping subset of the source data, there continues to be a difference in performance of the tools. One factor contributing to this is that even in cases where the same original source database was used, e.g. BioGRID, there can be different interpretation and hence consideration of edge data in the different tools. However, STRING and GeneMANIA performed similarly when co-expression data is used as the only data source (supplementary Fig. 6.4c). This suggested that the experimental protein-protein interaction database in STRING may play a big role in its performance.

Figure 6.4: Comparison grown nodes and their edges.

a) Detection performance comparison of the three network growing tools b) pair wise analysis of grown genes c) performance comparison using closest possible data sources d) the effect of edges in STRING e) the effect of edges in GeneMANIA f) comparing the mean score of intersecting IHFs in STRING g) comparing the mean score of intersecting IHFs in GeneMANIA

To study this further, we compared the contribution of each annotated edge type on the respective tools' performance. This analysis was limited to STRING and GeneMANIA as IPA didn't provide results by edge type as direct output. The result showed that analyzing with only "experimental" edge was almost as powerful as the performance of STRING default considering all edge types (Fig. 6.4d). Similarly, using only co-expression had similar performance as the complete GeneMANIA suite, with additional strong individual performances of edge types "Predicted" (protein-protein interaction including annotation transfer from orthologs in other organisms) and "Pathway" (Fig. 6.4e).

Comparison of the mean scores of grown genes intersecting or non-intersecting with known IHFs somewhat reflects comparing known true positives with unknown true plus false positives. Interestingly, GeneMANIA was significantly better than STRING in separating these two sets when intersecting with the smaller set of 153 known IHFs that are shared at least by two studies, while it is the other way round in favour of STRING when intersecting with the complete 1463 known set. This occurs because the genes non-intersecting with the small known set, but predicted by STRING, are in fact good candidates intersecting with the bigger known set (Fig. 6.4f and g).

6.2.5 Network topology analysis shows characteristic differences between networks from different tools.

Network topology analysis indicated that the mean values of several network topological parameters were significantly different between STRING and GeneMANIA. The mean values of average shortest path length, clustering coefficient, eccentricity, radiality and topological coefficient were significantly higher in STRING when compared with GeneMANIA (Appendix F). In contrast, closeness centrality and degree averages were higher in GeneMANIA when compared with STRING (Fig. 6.5a). Upon investigation of the degree distribution, the higher degree average in GeneMANIA originates from a long tail made up of several very high degree nodes (Fig. 6.5a). It was interesting to note higher degrees in GeneMANIA but higher clustering coefficient in STRING. This was, because the degree of a node represents simply the number of immediate neighbours it has, while the clustering coefficient measures the same plus further interconnectivity of these immediate neighbours with each other. The latter could be high when the edge data comprises physical interaction clusters typical for protein complexes which may be a dominant edge source for STRING.

Figure 6.5: Network topology parameters, GOBP and pathways of grown genes.

a) Average network topology parameters of nodes b) Degree distribution of nodes in GeneMANIA and STRING networks c) Significant pathways of the grown genes overlapping with significant pathways of IHFs d) Significant GOBP of the grown genes overlapping with significant GOBP of IHFs

We also computed the topological properties of the intersecting and non-intersecting grown genes, and the result showed that the mean score of betweenness centrality and stress in intersecting grown genes (known IHFs) were consistently higher compared to the non-intersecting grown genes (Appendix F). Nodes with higher stress and betweenness centrality values have a critical role in overall network connectivity and often connect clusters or cellular processes. Future detailed follow-up work will have to establish if this can be exploited to

improve identification of known host factors or it simply reflects that critical hubs are more likely to be found in phenotypic screens which are the source of our known set.

6.2.6 Interchanging and combination of genes didn't improve the performance of the free tools

As many grown genes by the network growing tools are distinct, we tested if combination and interchange of the first 50 genes from GeneMANIA and STRING (free web services/tools) could improve the detection rate of known IHFs. First, we tried combining the first 50 genes of all pairs of the methods to get a total of 100 genes. However, the performance of the combination of genes from the two network growing tools doesn't exceed the performance of 100 genes from STRING alone (Fig. 6.6).

Figure 6.6: The detection rate of STRING and GeneMANIA after combination and interchange of the 1st 50 grown genes.

Next, we tried interchanging gene sets by taking the first 50 genes from one method as a seed along with the randomized 30 IHFs to a second method for growing by another 50 genes to complete 100 grown genes (see materials and methods for detail). The result showed that GeneMANIA with the first 50 from STRING as input seems better compared to GeneMANIA

alone (Fig. 6.6). However, it was again lower than the detection rate of STRING alone, and it appeared that the number of genes contributed by STRING correlated with the strength of performance.

6.2.7 Network growing algorithms detect GO biological process and pathways of the known IHFs

As described above, the network growing tools have the potential to detect IHFs known from siRNA screens at rates as high as 60%. We extended the analysis into benchmarking the performance of the network growing tools at the GO biological process (BP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis level. Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8³⁶⁸ was used for analysis using the seed genes before and after growing 100 genes in the respective network growing tools. As the true positive list of pathways, we defined all those found when using either 153 or the combined set 1463 IHFs as input. We found that the mean percentage of correctly identified pathways after growing the networks was higher in STRING (both from 153 and 1463 IHFs pathways) and GeneMANIA (for 153 IHF pathways) than before growing networks, but not statistically significant (Fig. 6.7a). In contrast, the intersection of GO BP after growing a network was significantly higher compared to before growing the networks (Fig. 6.7b). This is not too surprising, since the tools may include edge sources correlating with GO similarity in the growing process.

When comparing the overlap of KEGG pathways (Fig. 6.7c) and GO BP (Fig. 6.7d) identified after network growing, we observed several shared but also some unique pathways or BPs suggested by the different tools. Matching the underlying biology, several of the intersecting pathways were implicated in supporting different stages of the IAV replication cycle^{279,350,354}. In addition, the three network growing together identified 5 and 40 novel pathways and GO BP respectively (Fig. 6.7c and d; Appendix G) that could be investigated further. It should also be noted that more than 70% pathways from the grown 100 genes in GeneMANIA and STRING were true positives (Fig. 6.7e), which was considerably better than the recall rates of genes (31.4 to 48.4%, Fig. 6.4a). Similar to the KEGG pathways more than 60% of the GO BP of the grown genes in GeneMANIA and STRING were intersecting with true positive set pathways (Fig. 6.7f).

Figure 6.7: Performance comparison of the three network growing tools in terms of KEGG pathways and GO BP.

a) Rate of overlapping KEGG pathways before and after network growing relative to pathways of the 153 and 1463 known IHFs. b) Rate of overlapping GO BP of the grown genes relative to GO BP of the 153 and 1463 known IHFs. c) Pairwise analysis of KEGG pathways from the grown genes in the network growing tools and KEGG pathways of the positive sets (1463 IHFs) d) Pairwise analysis of GO BP from the grown genes in the network growing tools and GO BP of the positive sets (1463 IHFs). e) Significant pathways of the grown genes overlapping with significant pathways of IHFs, f) Significant GOBP of the grown genes overlapping with significant GOBP of IHFs.

6.2.8 Network growing approach on existing data could be as powerful as new siRNA screens for finding relevant additional IHFs

The overlap of IHFs from a single siRNA screen with any of the other siRNA screens ranged between 11.1% and 43.1%, with a mean percentage of 21.5% (Appendix F). However, the mean overlap of the IHFs retrieved from growing networks with any other siRNA screen was 46.7, 33.1 and 18.9 for STRING, GeneMANIA and IPA (Fig. 6.8) respectively. This leads to the interesting conclusion that computational network growing approaches on existing data could be as powerful as new experimental siRNA screens in finding additional relevant IHFs.

Figure 6.8: Comparison of siRNA experimental studies and network growing tools in detecting known IHFs.

6.2.9 The known and the new candidate host factors could be targeted by FDA-approved drugs

We analyzed 1,445 known IHFs and 1,538 newly identified candidate host targets from network growing (Fig. 6.9a) using MetaCore™ for therapeutic and secondary drug interactions³⁵. All drugs that activate any host factor or have both inhibitory and activation effect were excluded from the analysis. A total of 343 FDA-approved drugs that inhibit 218 host factors were identified (Appendix B). Of these, 288 and 108 drugs targeted 147 known and 71 new host factors, respectively (Appendix B). Among the 343 drugs, 53 target both known and new host

factors (Fig. 6.9b). Many drugs (e.g. kinase inhibitors) can target more than one host factor or one host factor can be targeted by many drugs. Overall, 230 host factor targeting FDA approved drugs have been either predicted *in silico* or screened against IAV in previous studies, of which 75 were also predicted by the present study (Fig. 6.9c; Appendix B). Six of the drugs that have been shown with detailed experimental follow-up to have direct or indirect effect on IAV replication through targeting the IHFs have been also predicted by the current study, suggesting the potential of the approach to identify valid new drugs that could be tested (Fig. 6.9c; Appendix B).

Figure 6.9: The known and the new candidate host factors could be targeted by FDA-approved drugs.

a) Recovery rate of anti-viral host factors in known and new candidate IHFs. b) Number of FDA approved drugs that target the known IHFs, candidate new IHFs and both. c) Number of FDA approved drugs that were shared with previous studies and unique to this study. The red circle shows the number of drugs that have already been experimentally validated to have an effect against influenza.

We have to be aware that network growing provides the functional context of genes being connected but not exactly how they influence each other. For viral infections, these genes could be true host factors (required for the virus) or suppressors of host factors acting as antivirals which would have the opposite effect of helping the virus when targeted. We compiled a list of experimentally identified antiviral genes from the siRNA screens (Appendix B) and found that network growing as well as also experimental screens themselves would identify 2.3% or 1.2% of these antiviral genes with their respective approaches (Fig. 6.9a). This serves as a warning that although small in proportions, host networks may include antivirals that should not be targeted. In addition, although the pathways from the new IHFs were indicated to support different stages of IAV replication cycle^{279,350,354}, it is worth to validate their effect on IAV. Among the 204 predicted FDA approved drugs by De Chassey et al., 2012²⁷⁶ and 60 molecules by Watanabe et

al., 2014³⁵⁴ (in silico against IAV); 71 and eight drugs were also suggested by the MetaCore analysis respectively (Appendix B). While this shows that the approach is in agreement with other studies, in general, it also highlights the diversity of available drug to gene target annotations in different databases (MetaCore in our case, different sources in others). It has been suggested that integration of different databases could improve drug discovery⁶⁰⁸.

6.2.10 Ranking the FDA-approved drugs identified in this study

Finally, we ranked the 343 FDA-approved drugs based on the seven criteria: killing index, number of side effects, cost of the drugs; number of the interactions between IHFs, number of IHFs targeted by a drug; and the co-occurrence likelihood ratio of drug-target and target-influenza in PubMed abstracts (see materials & methods). The top 10 ranked drugs were described in Table 6.1, among which the anti-hypertensive (Losartan), anti-malaria (Chloroquine), and the anti-influenza (Zanamivir) have been shown to have effect on IAV replication or virion release, respectively. The angiotensin II receptor, type 1 (AGTR1) which is negatively regulated by ACE2 (angiotensin I converting enzyme 2) was highly induced in IAV (H7N9) infected mice and patients, and its inhibition by losartan (also predicted drug in this study) reduced severity, lung injury and virus titer in H7N9 infected mice⁶⁰⁹. Losartan also reduced lung injury in H5N1 infected mice through inhibition of inflammatory cytokines (IL-6)⁶¹⁰. Similarly, chloroquine has been shown to have significant anti-influenza activity, however the reports are controversial. Ooi et al., 2006 indicated that chloroquine inhibited the replication of H1N1 and H3N2 viruses in MDCK cells⁶¹¹. Similarly, Yan et al., 2013 reported the effectiveness of chloroquine in treating H5N1 virus infections in A549 cells and mouse model without significantly affecting the viral load and inflammatory cytokines⁶¹². Conversely, Wu et al., 2015 demonstrated that chloroquine enhances the replication of H1N1/WSN virus in A549 cells⁶¹³. These examples suggest that the new candidate IHFs have reasonable potential to have similar phenotypic effects on IAV replication as the known IHFs they are connected to, and both can be targeted by the existing FDA approved drugs. Finally, our network growing suggests a list of ranked drugs for further investigations as potential anti-IAV therapy.

Table 6.1: The top 10 FDA-approved drugs of the 343 drugs that target the new candidate and known IHFs (ranked based on different criteria)

Drug	Target	Target_name	Group	Z_score_ Total
Losartan*	<i>AGTR1</i>	angiotensin II receptor, type 1	Antihypertensive	18.2258
Piperacillin	<i>LACTB</i> ⁺	lactamase, beta	Anti-Bacterial	10.19687
Valsartan	<i>AGTR1</i>	angiotensin II receptor, type 1	Antihypertensive	8.075544
Telmisartan	<i>AGTR1</i>	angiotensin II receptor, type 1	Cardiovascular System	5.988901
Chloroquine*	<i>TLR9</i>	toll-like receptor 9	Antimalarials	5.617846
Milrinone	<i>PDE3A</i>	phosphodiesterase 3A, cGMP-inhibited	Platelet Aggregation Inhibitors	4.548284
Zanamivir*	<i>NEUI</i> ⁺	sialidase 1	Antiviral (Influenza)	4.108547
Prednisolone	<i>P2RY12</i> ⁺	purinergic receptor P2Y G-protein coupled 12	Anti-Inflammatory	3.88681
Methadone	<i>CYP2J2</i> ⁺	cytochrome P450, family 2, subfamily J, polypeptide 2	Analgesics	3.675142
Diflunisal	<i>TTR</i>	transthyretin	Anti-Inflammatory, Non-Steroidal	3.490854

*previously tested against IAV, ⁺Known IHFs

6.3 Discussion

In this study, we compared the performance of selected computational gene network growing tools with each other using a large set of 1,463 experimentally determined influenza host factors. This set is well-suited since the siRNA screen data has not yet directly been exploited by the tools, and the genes were found to cover a broad range of representative cellular functions. Considering the experimentally determined genes as true positives, the detection percentages are equivalent to a measure of precision or positive predictive value of the methods.

We need to emphasize that network growing tool performance is intrinsically linked to the respective underlying data types and data sources used, which is also apparent in our benchmark results. Besides the differences in underlying databases between the tools, we have also observed considerable performance fluctuations of different versions of the same tool after database updates which highlight the clear dependency of this approach on the available search space.

Furthermore, another main difference between the tools was that they not only use protein-protein interactions, but they also use other interaction/edge types such as co-expression, gene neighbourhood, co-localization, shared domains, text-mining, genetic interaction etc which are mostly unique to the different tools. Since the contribution of the different edge type sources may be of interest to understand why certain tools perform better, we provided a detailed comparison of the individual edge type contributions to prediction performance and found that

protein-protein interaction data in STRING seems to play a dominant role in overall performance while for GeneMANIA co-expression is the main contributor. Interestingly, using co-expression alone performed equally well when used through STRING or GeneMANIA which suggested that the basic algorithms used to recruit genes may perform similar if the same underlying data type and set would be used. However, each tool also uses different interpretations for reliability and rules applied to define edges even from the same data set. Network topology analysis further revealed characteristic differences of the networks produced by the two tools in parameters such as average degree, closeness centrality and cluster coefficient mostly reflecting the differences in approach and underlying source data. Interestingly, some network parameters seem to distinguish between true hits and assumed false positives, which will require further testing, but could inspire future method development.

There may be concerns that the network growing approach could select highly connected genes in general and not necessarily those specifically related to the input set. To confirm that this is not the case, we compared the network growing performance of the influenza-related host seed genes with that of random human genes and find a significant difference in favour of growing influenza-related rather than the highly connected but random genes.

Although there are considerable numbers of distinct genes suggested uniquely by some methods but not others, the interchange and combination of genes recruited from GeneMANIA and STRING did not improve the performance of host factor detection rates. Although the performance was not improved, interpretation of results from a single method may lead to biases in interpretations and considering multiple methods could uncover a more diverse and complete set of host factors albeit at the same time diluted with false positives.

Overall, the network growing tools had better recall rates at GO BP and KEGG pathway level than for genes. This was expected, since several of the tools directly use this information for growing networks and it was also reported previously that influenza host factors (IHF) identified with siRNA screens in different studies have higher overlap at functional categories than at gene level²⁷⁷.

The main goal of the previous influenza siRNA screens was to identify druggable host targets. There is currently a limited availability of drugs to treat and prevent influenza virus infection, which is compounded by high mutation rates that lead to drug resistance. Drug repurposing is a recent strategy that has been proposed as one possible solution to this problem.

In this strategy, existing drugs that are in clinical use and that have known safety profiles are evaluated for either anti-viral activity or as therapeutics to treat virus infection. System-wide phenotypic data together with bioinformatics methods have been used for *in silico* prediction for repurposing FDA approved drugs as alternative therapeutics against infectious diseases and cancer^{343,344}. Therefore, it would be plausible that network growing could identify new host factors that could be suggested as anti-influenza targets. Indeed, applying this approach on newly identified candidate host targets from network growing suggested 258 new predictions for existing FDA-approved drugs for further investigations as potential anti-IAV therapy (Appendix B). Importantly, six drugs that have previously been shown experimentally to have an effect on influenza virus were also identified by our investigation, thereby adding support to the validity of the approach (Fig. S4c, supplementary Data 3)³⁵⁴. In addition, our systematic ranking of the drugs using different criteria would help to focus on certain relevant drugs for further experimental screening.

6.4 Conclusion

Our aim was to establish the performance of selected network growing tools in a typical real usage scenario to recover known influenza host factors from siRNA screens not included as source for the tools. All 3 tools tested are able to do so significantly better than expected by chance and even to a similar extent as conducting a new experimental siRNA screen. A detailed comparison of network topology and performance contribution of different data types used by individual tools highlighted that the network growing algorithms crucially depend on their underlying databases and data types that are used. While we used the system of influenza host interactions here, we believe that these network growing tools would be similarly able to recover relevant gene/protein connections also for other biological systems, viruses and diseases.

Chapter 7: General conclusions

In this study, we highlighted the gene expression signatures of different cell types (A549, CEF, MDCK and PM Φ) during infections with different IAV isolates (H1N1/WSN, pH1N1, H5N2, H5N3, H7N9 and H9N2). The result showed that although there were few exceptions, generally we observed that the expression signatures of different IAV isolates in a single cell type are more close to each other than the expression signature of a single isolate in different cell types. Type I and III IFN responses were prominent in A549 cells virus infections, while the *IFNB* response was strain-specific in CEF and MDCK cells. No type III IFN responses were induced in CEF and MDCK cells. However, H5N3 virus showed unique expression characteristics that it induced type I IFN in MDCK cells, but not in A549 and CEF cells infections. Regardless of the IFN responses in A549 cells, H5N3 virus is able to induce ISGs and inflammatory cytokines. The unique characteristics of the H5N3 virus were also observed in PM Φ infections. During PM Φ infections, IFN responses were not differentially regulated in H5N2 and H5N3 virus infections, however, increased expression of ISGs and inflammatory cytokines being prominent in H5N3 virus infections. ISGs are mainly induced by increased level of IFNs, but our observation suggested that there could be another alternative pathway for induction of ISGs.

Looking into details, we have identified highly expressed cell-specific signatures in CEF cells (e.g. *OSBPL1A*, *ARHGAP21*, *MRPS27*, *COPS2*) which would probably increase IAV replication while maintaining the cellular physiology of CEF cells. Conversely, increased antiviral, pro-apoptotic and inflammatory signatures have been identified in A549 cells. The antiviral and inflammatory responses inhibit IAV replication, but partially affecting the survival of the cell. Except in H5N3 and H9N2 viruses, infection of the other viruses down-regulates genes associated with cell cycle and metabolism with the strongest effect in MDCK cells, followed by A549 cells in strain dependent manner, but not in CEF cells. To the reverse, H5N3 virus induces metabolic pathways and ribosome pathway in MDCK cell infections. Integration of the DEGs with siRNA screens and interactome showed that cell cycle, ribosome and metabolic pathways were important for the survival of the cell during IAV infections and virion productions.

While H5N3 virus maintains or induces metabolic and pro-viral pathways in mammalian cell lines (A549 and MDCK cells), H1N1/WSN virus dampen these pathways in the same cell lines. Lack of IFN induction in A549 cells, activation of pro-viral metabolic and other pathways, and lack of PB1-F2 which induced apoptosis might contribute for increased replication of the H5N3 virus in the three cell lines (A549, CEF, and MDCK cells). Although the H5N3 virus showed higher replication rate in the three cell lines and induced a higher level of inflammatory cytokines in PM Φ , the virus caused asymptomatic infections in mice. Hence, we hypothesise that during IAV infections overexpression of mitochondrial proteins that maintain the bioenergetics and prevent apoptosis supports cell survival and IAV replication without pathologic effects. To the reverse, IAV that destruct mitochondrial bioenergetics and induce apoptosis may result in pathologic infections.

The current study also highlighted plausible IHFs that could be targeted against IAV through the integration of conserved up-regulated genes in the A549 cells infections with host factors from the siRNA screens and interactome. These plausible host factors could support replication of the virus at different stages of life cycle (e.g. immune response, endocytosis). It was expected that IAV infections up-regulate genes or pathways related to innate immune response, however, we have noted few pro-viral genes were also part of the innate immune responses (e.g. TNF signalling, cytokine-cytokine receptor interactions and apoptosis). Among the plausible IHFs the *TRIM21*, *TRIM26*, *SAMHD1*, *IRF2*, *PIK3R3*, *JAK2* and *EIF2AK2* negatively regulates IFN or immune responses, while *TIMM17A*, *IL15RA*, *BCL3*, and *CARD16* prevents apoptosis. Thus we suggest that host responses that support cell survival either through negative regulation of immune response or apoptosis are important for the replication of IAVs. The detailed mechanism of these plausible host factors how they support IAV replication should be validated in future.

Benchmarking of the gene network growing tools using IHFs revealed that network growing tools detect significantly more known host factors than random human genes suggesting functional interconnectedness of the IHFs. STRING performs better than GeneMANIA and IPA in detecting the known IHFs and their associated pathways. Differences in the performance of the three tools were due to different data types used by individual tools and their underlying databases. Hence, we recommend that in future, the database or gene network growing tool

developers should consider such issues, so users will benefit from more comprehensive and consistent results. Interestingly, the rate of identification of true host factors using computational network growing is equal to or better than doing another experimental siRNA screen. It may be interesting to apply the approach also to study other biological pathways/processes.

Through growing gene networks using the known IHFs, we have also identified hundreds of other genes that were closely connected to the known IHFs. As closely connected genes are functionally related, we claim that these genes are "new candidate" IHFs that would have been missed by experimental methods. In future, experimental validation of this new candidate host factors towards the replication of IAV would maximise the benefit of network growing tools. In addition, experimental validation of the FDA approved drugs that targeted the known and new candidate IHFs will accelerate the discovery of alternative drugs for IAV therapy.

Chapter 8: Future work

8.1 Wet lab validations

Our computational analysis of the IAV-host interactions provided insights for future wet lab validation of the following:

a) Validation of host gene expression changes in primary cell lines of the upper and lower respiratory tract (e.g. human airway epithelial (HAE) and human bronchial epithelial (NHBE) cells lines). In addition, recent advances in stem cell technology can provide the opportunity to investigate the host gene expression changes in lung cell lines. Furthermore it is also important to investigate the expression changes in 2 dimension (2D) cell culture vs 3D cell culture as recent studies showed expression differences between the two systems. On top of this, single cell transcriptome analysis is also important as the number of viruses infecting a single cell may vary in cell culture.

b) From this study, we have observed that in some cases, although there is an absence or reduced level of IFN responses, the expressions of ISGs were very high in most of the virus infections (in A549 and Macrophage infections). This suggests that there could be an alternative pathway that could activate ISGs in IAV infections. Hence, this alternative pathway should be verified in future.

c) From the integration analysis (Chapter 3) we have identified plausible host factors that could be targeted as IAV therapeutics. The mechanism of these host factors how they support or affect replication of IAV needs validation. This can be achieved by knocking in/out of the identified plausible host factors in A549 cells, primary cell lines and also in animal models (e.g. mice) and measuring the level of the viruses through different virus assays.

d) In chapter 4, we have also identified important host species signatures that could help the replication of IAV while not affecting the overall cellular physiology (e.g. apoptosis and metabolic pathways). These shortlisted expression signatures should also be validated using knock in/out experiments in the three cell types, primary cell lines of the URT and LRT of different species. Although we were able to identify virus and species-specific gene expression signatures, our study is not without limitation. We have used different cell types of different species. Therefore, studies that involve similar cell types (lung or blood) across several host species (e.g. human, chicken, duck, swine etc), will be beneficial for a better understanding of

the host adaptation mechanisms. On top of these, we suggest detailed experimental investigation on the role of the identified species-specific signatures.

e) This study demonstrated that H5N3 virus had unique characteristics in that in A549 and MΦ it induces higher inflammatory responses compared to other viruses. In addition, it had unique property that it up-regulates most of the metabolic pathways in MDCK cells, unlike the other tested virus strains. In addition, H9N2 virus had weak expression signatures in all of the cell types. Hence, in future, it is important to characterise the sequence motifs that are responsible for this unique property of these viruses.

f) The network growing analysis (Chapter 6), suggested two main hypotheses that could be validated experimentally. Firstly, the new candidate IHFs that were identified through growing network should be validated using siRNA screen targeting them. Secondly, FDA-approved drugs which are not tested in the context of IAV before should be screened in cell lines (e.g. A549, primary cell lines) and animal models for their efficacy against IAVs.

g) The inclusion of other clinical isolates of IAVs (seasonal H1N1, H3N2, HPAI H5N1, H7N9) in all of the above-mentioned experimental validations needs to be considered.

8.2 Computational future work

a) Computational investigation of host-specific gene expression signatures during IAVs in different species using the public available data in NCBI GEO data set (lung or blood) across several host species (e.g. human, chicken, duck, swine etc). This will validate the shortlisted signatures from the cell lines (our study) and give insight for a better understanding of the host adaptation mechanisms.

b) Another important aspect to consider in the future would be population genetics study based on the IHFs (pro-viral) and antiviral host factors. Determination of populations with single nucleotide polymorphisms (SNPs) that reduce the activity of IHFs would identify parts of the population that would be resistant to IAV infections (diseases). On the other hand, people with SNPs that deactivate antiviral host factors could be highly susceptible to influenza infections and severe disease. Therefore, determining the prevalence of SNPs in these host factors along with the distribution of influenza A virus would help in understanding host genetic determinants for the epidemiology of the virus.

References

- 1 Ng, W. C. *et al.* The C-type lectin langerin functions as a receptor for attachment and infectious entry of influenza A virus. *Journal of Virology* **90**, 206-221, doi:10.1128/JVI.01447-15 (2016).
- 2 Fraaij, P. L. A. & Osterhaus, A. D. M. E. in *Hot Topics in Respiratory Medicine* 7-14 (2010).
- 3 Koutsakos, M., Nguyen, T. H., Barclay, W. S. & Kedzierska, K. Knowns and unknowns of influenza B viruses. *Future Microbiology* **11**, 119-135, doi:10.2217/fmb.15.120 (2016).
- 4 Matsuzaki, Y. & Nakamura, K. The virological, epidemiological and clinical features of influenza A, B and C viruses. *Nippon rinsho. Japanese journal of clinical medicine* **55**, 2515-2520 (1997).
- 5 Muraki, Y. & Hongo, S. The molecular virology and reverse genetics of influenza C virus. *Japanese Journal of Infectious Diseases* **63**, 157-165 (2010).
- 6 Markowska-Daniel, I., Mickiewicz, M., Witkowski, L. & Kita, J. Characterization of a new influenza virus type D. *Medycyna Weterynaryjna* **72**, 531-535, doi:10.21521/mw.5565 (2016).
- 7 Killingley, B. & Nguyen-Van-Tam, J. Routes of influenza transmission. *Influenza and other respiratory viruses* **7 Suppl 2**, 42-51, doi:10.1111/irv.12080 (2013).
- 8 Lessler, J. *et al.* Incubation periods of acute respiratory viral infections: a systematic review. *The Lancet. Infectious diseases* **9**, 291-300, doi:10.1016/S1473-3099(09)70069-6 (2009).
- 9 Saxena, S. *et al.* Clinical characterization of influenza A and human respiratory syncytial virus among patients with influenza like illness. *Journal of Medical Virology* **89**, 49-54, doi:10.1002/jmv.24607 (2017).
- 10 Li, H. & Cao, B. Pandemic and Avian Influenza A Viruses in Humans: Epidemiology, Virology, Clinical Characteristics, and Treatment Strategy. *Clinics in Chest Medicine* **38**, 59-70, doi:10.1016/j.ccm.2016.11.005 (2017).
- 11 Vemula, S. V. *et al.* Current Approaches for Diagnosis of Influenza Virus Infections in Humans. *Viruses* **8**, 96, doi:10.3390/v8040096 (2016).
- 12 Simonsen, L., Viboud, C., Taylor, R. J. & Miller, M. A. in *Birkhauser Advances in Infectious Diseases* 27-54 (2011).
- 13 Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host and Microbe* **7**, 440-451, doi:10.1016/j.chom.2010.05.009 (2010).
- 14 Dawood, F. S. *et al.* Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *The Lancet. Infectious diseases* **12**, 687-695, doi:10.1016/S1473-3099(12)70121-4 (2012).
- 15 Westgeest, K. B. *et al.* Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011. *Journal of Virology* **88**, 2844-2857, doi:10.1128/JVI.02163-13 (2014).
- 16 Jhung, M. A. *et al.* Outbreak of variant influenza A(H3N2) virus in the United States. *Clinical Infectious Diseases* **57**, 1703-1712 (2013).

- 17 Chutinimitkul, S. *et al.* The spread of avian influenza H5N1 virus; a pandemic threat to mankind. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet* **89 Suppl 3**, S218-233 (2006).
- 18 Duan, L. *et al.* Characterization of low-pathogenic H5 subtype influenza viruses from Eurasia: implications for the origin of highly pathogenic H5N1 viruses. *J Virol* **81**, 7529-7539, doi:10.1128/JVI.00327-07 (2007).
- 19 Richard, M., De Graaf, M. & Herfst, S. Avian influenza A viruses: From zoonosis to pandemic. *Future Virology* **9**, 513-524 (2014).
- 20 Lindh, E. *et al.* Molecular epidemiology of H9N2 influenza viruses in Northern Europe. *Veterinary Microbiology* **172**, 548-554 (2014).
- 21 DeLay, P. D., Casey, H. L. & Tubiash, H. S. Comparative study of fowl plague virus and a virus isolated from man. *Public health reports* **82**, 615-620 (1967).
- 22 Campbell, C. H., Webster, R. G. & Breese, S. S., Jr. Fowl plague virus from man. *Journal of Infectious Diseases* **122**, 513-516, doi:10.1093/infdis/122.6.513 (1970).
- 23 Iuliano, A. D. *et al.* Increase in Human Infections with Avian Influenza A(H7N9) Virus During the Fifth Epidemic - China, October 2016-February 2017. *MMWR. Morbidity and mortality weekly report* **66**, 254-255, doi:10.15585/mmwr.mm6609e2 (2017).
- 24 Horimoto, T. & Kawaoka, Y. Pandemic threat posed by avian influenza A viruses. *Clinical Microbiology Reviews* **14**, 129-149 (2001).
- 25 Gong, Z. *et al.* Epidemiology of the avian influenza A (H7N9) outbreak in Zhejiang Province, China. *BMC Infectious Diseases* **14** (2014).
- 26 Li, Q. *et al.* Epidemiology of human infections with avian influenza A(H7N9) virus in China. *New England Journal of Medicine* **370**, 520-532 (2014).
- 27 Mao, L., Yang, Y., Qiu, Y. & Yang, Y. Annual economic impacts of seasonal influenza on US counties: Spatial heterogeneity and patterns. *International Journal of Health Geographics* **11** (2012).
- 28 Akunzule, A. N., Koney, E. B. M. & Tiongco, M. Economic impact assessment of highly pathogenic avian influenza on the poultry industry in Ghana. *World's Poultry Science Journal* **65**, 517-528 (2009).
- 29 Basuno, E., Yusdja, Y. & Ilham, N. Socio-economic impacts of avian influenza outbreaks on small-scale producers in Indonesia. *Transboundary and Emerging Diseases* **57**, 7-10 (2010).
- 30 Buetre, B., Kim, Y., Tran, Q. T., Thomson, J. & Gunasekera, D. Avian influenza: Potential economic impact of a pandemic on Australia. *Australian Commodities* **13**, 351-359 (2006).
- 31 Alexander, D. J. in *Poultry Diseases* 317-332 (2008).
- 32 Palese, P. & Young, J. F. Variation of influenza A, B, and C viruses. *Science* **215**, 1468-1474 (1982).
- 33 Zhirnov, O. P. & Grigoriev, V. B. Disassembly of influenza C viruses, distinct from that of influenza A and B viruses requires neutral-alkaline pH. *Virology* **200**, 284-291 (1994).
- 34 Yamashita, M., Krystal, M. & Palese, P. Comparison of the three large polymerase proteins of influenza A, B, and C viruses. *Virology* **171**, 458-466 (1989).
- 35 Hause, B. M. *et al.* Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *mBio* **5**, e00031-00014, doi:10.1128/mBio.00031-14 (2014).

- 36 Ferguson, L. *et al.* Pathogenesis of Influenza D Virus in Cattle. *J Virol* **90**, 5636-5642, doi:10.1128/JVI.03122-15 (2016).
- 37 Noda, T. Native morphology of influenza virions. *Frontiers in Microbiology* **2** (2012).
- 38 Dadonaite, B., Vijayakrishnan, S., Fodor, E., Bhella, D. & Hutchinson, E. C. Filamentous influenza viruses. *Journal of General Virology* **97**, 1755-1764, doi:10.1099/jgv.0.000535 (2016).
- 39 Campbell, P. J. *et al.* The M segment of the 2009 pandemic influenza virus confers increased neuraminidase activity, filamentous morphology, and efficient contact transmissibility to a/puerto rico/8/1934-based reassortant viruses. *Journal of Virology* **88**, 3802-3814 (2014).
- 40 Elton, D. *et al.* The genetics of virus particle shape in equine influenza A virus. *Influenza and other Respiratory Viruses* **7**, 81-89 (2013).
- 41 Louisirirothanakul, S. *et al.* Electron micrographs of human and avian influenza viruses with high and low pathogenicity. *Asian Biomedicine* **7**, 155-167 (2013).
- 42 Roberts, K. L., Leser, G. P., Ma, C. & Lamb, R. A. The amphipathic helix of influenza A virus M2 protein is required for filamentous bud formation and scission of filamentous and spherical particles. *Journal of Virology* **87**, 9973-9982 (2013).
- 43 Rossman, J. S. *et al.* Influenza virus M2 ion channel protein is necessary for filamentous virion formation. *Journal of Virology* **84**, 5078-5088 (2010).
- 44 Roberts, P. C., Lamb, R. A. & Compans, R. W. The M1 and M2 proteins of influenza A virus are important determinants in filamentous particle formation. *Virology* **240**, 127-137 (1998).
- 45 Bruce, E. A., Digard, P. & Stuart, A. D. The Rab11 pathway is required for influenza A virus budding and filament formation. *Journal of Virology* **84**, 5848-5859, doi:10.1128/JVI.00307-10 (2010).
- 46 Nelson, M. I. & Holmes, E. C. The evolution of epidemic influenza. *Nature Reviews Genetics* **8**, 196-205, doi:10.1038/nrg2053 (2007).
- 47 Zheng, W. & Tao, Y. J. Structure and assembly of the influenza A virus ribonucleoprotein complex. *FEBS Letters* **587**, 1206-1214 (2013).
- 48 Vasin, A. V. *et al.* Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins. *Virus Research* **185**, 53-63 (2014).
- 49 Abdelwhab, E. M., Veits, J. & Mettenleiter, T. C. Avian influenza virus NS1: A small protein with diverse and versatile functions. *Virulence* **4**, 583-588 (2013).
- 50 Li, M. L., Rao, P. & Krug, R. M. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. *EMBO Journal* **20**, 2078-2086 (2001).
- 51 Nilsson, B. E., te Velhuis, A. J. W. & Fodor, E. Role of the PB2 627 domain in influenza A virus polymerase function. *Journal of Virology* **91**, doi:10.1128/JVI.02467-16 (2017).
- 52 Hatakeyama, D. *et al.* A novel functional site in the PB2 subunit of influenza A virus essential for acetyl-CoA interaction, RNA polymerase activity, and viral replication. *Journal of Biological Chemistry* **289**, 24980-24994, doi:10.1074/jbc.M114.559708 (2014).
- 53 Long, J. C. D. & Fodor, E. The PB2 subunit of the influenza A virus RNA polymerase is imported into the mitochondrial matrix. *Journal of Virology* **90**, 8729-8738, doi:10.1128/JVI.01384-16 (2016).

- 54 Toyoda, T., Adyshev, D. M., Kobayashi, M., Iwata, A. & Ishihama, A. Molecular assembly of the influenza virus RNA polymerase: Determination of the subunit-subunit contact sites. *Journal of General Virology* **77**, 2149-2157 (1996).
- 55 Binh, N. T., Wakai, C., Kawaguchi, A. & Nagata, K. The N-terminal region of influenza virus polymerase PB1 adjacent to the PA binding site is involved in replication but not transcription of the viral genome. *Frontiers in Microbiology* **4** (2013).
- 56 Binh, N. T., Wakai, C., Kawaguchi, A. & Nagata, K. Involvement of the N-terminal portion of influenza virus RNA polymerase subunit PB1 in nucleotide recognition. *Biochemical and Biophysical Research Communications* **443**, 975-979 (2014).
- 57 Yoshizumi, T. *et al.* Influenza A virus protein PB1-F2 translocates into mitochondria via Tom40 channels and impairs innate immunity. *Nature Communications* **5**, doi:10.1038/ncomms5713 (2014).
- 58 Vidic, J. *et al.* Amyloid assemblies of influenza A virus PB1-F2 protein damage membrane and induce cytotoxicity. *Journal of Biological Chemistry* **291**, 739-751, doi:10.1074/jbc.M115.652917 (2016).
- 59 Shin, N., Pyo, C. W., Jung, K. I. & Choi, S. Y. Influenza A virus PB1-F2 is involved in regulation of cellular redox state in alveolar epithelial cells. *Biochemical and Biophysical Research Communications* **459**, 699-705, doi:10.1016/j.bbrc.2015.03.010 (2015).
- 60 Pinar, A. *et al.* PB1-F2 peptide derived from avian influenza A virus H7N9 induces inflammation via activation of the NLRP3 inflammasome. *Journal of Biological Chemistry* **292**, 826-836, doi:10.1074/jbc.M116.756379 (2017).
- 61 Obayashi, E. *et al.* The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* **454**, 1127-1131 (2008).
- 62 Gao, H. *et al.* Twenty amino acids at the C-terminus of PA-X are associated with increased influenza A virus replication and pathogenicity. *Journal of General Virology* **96**, 2036-2049, doi:10.1099/vir.0.000143 (2015).
- 63 Hayashi, T., Chaimayo, C., McGuinness, J. & Takimoto, T. Critical Role of the PA-X C-Terminal Domain of Influenza A Virus in Its Subcellular Localization and Shutoff Activity. *Journal of Virology* **90**, 7131-7141, doi:10.1128/JVI.00954-16 (2016).
- 64 Hu, J. *et al.* PA-X-associated early alleviation of the acute lung injury contributes to the attenuation of a highly pathogenic H5N1 avian influenza virus in mice. *Medical Microbiology and Immunology* **205**, 381-395, doi:10.1007/s00430-016-0461-2 (2016).
- 65 Hu, J. *et al.* PA-X decreases the pathogenicity of highly pathogenic H5N1 influenza A virus in avian species by inhibiting virus replication and host response. *Journal of Virology* **89**, 4126-4142, doi:10.1128/JVI.02132-14 (2015).
- 66 Zhang, W. Q., Song, S. X. & Wang, T. Z. [Advances in the structure and function of pandemic A/H1N1/2009 influenza virus HA protein]. *Bing du xue bao = Chinese journal of virology / [bian ji, Bing du xue bao bian ji wei yuan hui]* **28**, 444-452 (2012).
- 67 Bertram, S., Glowacka, I., Steffen, I., Köhl, A. & Pöhlmann, S. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. *Reviews in Medical Virology* **20**, 298-310 (2010).
- 68 Colman, P. M., Varghese, J. N. & Laver, W. G. Structure of the catalytic and antigenic site in influenza virus neuraminidase. *Nature* **303**, 41-44 (1983).
- 69 Wan, H. *et al.* Molecular basis for broad neuraminidase immunity: Conserved epitopes in seasonal and pandemic H1N1 as well as H5N1 influenza viruses. *Journal of Virology* **87**, 9290-9300 (2013).

- 70 Hooper, K. A. & Bloom, J. D. A mutant influenza virus that uses an N1 neuraminidase as the receptor-binding protein. *Journal of Virology* **87**, 12531-12540 (2013).
- 71 Noton, S. L. *et al.* Identification of the domains of the influenza A virus M1 matrix protein required for NP binding, oligomerization and incorporation into virions. *Journal of General Virology* **88**, 2280-2290 (2007).
- 72 Bui, M., Wills, E. G., Helenius, A. & Whittaker, G. R. Role of the influenza virus M1 protein in nuclear export of viral ribonucleoproteins. *Journal of Virology* **74**, 1781-1786 (2000).
- 73 Chen, B. J., Leser, G. P., Jackson, D. & Lamb, R. A. The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding. *Journal of Virology* **82**, 10059-10070 (2008).
- 74 Jing, X. *et al.* Functional studies indicate amantadine binds to the pore of the influenza A virus M2 proton-selective ion channel. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10967-10972 (2008).
- 75 Rossman, J. S., Jing, X., Leser, G. P. & Lamb, R. A. Influenza Virus M2 Protein Mediates ESCRT-Independent Membrane Scission. *Cell* **142**, 902-913 (2010).
- 76 McCown, M. F. & Pekosz, A. Distinct domains of the influenza A virus M2 protein cytoplasmic tail mediate binding to the M1 protein and facilitate infectious virus production. *Journal of Virology* **80**, 8178-8189 (2006).
- 77 Sugrue, R. J. & Hay, A. J. Structural characteristics of the M2 protein of influenza A viruses: Evidence that it forms a tetrameric channel. *Virology* **180**, 617-624 (1991).
- 78 Hale, B. G., Randall, R. E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *Journal of General Virology* **89**, 2359-2376 (2008).
- 79 Kim, J. I. *et al.* The PDZ-binding motif of the avian NS1 protein affects transmission of the 2009 influenza A(H1N1) virus. *Biochemical and Biophysical Research Communications* **449**, 19-25 (2014).
- 80 Chen, G., Liu, C. H., Zhou, L. & Krug, R. M. Cellular DDX21 RNA helicase inhibits influenza A virus replication but is counteracted by the viral NS1 protein. *Cell Host and Microbe* **15**, 484-493 (2014).
- 81 Aragon, T. *et al.* Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus. *Molecular and Cellular Biology* **20**, 6259-6268 (2000).
- 82 O'Neill, R. E., Talon, J. & Palese, P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO Journal* **17**, 288-296 (1998).
- 83 Huang, S. *et al.* A second CRM1-dependent nuclear export signal in the influenza A virus NS2 protein contributes to the nuclear export of viral ribonucleoproteins. *Journal of Virology* **87**, 767-778 (2013).
- 84 Chen, J., Huang, S. & Chen, Z. Human cellular protein nucleoporin hNup98 interacts with influenza A virus NS2/nuclear export protein and overexpression of its GLFG repeat domain can inhibit virus propagation. *Journal of General Virology* **91**, 2474-2484 (2010).
- 85 Akarsu, H. *et al.* Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO Journal* **22**, 4646-4655 (2003).
- 86 Paterson, D. & Fodor, E. Emerging Roles for the Influenza A Virus Nuclear Export Protein (NEP). *PLoS Pathogens* **8** (2012).
- 87 Odagiri, T., Tominaga, K., Tobita, K. & Ohta, S. An amino acid change in the non-structural NS2 protein of an influenza A virus mutant is responsible for the generation of

- defective interfering (DI) particles by amplifying DI RNAs and suppressing complementary RNA synthesis. *Journal of General Virology* **75**, 43-53 (1994).
- 88 Selman, M., Dankar, S. K., Forbes, N. E., Jia, J. J. & Brown, E. G. Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerging Microbes and Infections* **1**, doi:10.1038/emi.2012.38 (2012).
- 89 Ng, A. K. L. *et al.* Structure of the influenza virus A H5N1 nucleoprotein: Implications for RNA binding, oligomerization, and vaccine design. *FASEB Journal* **22**, 3638-3647 (2008).
- 90 Ortega, J. *et al.* Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *Journal of Virology* **74**, 156-163 (2000).
- 91 Chan, W. H. *et al.* Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *Journal of Virology* **84**, 7337-7345 (2010).
- 92 Marklund, J. K., Ye, Q., Dong, J., Tao, Y. J. & Krug, R. M. Sequence in the influenza A virus nucleoprotein required for viral polymerase binding and RNA synthesis. *Journal of Virology* **86**, 7292-7297 (2012).
- 93 Yu, M. *et al.* Identification and characterization of three novel nuclear export signals in the influenza A virus nucleoprotein. *Journal of Virology* **86**, 4970-4980 (2012).
- 94 Liu, Y., Lou, Z., Bartlam, M. & Rao, Z. Structure-function studies of the influenza virus RNA polymerase PA subunit. *Science in China, Series C: Life Sciences* **52**, 450-458 (2009).
- 95 Chen, W. *et al.* A novel influenza A virus mitochondrial protein that induces cell death. *Nature Medicine* **7**, 1306-1312, doi:10.1038/nm1201-1306 (2001).
- 96 Kamal, R. P. *et al.* Emergence of highly pathogenic avian influenza A(H5N1) virus PB1-F2 variants and their virulence in BALB/c mice. *Journal of Virology* **89**, 5835-5846, doi:10.1128/JVI.03137-14 (2015).
- 97 Ye, Q., Krug, R. M. & Tao, Y. J. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* **444**, 1078-1082, doi:10.1038/nature05379 (2006).
- 98 Te Velthuis, A. J. W. & Fodor, E. Influenza virus RNA polymerase: Insights into the mechanisms of viral RNA synthesis. *Nature Reviews Microbiology* **14**, 479-493, doi:10.1038/nrmicro.2016.87 (2016).
- 99 Stevens, J. *et al.* Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* **312**, 404-410, doi:10.1126/science.1124513 (2006).
- 100 Russell, R. J. *et al.* The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* **443**, 45-49, doi:10.1038/nature05114 (2006).
- 101 Harris, A., Forouhar, F., Qiu, S., Sha, B. & Luo, M. The crystal structure of the influenza matrix protein M1 at neutral pH: M1-M1 protein: Interfaces can rotate in the oligomeric structures of M1. *Virology* **289**, 34-44, doi:10.1006/viro.2001.1119 (2001).
- 102 Martinez-Gil, L. & Mingarro, I. Viroporins, examples of the two-stage membrane protein folding model. *Viruses* **7**, 3462-3482, doi:10.3390/v7072781 (2015).
- 103 Hale, B. G. *et al.* Structural insights into phosphoinositide 3-kinase activation by the influenza A virus NS1 protein. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1954-1959, doi:10.1073/pnas.0910715107 (2010).

- 104 Poole, E., Elton, D., Medcalf, L. & Digard, P. Functional domains of the influenza A virus PB2 protein: Identification of NP- and PB1-binding sites. *Virology* **321**, 120-133 (2004).
- 105 Hemerka, J. N. *et al.* Detection and characterization of influenza A virus PA-PB2 interaction through a bimolecular fluorescence complementation assay. *Journal of Virology* **83**, 3944-3955 (2009).
- 106 Kuzuhara, T. *et al.* Structural basis of the influenza A virus RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. *Journal of Biological Chemistry* **284**, 6855-6860 (2009).
- 107 Yuan, P. *et al.* Crystal structure of an avian influenza polymerase PA N reveals an endonuclease active site. *Nature* **458**, 909-913 (2009).
- 108 Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914-918 (2009).
- 109 Bavagnoli, L. *et al.* The novel influenza A virus protein PA-X and its naturally deleted variant show different enzymatic properties in comparison to the viral endonuclease PA. *Nucleic Acids Research* **43**, 9405-9417, doi:10.1093/nar/gkv926 (2015).
- 110 Muramoto, Y., Noda, T., Kawakami, E., Akkina, R. & Kawaoka, Y. Identification of novel influenza A virus proteins translated from PA mRNA. *Journal of Virology* **87**, 2455-2462 (2013).
- 111 Tong, S. *et al.* New World Bats Harbor Diverse Influenza A Viruses. *PLoS Pathogens* **9** (2013).
- 112 Foucalt, M. L., Moules, V., Rosa-Calatrava, M. & Riteau, B. Role for proteases and HLA-G in the pathogenicity of influenza A viruses. *Journal of Clinical Virology* **51**, 155-159 (2011).
- 113 Varghese, J. N., Laver, W. G. & Colman, P. M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* **303**, 35-40 (1983).
- 114 Y.A. Shtyrya, L. V. M., N.V. Bovin*. Influenza Virus Neuraminidase: Structure and Function. *Acta Naturae* **26** (2009).
- 115 Yang, J., Liu, S., Du, L. & Jiang, S. A new role of neuraminidase (NA) in the influenza virus life cycle: implication for developing NA inhibitors with novel mechanism of action. *Reviews in Medical Virology* **26**, 242-250, doi:10.1002/rmv.1879 (2016).
- 116 Lamb, R. A. & Lai, C. J. Conservation of the influenza virus membrane protein (M1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains. *Virology* **112**, 746-751 (1981).
- 117 Sha, B. & Luo, M. Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein M1. *Nature Structural Biology* **4**, 239-244 (1997).
- 118 Zhang, K. *et al.* Dissection of influenza A virus M1 protein: PH-dependent oligomerization of N-terminal domain and dimerization of C-terminal domain. *PLoS ONE* **7** (2012).
- 119 Pinto, L. H. & Lamb, R. A. The M2 proton channels of influenza A and B viruses. *Journal of Biological Chemistry* **281**, 8997-9000 (2006).
- 120 Carrillo, B. *et al.* The influenza A virus protein NS1 displays structural polymorphism. *Journal of Virology* **88**, 4113-4122 (2014).
- 121 Veit, M., Engel, S., Thaa, B., Scolari, S. & Herrmann, A. Lipid domain association of influenza virus proteins detected by dynamic fluorescence microscopy techniques. *Cellular Microbiology* **15**, 179-189 (2013).

- 122 Jiang, W. *et al.* Influenza A virus NS1 induces G0/G1 cell cycle arrest by inhibiting the expression and activity of RhoA protein. *Journal of Virology* **87**, 3039-3052 (2013).
- 123 de Chassey, B. *et al.* The Interactomes of Influenza Virus NS1 and NS2 Proteins Identify New Host Factors and Provide Insights for ADAR1 Playing a Supportive Role in Virus Replication. *PLoS Pathogens* **9** (2013).
- 124 de Vries, W., Haasnoot, J., Fouchier, R., de Haan, P. & Berkhout, B. Differential RNA silencing suppression activity of NS1 proteins from different influenza A virus strains. *Journal of General Virology* **90**, 1916-1922 (2009).
- 125 Ha, Y., Stevens, D. J., Skehel, J. J. & Wiley, D. C. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11181-11186 (2001).
- 126 Edinger, T. O., Pohl, M. O. & Stertz, S. Entry of influenza A virus: Host factors and antiviral targets. *Journal of General Virology* **95**, 263-277, doi:10.1099/vir.0.059477-0 (2014).
- 127 Lakadamyali, M., Rust, M. J. & Zhuang, X. Endocytosis of influenza viruses. *Microbes and Infection* **6**, 929-936 (2004).
- 128 Chen, C. & Zhuang, X. Epsin 1 is a cargo-specific adaptor for the clathrin-mediated endocytosis of the influenza virus. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 11790-11795 (2008).
- 129 Elbahesh, H. *et al.* Novel roles of focal adhesion kinase in cytoplasmic entry and replication of influenza A viruses. *Journal of Virology* **88**, 6714-6728, doi:10.1128/JVI.00530-14 (2014).
- 130 Nunes-Correia, I., Eulálio, A., Nir, S. & Pedroso De Lima, M. C. Caveolae as an additional route for influenza virus endocytosis in MDCK cells. *Cellular and Molecular Biology Letters* **9**, 47-60 (2004).
- 131 Sieczkarski, S. B. & Whittaker, G. R. Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis. *Journal of Virology* **76**, 10455-10464 (2002).
- 132 de Vries, E. *et al.* Dissection of the influenza A virus endocytic routes reveals macropinocytosis as an alternative entry pathway. *PLoS Pathogens* **7** (2011).
- 133 Sieczkarski, S. B. & Whittaker, G. R. Differential requirements of Rab5 and Rab7 for endocytosis of influenza and other enveloped viruses. *Traffic* **4**, 333-343 (2003).
- 134 Lakadamyali, M., Rust, M. J., Babcock, H. P. & Zhuang, X. Visualizing infection of individual influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 9280-9285 (2003).
- 135 Li, S. *et al.* PH-controlled two-step uncoating of influenza virus. *Biophysical Journal* **106**, 1447-1456 (2014).
- 136 Fontana, J. & Steven, A. C. At Low pH, influenza virus matrix protein M1 undergoes a conformational change prior to dissociating from the membrane. *Journal of Virology* **87**, 5621-5628 (2013).
- 137 Hamilton, B. S., Whittaker, G. R. & Daniel, S. Influenza virus-mediated membrane fusion: Determinants of hemagglutinin fusogenic activity and experimental approaches for assessing virus fusion. *Viruses* **4**, 1144-1168 (2012).
- 138 Banerjee, I. *et al.* Influenza A virus uses the aggresome processing machinery for host cell entry. *Science* **346**, 473-477, doi:10.1126/science.1257037 (2014).

- 139 Kemler, I., Whittaker, G. & Helenius, A. Nuclear import of microinjected influenza virus
ribonucleoproteins. *Virology* **202**, 1028-1033 (1994).
- 140 Wu, W. W. H., Sun, Y. H. B. & Panté, N. Nuclear import of influenza A viral
ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral
nucleoprotein. *Virology Journal* **4** (2007).
- 141 Batra, J. *et al.* Human Heat shock protein 40 (Hsp40/DnaJB1) promotes influenza A virus
replication by assisting nuclear import of viral ribonucleoproteins. *Scientific Reports* **6**,
doi:10.1038/srep19063 (2016).
- 142 Li, C. *et al.* Selection of antigenically advanced variants of seasonal influenza viruses.
Nature microbiology **1**, 16058, doi:10.1038/nmicrobiol.2016.58 (2016).
- 143 Fodor, E. The RNA polymerase of influenza A virus: Mechanisms of viral transcription
and replication. *Acta Virologica* **57**, 113-122 (2013).
- 144 Dubois, J., Terrier, O. & Rosa-Calatrava, M. Influenza viruses and mRNA splicing:
Doing more with less. *mBio* **5** (2014).
- 145 Mohr, I. Closing in on the causes of host shutoff: Influenza a virus suppresses the
translation of host mRNA by selectively remodeling and dominating the pool of mRNA
in infected cells. *eLife* **5**, doi:10.7554/eLife.20755 (2016).
- 146 Jorba, N., Coloma, R. & Ortín, J. Genetic trans-complementation establishes a new
model for influenza virus RNA transcription and replication. *PLoS Pathogens* **5** (2009).
- 147 Green, T. J. *et al.* Common mechanism for RNA encapsidation by negative-strand RNA
viruses. *Journal of Virology* **88**, 3766-3775 (2014).
- 148 York, A., Hengrung, N., Vreede, F. T., Huiskonen, J. T. & Fodor, E. Isolation and
characterization of the positive-sense replicative intermediate of a negative-strand RNA
virus. *Proceedings of the National Academy of Sciences of the United States of America*
110, E4238-E4245 (2013).
- 149 Martín-Benito, J. & Ortín, J. in *Advances in Virus Research* Vol. 87 113-137 (2013).
- 150 Rossman, J. S. & Lamb, R. A. Influenza virus assembly and budding. *Virology* **411**, 229-
236 (2011).
- 151 Veit, M. & Thaa, B. Association of influenza virus proteins with membrane rafts.
Advances in Virology **2011** (2011).
- 152 Schrauwen, E. J. A. & Fouchier, R. A. M. Host adaptation and transmission of influenza
A viruses in mammals. *Emerging Microbes and Infections* **3**, doi:10.1038/emi.2014.9
(2014).
- 153 De Graaf, M. & Fouchier, R. A. M. Role of receptor binding specificity in influenza A
virus transmission and pathogenesis. *EMBO Journal* **33**, 823-841,
doi:10.1002/emboj.201387442 (2014).
- 154 Shi, Y., Wu, Y., Zhang, W., Qi, J. & Gao, G. F. Enabling the 'host jump': Structural
determinants of receptor-binding specificity in influenza A viruses. *Nature Reviews
Microbiology* **12**, 822-831, doi:10.1038/nrmicro3362 (2014).
- 155 Tate, M. D. *et al.* Playing hide and seek: how glycosylation of the influenza virus
hemagglutinin can modulate the immune response to infection. *Viruses* **6**, 1294-1316,
doi:10.3390/v6031294 (2014).
- 156 Job, E. R. *et al.* A single amino acid substitution in the hemagglutinin of H3N2 subtype
influenza a viruses is associated with resistance to the long pentraxin PTX3 and enhanced
virulence in mice. *Journal of Immunology* **192**, 271-281, doi:10.4049/jimmunol.1301814
(2014).

- 157 Tate, M. D., Job, E. R., Brooks, A. G. & Reading, P. C. Glycosylation of the hemagglutinin modulates the sensitivity of H3N2 influenza viruses to innate proteins in airway secretions and virulence in mice. *Virology* **413**, 84-92, doi:10.1016/j.virol.2011.01.036 (2011).
- 158 Guo, H. *et al.* Highly pathogenic influenza A(H5Nx) viruses with altered H5 receptor-binding specificity. *Emerging Infectious Diseases* **23**, 220-230, doi:10.3201/eid2302.161072 (2017).
- 159 Schmier, S. *et al.* In silico prediction and experimental confirmation of HA residues conferring enhanced human receptor specificity of H5N1 influenza A viruses. *Scientific Reports* **5**, doi:10.1038/srep11434 (2015).
- 160 Behera, A. K., Chandra, I. & Cherian, S. S. Molecular dynamics simulation of the effects of single (S221P) and double (S221P and K216E) mutations in the hemagglutinin protein of influenza A H5N1 virus: a study on host receptor specificity. *Journal of Biomolecular Structure and Dynamics* **34**, 2054-2067, doi:10.1080/07391102.2015.1106341 (2016).
- 161 Diederich, S. *et al.* Hemagglutinin-neuraminidase balance influences the virulence phenotype of a recombinant H5N3 influenza A virus possessing a polybasic HA0 cleavage site. *Journal of Virology* **89**, 10724-10734, doi:10.1128/JVI.01238-15 (2015).
- 162 Yen, H. L. *et al.* Changes in H5N1 influenza virus hemagglutinin receptor binding domain affect systemic spread. *Proc Natl Acad Sci U S A* **106**, 286-291, doi:10.1073/pnas.0811052106 (2009).
- 163 Zhang, X. *et al.* Hemagglutinin glycosylation modulates the pathogenicity and antigenicity of the H5N1 avian influenza virus. *Vet Microbiol* **175**, 244-256, doi:10.1016/j.vetmic.2014.12.011 (2015).
- 164 Smirnov, Y. A. *et al.* Characterization of adaptation of an avian influenza A (H5N2) virus to a mammalian host. *Acta Virologica* **44**, 1-8 (2000).
- 165 Su, S. *et al.* Epidemiology, Evolution, and Recent Outbreaks of Avian Influenza Virus in China. *J Virol* **89**, 8671-8676, doi:10.1128/JVI.01034-15 (2015).
- 166 Tan, L. *et al.* A combination of HA and PA mutations enhances virulence in a mouse-adapted H6N6 influenza A virus. *Journal of Virology* **88**, 14116-14125, doi:10.1128/JVI.01736-14 (2014).
- 167 Yu, Z. *et al.* Adaptive amino acid substitutions enhance the virulence of a reassortant H7N1 avian influenza virus isolated from wild waterfowl in mice. *Virology* **476**, 233-239, doi:10.1016/j.virol.2014.11.031 (2015).
- 168 Schrauwen, E. J. A. *et al.* Determinants of virulence of influenza A virus. *European Journal of Clinical Microbiology and Infectious Diseases* **33**, 479-490, doi:10.1007/s10096-013-1984-8 (2014).
- 169 Abdelwhab, E. S. M. *et al.* A unique multibasic proteolytic cleavage site and three mutations in the HA2 domain confer high virulence of H7N1 avian influenza virus in chickens. *Journal of Virology* **90**, 400-411, doi:10.1128/JVI.02082-15 (2016).
- 170 Sun, X., Belser, J. A. & Tumpey, T. M. A novel eight amino acid insertion contributes to the hemagglutinin cleavability and the virulence of a highly pathogenic avian influenza A (H7N3) virus in mice. *Virology* **488**, 120-128, doi:10.1016/j.virol.2015.10.032 (2016).
- 171 Zhao, Y. *et al.* Adaptive amino acid substitutions enhance the virulence of a novel human H7N9 influenza virus in mice. *Veterinary Microbiology* **187**, 8-14, doi:10.1016/j.vetmic.2016.02.027 (2016).

- 172 Hu, J. & Liu, X. Crucial role of PA in virus life cycle and host adaptation of influenza A virus. *Medical Microbiology and Immunology* **204**, 137-149, doi:10.1007/s00430-014-0349-y (2015).
- 173 Boivin, S. & Hart, D. J. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin α isoforms provides insights into host adaptation and polymerase assembly. *Journal of Biological Chemistry* **286**, 10439-10448, doi:10.1074/jbc.M110.182964 (2011).
- 174 Aggarwal, S., Dewhurst, S., Takimoto, T. & Kim, B. Biochemical impact of the host adaptation-associated PB2 E627K mutation on the temperature-dependent RNA synthesis kinetics of influenza A virus polymerase complex. *Journal of Biological Chemistry* **286**, 34504-34513, doi:10.1074/jbc.M111.262048 (2011).
- 175 Czudai-Matwich, V., Otte, A., Matrosovich, M., Gabriel, G. & Klenk, H. D. PB2 mutations D701N and S714R promote adaptation of an influenza H5N1 virus to a mammalian host. *Journal of Virology* **88**, 8735-8742, doi:10.1128/JVI.00422-14 (2014).
- 176 Zhao, D. *et al.* Molecular determinants of virulence and stability of a reporter- expressing H5N1 influenza A virus. *Journal of Virology* **89**, 11337-11346, doi:10.1128/JVI.01886-15 (2015).
- 177 Hu, M. *et al.* PB2 substitutions V598T/I increase the virulence of H7N9 influenza A virus in mammals. *Virology* **501**, 92-101, doi:10.1016/j.virol.2016.11.008 (2017).
- 178 Gao, H. *et al.* PA-X is a virulence factor in avian H9N2 influenza virus. *Journal of General Virology* **96**, 2587-2594, doi:10.1099/jgv.0.000232 (2015).
- 179 Hu, M. *et al.* Amino acid substitutions V63I or A37S/I61T/V63I/V100A in the PA N-terminal domain increase the virulence of H7N7 influenza A virus. *Scientific Reports* **6**, doi:10.1038/srep37800 (2016).
- 180 Liedmann, S. *et al.* New virulence determinants contribute to the enhanced immune response and reduced virulence of an influenza a virus A/PR8/34 variant. *Journal of Infectious Diseases* **209**, 532-541, doi:10.1093/infdis/jit463 (2014).
- 181 Mänz, B., Schwemmle, M. & Brunotte, L. Adaptation of avian influenza a virus polymerase in mammals to overcome the host species barrier. *Journal of Virology* **87**, 7200-7209, doi:10.1128/JVI.00980-13 (2013).
- 182 Neumann, G. H5N1 influenza virulence, pathogenicity and transmissibility: What do we know? *Future Virology* **10**, 971-980, doi:10.2217/fvl.15.62 (2015).
- 183 Kaewborisuth, C. *et al.* G45R on nonstructural protein 1 of influenza A virus contributes to virulence by increasing the expression of proinflammatory cytokines in mice. *Archives of Virology* **162**, 45-55, doi:10.1007/s00705-016-3072-8 (2017).
- 184 DeDiego, M. L., Nogales, A., Lambert-Emo, K., Martinez-Sobrido, L. & Topham, D. J. NS1 protein mutation I64T affects interferon responses and virulence of circulating H3N2 human influenza A viruses. *Journal of Virology* **90**, 9693-9711, doi:10.1128/JVI.01039-16 (2016).
- 185 Nogales, A., Martinez-Sobrido, L., Topham, D. J. & DeDiego, M. L. NS1 protein amino acid changes D189N and V194I affect interferon responses, thermosensitivity, and virulence of circulating H3N2 human influenza A viruses. *Journal of Virology* **91**, doi:10.1128/JVI.01930-16 (2017).
- 186 Ayllon, J. *et al.* A single amino acid substitution in the novel H7N9 influenza a virus NS1 protein increases CPSF30 binding and virulence. *Journal of Virology* **88**, 12146-12151, doi:10.1128/JVI.01567-14 (2014).

- 187 Di Lella, S., Herrmann, A. & Mair, C. M. Modulation of the pH Stability of Influenza Virus Hemagglutinin: A Host Cell Adaptation Strategy. *Biophysical Journal* **110**, 2293-2301, doi:10.1016/j.bpj.2016.04.035 (2016).
- 188 Job, E. R. *et al.* Neutralizing inhibitors in the airways of naïve ferrets do not play a major role in modulating the virulence of H3 subtype influenza A viruses. *Virology* **494**, 143-157, doi:10.1016/j.virol.2016.01.024 (2016).
- 189 Shen, Q. *et al.* The Molecular Mechanisms Underlying the Hijack of Host Proteins by the 1918 Spanish Influenza Virus. *ACS chemical biology*, doi:10.1021/acschembio.7b00168 (2017).
- 190 Tripathi, S., White, M. R. & Hartshorn, K. L. The amazing innate immune response to influenza A virus infection. *Innate Immunity* **21**, 73-98, doi:10.1177/1753425913508992 (2015).
- 191 Kreijtz, J. H. C. M., Fouchier, R. A. M. & Rimmelzwaan, G. F. Immune responses to influenza virus infection. *Virus Research* **162**, 19-30, doi:10.1016/j.virusres.2011.09.022 (2011).
- 192 Zanin, M., Baviskar, P., Webster, R. & Webby, R. The Interaction between Respiratory Pathogens and Mucus. *Cell Host and Microbe* **19**, 159-168, doi:10.1016/j.chom.2016.01.001 (2016).
- 193 Yang, X. *et al.* A beneficiary role for neuraminidase in influenza virus penetration through the respiratory mucus. *PLoS ONE* **9**, doi:10.1371/journal.pone.0110026 (2014).
- 194 Cohen, M. *et al.* Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virology Journal* **10**, doi:10.1186/1743-422X-10-321 (2013).
- 195 Hartshorn, K. L. *et al.* Mechanism of binding of surfactant protein D to influenza A viruses: Importance of binding to haemagglutinin to antiviral activity. *Biochemical Journal* **351**, 449-458, doi:10.1042/0264-6021:3510449 (2000).
- 196 Tecle, T. *et al.* Critical role for cross-linking of trimeric lectin domains of surfactant protein D in antiviral activity against influenza A virus. *Biochemical Journal* **412**, 323-329, doi:10.1042/BJ20071663 (2008).
- 197 Van Eijk, M. *et al.* Interactions of influenza A virus with sialic acids present on porcine surfactant protein D. *American Journal of Respiratory Cell and Molecular Biology* **30**, 871-879, doi:10.1165/rcmb.2003-0355OC (2004).
- 198 LeVine, A. M., Whitsett, J. A., Hartshorn, K. L., Crouch, E. C. & Korfhagen, T. R. Surfactant protein D enhances clearance of influenza A virus from the lung in vivo. *Journal of Immunology* **167**, 5868-5873 (2001).
- 199 Herrera-Ramos, E. *et al.* Surfactant protein A genetic variants associate with severe respiratory insufficiency in pandemic influenza A virus infection. *Critical Care* **18**, doi:10.1186/cc13934 (2014).
- 200 Qi, L. *et al.* The ability of pandemic influenza virus hemagglutinins to induce lower respiratory pathology is associated with decreased surfactant protein D binding. *Virology* **412**, 426-434, doi:10.1016/j.virol.2011.01.029 (2011).
- 201 Hillaire, M. L. B. *et al.* The number and position of N-linked glycosylation sites in the hemagglutinin determine differential recognition of seasonal and 2009 pandemic H1N1 influenza virus by porcine surfactant protein D. *Virus Research* **169**, 301-305, doi:10.1016/j.virusres.2012.08.003 (2012).
- 202 Nikolaidis, N. M. *et al.* Mutations flanking the carbohydrate binding site of surfactant protein D confer antiviral activity for pandemic influenza A viruses. *American Journal of*

- Physiology - Lung Cellular and Molecular Physiology* **306**, L1036-L1044, doi:10.1152/ajplung.00035.2014 (2014).
- 203 Chang, W. C. *et al.* Lack of the pattern recognition molecule mannose-binding lectin increases susceptibility to influenza A virus infection. *BMC Immunology* **11**, doi:10.1186/1471-2172-11-64 (2010).
- 204 Chang, W. C. *et al.* Recombinant chimeric lectins consisting of mannose-binding lectin and L-ficolin are potent inhibitors of influenza A virus compared with mannose-binding lectin. *Biochemical Pharmacology* **81**, 388-395, doi:10.1016/j.bcp.2010.10.012 (2011).
- 205 Nelson, B. *et al.* Recombinant human mannose-binding lectin dampens human alveolar macrophage inflammatory responses to influenza A virus in vitro. *Journal of Leukocyte Biology* **95**, 715-722, doi:10.1189/jlb.0313161 (2014).
- 206 Tokunaga, H., Ushirogawa, H. & Ohuchi, M. The pandemic (H1N1) 2009 influenza virus is resistant to mannose-binding lectin. *Virology Journal* **8**, doi:10.1186/1743-422X-8-50 (2011).
- 207 Guo, J. *et al.* Limited effect of recombinant human mannose-binding lectin on the infection of novel influenza A (H7N9) virus in vitro. *Biochemical and Biophysical Research Communications* **458**, 77-81, doi:10.1016/j.bbrc.2015.01.070 (2015).
- 208 Pan, Q. *et al.* L-ficolin binds to the glycoproteins hemagglutinin and neuraminidase and inhibits influenza a virus infection both in vitro and in vivo. *Journal of Innate Immunity* **4**, 312-324, doi:10.1159/000335670 (2012).
- 209 Endo, Y., Matsushita, M. & Fujita, T. in *International Review of Cell and Molecular Biology* Vol. 316 49-110 (2015).
- 210 Verma, A. *et al.* Human H-ficolin inhibits replication of seasonal and pandemic influenza A viruses. *Journal of Immunology* **189**, 2478-2487, doi:10.4049/jimmunol.1103786 (2012).
- 211 Yang, M. L. *et al.* Galectin-1 binds to influenza virus and ameliorates influenza virus pathogenesis. *Journal of Virology* **85**, 10010-10020, doi:10.1128/JVI.00301-11 (2011).
- 212 Li, S. W. *et al.* Antiviral activity of aloe-emodin against influenza A virus via galectin-3 up-regulation. *European Journal of Pharmacology* **738**, 125-132, doi:10.1016/j.ejphar.2014.05.028 (2014).
- 213 Sharma, S. *et al.* T cell immunoglobulin and mucin protein-3 (Tim-3)/Galectin-9 interaction regulates influenza A virus-specific humoral and CD8 T-cell responses. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 19001-19006, doi:10.1073/pnas.1107087108 (2011).
- 214 Hattori, T. *et al.* Inhibition of influenza A virus infection by Galectin-9. *Japanese Journal of Veterinary Research* **61**, 5-18 (2013).
- 215 Katoh, S. *et al.* Increased levels of plasma galectin-9 in patients with influenza virus infection. *Tohoku Journal of Experimental Medicine* **232**, 263-267, doi:10.1620/tjem.232.263 (2014).
- 216 Nita-Lazar, M. *et al.* Desialylation of airway epithelial cells during influenza virus infection enhances pneumococcal adhesion via galectin binding. *Molecular Immunology* **65**, 1-16, doi:10.1016/j.molimm.2014.12.010 (2015).
- 217 Hsieh, I. N. & Hartshorn, K. L. The role of antimicrobial peptides in influenza virus infection and their potential as antiviral and immunomodulatory therapy. *Pharmaceuticals* **9**, doi:10.3390/ph9030053 (2016).

- 218 Weber-Gerlach, M. & Weber, F. Standing on three legs: Antiviral activities of RIG-I against influenza viruses. *Current Opinion in Immunology* **42**, 71-75, doi:10.1016/j.coi.2016.05.016 (2016).
- 219 van der Sandt, C. E., Kreijtz, J. H. C. M. & Rimmelzwaan, G. F. Evasion of influenza A viruses from innate and adaptive immune responses. *Viruses* **4**, 1438-1476, doi:10.3390/v4091438 (2012).
- 220 Lee, N. L. S., Wong, C. K., Hui, D. S. C. & Chan, P. K. S. Role of toll-like receptors in naturally occurring influenza virus infection. *Hong Kong Medical Journal* **20**, 11-15 (2014).
- 221 Wong, J. P. *et al.* Antiviral role of toll-like receptor-3 agonists against seasonal and avian influenza viruses. *Current Pharmaceutical Design* **15**, 1269-1274, doi:10.2174/138161209787846775 (2009).
- 222 Lee, S. M. Y. *et al.* Toll-like receptor 10 is involved in induction of innate immune responses to influenza virus infection. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 3793-3798, doi:10.1073/pnas.1324266111 (2014).
- 223 Allen, I. C. *et al.* The NLRP3 Inflammasome Mediates In Vivo Innate Immunity to Influenza A Virus through Recognition of Viral RNA. *Immunity* **30**, 556-565, doi:10.1016/j.immuni.2009.02.005 (2009).
- 224 Kuriakose, T. & Kanneganti, T. D. Regulation and functions of NLRP3 inflammasome during influenza virus infection. *Molecular Immunology*, doi:10.1016/j.molimm.2017.01.023 (2017).
- 225 Ramos, I. & Fernandez-Sesma, A. Modulating the innate immune response to influenza A virus: Potential therapeutic use of anti-inflammatory drugs. *Frontiers in Immunology* **6**, doi:10.3389/fimmu.2015.00361 (2015).
- 226 Killip, M. J., Fodor, E. & Randall, R. E. Influenza virus activation of the interferon system. *Virus Research* **209**, 11-22, doi:10.1016/j.virusres.2015.02.003 (2015).
- 227 Bahadoran, A. *et al.* Immune responses to influenza virus and its correlation to age and inherited factors. *Frontiers in Microbiology* **7**, doi:10.3389/fmicb.2016.01841 (2016).
- 228 Cox, R. J., Brokstad, K. A. & Ogra, P. Influenza Virus: Immunity and Vaccination Strategies. Comparison of the Immune Response to Inactivated and Live, Attenuated Influenza Vaccines. *Scandinavian Journal of Immunology* **59**, 1-15, doi:10.1111/j.0300-9475.2004.01382.x (2004).
- 229 Schmolke, M. & García-Sastre, A. Evasion of innate and adaptive immune responses by influenza A virus. *Cellular Microbiology* **12**, 873-880, doi:10.1111/j.1462-5822.2010.01475.x (2010).
- 230 Gustin, K. M., Belser, J. A., Katz, J. M., Tumpey, T. M. & Maines, T. R. Innovations in modeling influenza virus infections in the laboratory. *Trends in Microbiology* **20**, 275-281 (2012).
- 231 Bouvier, N. M. & Lowen, A. C. Animal models for influenza virus pathogenesis and transmission. *Viruses* **2**, 1530-1563 (2010).
- 232 Wang, J. *et al.* Innate immune response to influenza A virus in differentiated human alveolar type II cells. *American Journal of Respiratory Cell and Molecular Biology* **45**, 582-591 (2011).
- 233 Askovich, P. S. *et al.* Differential Host Response, Rather Than Early Viral Replication Efficiency, Correlates with Pathogenicity Caused by Influenza Viruses. *PLoS ONE* **8** (2013).

- 234 Ramírez-Martínez, G. *et al.* Seasonal and pandemic influenza H1N1 viruses induce differential expression of SOCS-1 and RIG-I genes and cytokine/chemokine production in macrophages. *Cytokine* **62**, 151-159 (2013).
- 235 Gerlach, R. L., Camp, J. V., Chu, Y. K. & Jonsson, C. B. Early host responses of seasonal and pandemic influenza A viruses in primary well-differentiated human lung epithelial cells. *PLoS ONE* **8** (2013).
- 236 Iparraguirre, A. *et al.* Two distinct activation states of plasmacytoid dendritic cells induced by influenza virus and CpG 1826 oligonucleotide. *Journal of Leukocyte Biology* **83**, 610-620 (2008).
- 237 Pommerenke, C. *et al.* Global transcriptome analysis in influenza-infected mouse lungs reveals the kinetics of innate and adaptive host immune responses. *PLoS ONE* **7** (2012).
- 238 Brandes, M., Klauschen, F., Kuchen, S. & Germain, R. N. XA systems analysis identifies a feedforward inflammatory circuit leading to lethal influenza infection. *Cell* **154**, X197-212 (2013).
- 239 Hsu, A. C. Y., Barr, I., Hansbro, P. M. & Wark, P. A. Human influenza is more effective than avian influenza at antiviral suppression in airway cells. *American Journal of Respiratory Cell and Molecular Biology* **44**, 906-913 (2011).
- 240 Ivan, F. X. *et al.* Neutrophils infected with highly virulent influenza H3N2 virus exhibit augmented early cell death and rapid induction of type I interferon signaling pathways. *Genomics* **101**, 101-112 (2013).
- 241 Ivan, F. X. *et al.* Differential pulmonary transcriptomic profiles in murine lungs infected with low and highly virulent influenza H3N2 viruses reveal dysregulation of TREM1 signaling, cytokines, and chemokines. *Functional and Integrative Genomics* **12**, 105-117 (2012).
- 242 Huang, Y. *et al.* Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza A infection. *PLoS Genetics* **7** (2011).
- 243 Delgado-Ortega, M. *et al.* Innate immune response to a H3N2 subtype swine influenza virus in newborn porcine trachea cells, alveolar macrophages, and precision-cut lung slices. *Veterinary Research* **45** (2014).
- 244 Paquette, S. G. *et al.* Pandemic H1N1 influenza A directly induces a robust and acute inflammatory gene signature in primary human bronchial epithelial cells downstream of membrane fusion. *Virology* **448**, 91-103 (2014).
- 245 Lee, S. M. Y. *et al.* Systems-level comparison of host responses induced by pandemic and seasonal influenza A H1N1 viruses in primary human type I-like alveolar epithelial cells in vitro. *Respiratory Research* **11** (2010).
- 246 Go, J. T. *et al.* 2009 pandemic H1N1 influenza virus elicits similar clinical course but differential host transcriptional response in mouse, macaque, and swine infection models. *BMC Genomics* **13** (2012).
- 247 Shoemaker, J. E. *et al.* Integrated network analysis reveals a novel role for the cell cycle in 2009 pandemic influenza virus-induced inflammation in macaque lungs. *BMC Systems Biology* **6** (2012).
- 248 Parnell, G. P. *et al.* A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia. *Critical Care* **16** (2012).
- 249 Bermejo-Martin, J. F. *et al.* Host adaptive immunity deficiency in severe pandemic influenza. *Critical Care* **14** (2010).

- 250 Herberg, J. A. *et al.* Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes. *Journal of Infectious Diseases* **208**, 1664-1668 (2013).
- 251 Josset, L. *et al.* Implication of inflammatory macrophages, nuclear receptors, and interferon regulatory factors in increased virulence of pandemic 2009 H1N1 influenza A virus after host adaptation. *Journal of Virology* **86**, 7192-7206 (2012).
- 252 Watanabe, T. *et al.* 1918 Influenza virus hemagglutinin (HA) and the viral RNA polymerase complex enhance viral pathogenicity, but only HA induces aberrant host responses in mice. *Journal of Virology* **87**, 5239-5254 (2013).
- 253 Li, C. *et al.* Host regulatory network response to infection with highly pathogenic H5N1 avian influenza virus. *Journal of Virology* **85**, 10955-10967 (2011).
- 254 Börgeling, Y. *et al.* Inhibition of p38 mitogen-activated protein kinase impairs influenza virus-induced primary and secondary host gene responses and protects mice from lethal H5N1 infection. *Journal of Biological Chemistry* **289**, 13-27 (2014).
- 255 Hu, J. *et al.* The PA-gene-mediated lethal dissemination and excessive innate immune response contribute to the high virulence of H5N1 avian influenza virus in mice. *Journal of Virology* **87**, 2660-2672 (2013).
- 256 Leymarie, O. *et al.* Kinetic Characterization of PB1-F2-Mediated Immunopathology during Highly Pathogenic Avian H5N1 Influenza Virus Infection. *PLoS ONE* **8** (2013).
- 257 Leymarie, O. *et al.* PB1-F2 attenuates virulence of highly pathogenic avian H5N1 influenza virus in chickens. *PLoS ONE* **9** (2014).
- 258 Massin, P. *et al.* Differential cellular gene expression in duck trachea infected with a highly or low pathogenic H5N1 avian influenza virus. *Virology Journal* **10** (2013).
- 259 Cameron, C. M. *et al.* Gene expression analysis of host innate immune responses during lethal H5N1 infection in ferrets. *Journal of Virology* **82**, 11308-11317 (2008).
- 260 Baskin, C. R. *et al.* Early and sustained innate immune response defines pathology and death in nonhuman primates infected by highly pathogenic influenza virus. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 3455-3460 (2009).
- 261 Muramoto, Y. *et al.* Disease severity is associated with differential gene expression at the early and late phases of infection in nonhuman primates infected with different H5N1 highly pathogenic Avian influenza viruses. *Journal of Virology* **88**, 8981-8997 (2014).
- 262 Friesenhagen, J. *et al.* Highly pathogenic avian influenza viruses inhibit effective immune responses of human blood-derived macrophages. *Journal of Leukocyte Biology* **92**, 11-20 (2012).
- 263 Dankar, S. K. *et al.* Influenza A/Hong Kong/156/1997(H5N1) virus NS1 gene mutations F103L and M106I both increase IFN antagonism, virulence and cytoplasmic localization but differ in binding to RIG-I and CPSF30. *Virology Journal* **10** (2013).
- 264 Sakabe, S. *et al.* Differences in cytokine production in human macrophages and in virulence in mice are attributable to the acidic polymerase protein of highly pathogenic influenza A virus subtype H5N1. *Journal of Infectious Diseases* **207**, 262-271 (2013).
- 265 Arilahti, V., Mäkelä, S. M., Tynell, J., Julkunen, I. & Österlund, P. Novel avian influenza a (H7N9) virus induces impaired interferon responses in human dendritic cells. *PLoS ONE* **9** (2014).
- 266 Josset, L., Zeng, H., Kelly, S. M., Tumpey, T. M. & Katze, M. G. Transcriptomic characterization of the novel avian-origin influenza A (H7N9) virus: Specific host

- response and responses intermediate between Avian (H5N1 and H7N7) and human (H3N2) viruses and implications for treatment options. *mBio* **5** (2014).
- 267 Morrison, J. *et al.* H7N9 and Other Pathogenic Avian Influenza Viruses Elicit a Three-
Pronged Transcriptomic Signature That Is Reminiscent of 1918 Influenza Virus and Is
Associated with Lethal Outcome in Mice. *J Virol* **88**, 10556-10568,
doi:10.1128/JVI.00570-14 (2014).
- 268 Josset, L. *et al.* Gene expression signature-based screening identifies new broadly
effective influenza A antivirals. *PLoS ONE* **5** (2010).
- 269 Sutejo, R. *et al.* Activation of type I and III interferon signalling pathways occurs in lung
epithelial cells infected with low pathogenic avian influenza viruses. *PLoS ONE* **7** (2012).
- 270 Calderón, A. J. & Lavergne, J. A. RNA interference: a novel and physiologic mechanism
of gene silencing with great therapeutic potential. *Puerto Rico health sciences journal* **24**,
27-33 (2005).
- 271 Tuzmen, S., Tuzmen, P., Arora, S., Mousses, S. & Azorsa, D. in *Methods in Molecular
Biology* Vol. 700 271-290 (2011).
- 272 Shan, G. RNA interference as a gene knockdown technique. *International Journal of
Biochemistry and Cell Biology* **42**, 1243-1251 (2010).
- 273 Kim, D. H. & Rossi, J. J. RNAi mechanisms and applications. *BioTechniques* **44**, 613-
616 (2008).
- 274 Watanabe, T. & Kawaoka, Y. Influenza virus-host interactomes as a basis for antiviral
drug development. *Current Opinion in Virology* **14**, 71-78,
doi:10.1016/j.coviro.2015.08.008 (2015).
- 275 Watanabe, T., Watanabe, S. & Kawaoka, Y. Cellular networks involved in the influenza
virus life cycle. *Cell Host and Microbe* **7**, 427-439, doi:10.1016/j.chom.2010.05.008
(2010).
- 276 De Chassey, B., Meyniel-Schicklin, L., Aublin-Gex, A., André, P. & Lotteau, V. Genetic
screens for the control of influenza virus replication: From meta-analysis to drug
discovery. *Molecular BioSystems* **8**, 1297-1303, doi:10.1039/c2mb05416g (2012).
- 277 Hao, L. *et al.* Limited Agreement of Independent RNAi Screens for Virus-Required Host
Genes Owes More to False-Negative than False-Positive Factors. *PLoS Computational
Biology* **9**, doi:10.1371/journal.pcbi.1003235 (2013).
- 278 Hao, L. *et al.* Drosophila RNAi screen identifies host genes important for influenza virus
replication. *Nature* **454**, 890-893, doi:10.1038/nature07151 (2008).
- 279 Karlas, A. *et al.* Genome-wide RNAi screen identifies human host factors crucial for
influenza virus replication. *Nature* **463**, 818-822, doi:10.1038/nature08760 (2010).
- 280 König, R. *et al.* Human host factors required for influenza virus replication. *Nature* **463**,
813-817, doi:10.1038/nature08699 (2010).
- 281 Bakre, A. *et al.* Identification of Host Kinase Genes Required for Influenza Virus
Replication and the Regulatory Role of MicroRNAs. *PLoS ONE* **8**,
doi:10.1371/journal.pone.0066796 (2013).
- 282 Su, W. C. *et al.* Pooled RNAi screen identifies ubiquitin ligase Itch as crucial for
influenza A virus release from the endosome during virus entry. *Proceedings of the
National Academy of Sciences of the United States of America* **110**, 17516-17521,
doi:10.1073/pnas.1312374110 (2013).

- 283 Tripathi, S. *et al.* Meta- and Orthogonal Integration of Influenza "oMICs" Data Defines a
Role for UBR4 in Virus Budding. *Cell Host and Microbe* **18**, 723-735,
doi:10.1016/j.chom.2015.11.002 (2015).
- 284 Shapira, S. D. *et al.* A Physical and Regulatory Map of Host-Influenza Interactions
Reveals Pathways in H1N1 Infection. *Cell* **139**, 1255-1267,
doi:10.1016/j.cell.2009.12.018 (2009).
- 285 Tran, A. T. *et al.* Knockdown of specific host factors protects against influenza virus-
induced cell death. *Cell Death and Disease* **4**, doi:10.1038/cddis.2013.296 (2013).
- 286 Sui, B. *et al.* The use of Random Homozygous Gene Perturbation to identify novel host-
oriented targets for influenza. *Virology* **387**, 473-481, doi:10.1016/j.virol.2009.02.046
(2009).
- 287 Brass, A. L. *et al.* The IFITM proteins mediate cellular resistance to influenza A H1N1
virus, West Nile virus, and dengue virus. *Cell* **139**, 1243-1254,
doi:10.1016/j.cell.2009.12.017 (2009).
- 288 Diot, C. *et al.* Influenza A Virus Polymerase Recruits the RNA Helicase DDX19 to
Promote the Nuclear Export of Viral mRNAs. *Scientific Reports* **6**,
doi:10.1038/srep33763 (2016).
- 289 Lee, J., Kim, J., Son, K., D'Alexandry D'Orengiani, A. L. P. H. & Min, J. Y. Acid
phosphatase 2 (ACP2) is required for membrane fusion during influenza virus entry.
Scientific Reports **7**, doi:10.1038/srep43893 (2017).
- 290 Jorba, N. *et al.* Analysis of the interaction of influenza virus polymerase complex with
human cell factors. *Proteomics* **8**, 2077-2088, doi:10.1002/pmic.200700508 (2008).
- 291 Tafforeau, L. *et al.* Generation and comprehensive analysis of an influenza virus
polymerase cellular interaction network. *Journal of Virology* **85**, 13010-13018,
doi:10.1128/JVI.02651-10 (2011).
- 292 Munier, S., Rolland, T., Diot, C., Jacob, Y. & Naffakh, N. Exploration of binary virus-
host interactions using an infectious protein complementation assay. *Molecular and
Cellular Proteomics* **12**, 2845-2855, doi:10.1074/mcp.M113.028688 (2013).
- 293 Mayer, D. *et al.* Identification of cellular interaction partners of the influenza virus
ribonucleoprotein complex and polymerase complex using proteomic-based approaches.
Journal of Proteome Research **6**, 672-682, doi:10.1021/pr060432u (2007).
- 294 York, A., Hutchinson, E. C. & Fodor, E. Interactome analysis of the influenza A virus
transcription/replication machinery identifies protein phosphatase 6 as a cellular factor
required for efficient virus replication. *Journal of Virology* **88**, 13284-13299,
doi:10.1128/JVI.01813-14 (2014).
- 295 Bortz, E. *et al.* Host- and strain-specific regulation of influenza virus polymerase activity
by interacting cellular proteins. *mBio* **2**, doi:10.1128/mBio.00151-11 (2011).
- 296 Fislová, T., Thomas, B., Graef, K. M. & Fodor, E. Association of the influenza virus
RNA polymerase subunit PB2 with the host chaperonin CCT. *Journal of Virology* **84**,
8691-8699, doi:10.1128/JVI.00813-10 (2010).
- 297 Fu, B. *et al.* TRIM32 Senses and Restricts Influenza A Virus by Ubiquitination of PB1
Polymerase. *PLoS Pathogens* **11**, doi:10.1371/journal.ppat.1004960 (2015).
- 298 Wang, L. *et al.* Comparative influenza protein interactomes identify the role of
plakophilin 2 in virus restriction. *Nature Communications* **8**, doi:10.1038/ncomms13876
(2017).

- 299 Guan, Z. H. *et al.* Identification of cellular proteins interacting with influenza A virus
PB1-F2 protein. *Acta Virologica* **56**, 199-207, doi:10.4149/av_2012_03_199 (2012).
- 300 Bradel-Tretheway, B. G. *et al.* Comprehensive proteomic analysis of influenza virus
polymerase complex reveals a novel association with mitochondrial proteins and RNA
polymerase accessory factors. *Journal of Virology* **85**, 8569-8581,
doi:10.1128/JVI.00496-11 (2011).
- 301 Kumar, D., Broor, S. & Rajala, M. S. Interaction of host nucleolin with influenza A virus
nucleoprotein in the early phase of infection limits the late viral gene expression. *PLoS*
ONE **11**, doi:10.1371/journal.pone.0164146 (2016).
- 302 Kuo, R. L. *et al.* Interactome analysis of the NS1 protein encoded by influenza A H1N1
virus reveals a positive regulatory role of host protein PRP19 in viral replication. *Journal*
of Proteome Research **15**, 1639-1648, doi:10.1021/acs.jproteome.6b00103 (2016).
- 303 Mochalova, L. V. *et al.* Synthetic polymeric inhibitors of influenza virus receptor-binding
activity suppress virus replication. *Antiviral Research* **23**, 179-190 (1994).
- 304 Hidari, K. I. P. J. *et al.* Chemoenzymatic synthesis, characterization, and application of
glycopolymers carrying lactosamine repeats as entry inhibitors against influenza virus
infection. *Glycobiology* **18**, 779-788 (2008).
- 305 Song, G. *et al.* Discovery of the first series of small molecule H5N1 entry inhibitors.
Journal of Medicinal Chemistry **52**, 7368-7371 (2009).
- 306 Sato, Y. *et al.* High mannose-binding lectin with preference for the cluster of α 1-2-
mannose from the green alga *Boodlea coacta* is a potent entry inhibitor of HIV-1 and
influenza viruses. *Journal of Biological Chemistry* **286**, 19446-19458 (2011).
- 307 Matsubara, T. *et al.* Sialic acid-mimic peptides as hemagglutinin inhibitors for anti-
influenza therapy. *Journal of Medicinal Chemistry* **53**, 4441-4449 (2010).
- 308 Zeng, L. Y., Yang, J. & Liu, S. Investigational hemagglutinin-targeted influenza virus
inhibitors. *Expert Opinion on Investigational Drugs* **26**, 63-73,
doi:10.1080/13543784.2017.1269170 (2017).
- 309 Koszalka, P., Tilmanis, D. & Hurt, A. C. Influenza antivirals currently in late-phase
clinical trial. *Influenza and other respiratory viruses*, doi:10.1111/irv.12446 (2017).
- 310 Cheng, C. K., Tsai, C. H., Shie, J. J. & Fang, J. M. From neuraminidase inhibitors to
conjugates: A step towards better anti-influenza drugs? *Future Medicinal Chemistry* **6**,
757-774 (2014).
- 311 Burnham, A. J., Baranovich, T. & Govorkova, E. A. Neuraminidase inhibitors for
influenza B virus infection: Efficacy and resistance. *Antiviral Research* **100**, 520-534,
doi:10.1016/j.antiviral.2013.08.023 (2013).
- 312 Muthuri, S. G. *et al.* Effectiveness of neuraminidase inhibitors in reducing mortality in
patients admitted to hospital with influenza A H1N1pdm09 virus infection: a meta-
analysis of individual participant data. *The Lancet. Respiratory medicine* **2**, 395-404,
doi:10.1016/S2213-2600(14)70041-4 (2014).
- 313 Nitsch-Osuch, A. & Brydak, L. B. Influenza viruses resistant to neuraminidase inhibitors.
Acta Biochimica Polonica **61**, 505-508 (2014).
- 314 Feng, E. *et al.* Recent Advances in Neuraminidase Inhibitor Development as Anti-
influenza Drugs. *ChemMedChem* **7**, 1527-1536 (2012).
- 315 Kim, C. U. Rational drug design of orally active influenza neuraminidase inhibitors:
Discovery and development of GS 4104. *Medicinal Chemistry Research* **8**, 392-399
(1998).

- 316 Rakers, C. *et al.* Inhibitory potency of flavonoid derivatives on influenza virus neuraminidase. *Bioorganic and Medicinal Chemistry Letters* **24**, 4312-4317 (2014).
- 317 Clementi, N., Criscuolo, E., Castelli, M. & Clementi, M. Broad-range neutralizing anti-influenza a human monoclonal antibodies: New perspectives in therapy and prophylaxis. *New Microbiologica* **35**, 399-406 (2012).
- 318 Doyle, T. M. *et al.* A monoclonal antibody targeting a highly conserved epitope in influenza B neuraminidase provides protection against drug resistant strains. *Biochemical and Biophysical Research Communications* **441**, 226-229 (2013).
- 319 Nachbagauer, R. & Krammer, F. Universal influenza virus vaccines and therapeutic antibodies. *Clinical Microbiology and Infection* **23**, 222-228, doi:10.1016/j.cmi.2017.02.009 (2017).
- 320 Shahzad, M. I., Naeem, K., Mukhtar, M. & Khanum, A. Passive immunization against highly pathogenic Avian Influenza Virus (AIV) strain H7N3 with antiserum generated from viral polypeptides protect poultry birds from lethal viral infection. *Virology Journal* **5** (2008).
- 321 Chames, P., Van Regenmortel, M., Weiss, E. & Baty, D. Therapeutic antibodies: Successes, limitations and hopes for the future. *British Journal of Pharmacology* **157**, 220-233 (2009).
- 322 Casadevall, A. Antibody-Based Therapies for Emerging Infectious Diseases. *Emerging Infectious Diseases* **2**, 200-208 (1996).
- 323 Richard, J. S. *et al.* Specific structural alteration of the influenza haemagglutinin by amantadine. *EMBO Journal* **9**, 3469-3476 (1990).
- 324 Sugrue, R. J., Tan, B. H., Yeo, D. S. Y. & Sutejo, R. Antiviral drugs for the control of pandemic influenza virus. *Annals of the Academy of Medicine Singapore* **37**, 518-524 (2008).
- 325 Gu, R. X., Liu, L. A. & Wei, D. Q. Structural and energetic analysis of drug inhibition of the influenza A M2 proton channel. *Trends in Pharmacological Sciences* **34**, 571-580 (2013).
- 326 Jefferson, T., Demicheli, V., Di Pietrantonj, C. & Rivetti, D. Amantadine and rimantadine for influenza A in adults. *Cochrane Database of Systematic Reviews* (2006).
- 327 Bright, R. A., Shay, D. K., Shu, B., Cox, N. J. & Klimov, A. I. Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. *Journal of the American Medical Association* **295**, 891-894 (2006).
- 328 Du, Q. S., Huang, R. B., Wang, S. Q. & Chou, K. C. Designing inhibitors of M2 proton channel against H1N1 swine influenza virus. *PLoS ONE* **5** (2010).
- 329 Zhao, X. *et al.* Design and synthesis of pinanamine derivatives as anti-influenza A M2 ion channel inhibitors. *Antiviral Research* **96**, 91-99 (2012).
- 330 Zhao, X. *et al.* Identification of camphor derivatives as novel M2 ion channel inhibitors of influenza A virus. *MedChemComm* **6**, 727-731, doi:10.1039/c4md00515e (2015).
- 331 Reviriego, C. RNA-directed RNA polymerase NS5B inhibitor treatment of influenza A. *Drugs of the Future* **35**, 713-717 (2010).
- 332 Furuta, Y. *et al.* Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral Research* **100**, 446-454 (2013).
- 333 Chan-Tack, K. M., Murray, J. S. & Birnkrant, D. B. Use of ribavirin to treat influenza. *New England Journal of Medicine* **361**, 1713-1714 (2009).

- 334 Byrn, R. A. *et al.* Preclinical activity of VX-787, a first-in-class, orally bioavailable inhibitor of the influenza virus polymerase PB2 subunit. *Antimicrobial Agents and Chemotherapy* **59**, 1569-1582, doi:10.1128/AAC.04623-14 (2015).
- 335 Yuan, S. *et al.* A novel small-molecule inhibitor of influenza A virus acts by suppressing PA endonuclease activity of the viral polymerase. *Scientific Reports* **6**, doi:10.1038/srep22880 (2016).
- 336 Engel, D. A. The influenza virus NS1 protein as a therapeutic target. *Antiviral Research* **99**, 409-416 (2013).
- 337 Woo, H. M. *et al.* Single-stranded DNA aptamer that specifically binds to the influenza virus NS1 protein suppresses interferon antagonism. *Antiviral Research* **100**, 337-345 (2013).
- 338 Kao, R. Y. *et al.* Identification of influenza A nucleoprotein as an antiviral target. *Nature Biotechnology* **28**, 600-605 (2010).
- 339 Amorim, M. J., Kao, R. Y. & Digard, P. Nucleozin targets cytoplasmic trafficking of viral ribonucleoprotein-Rab11 complexes in influenza A virus infection. *Journal of Virology* **87**, 4694-4703 (2013).
- 340 Zhu, J. *et al.* Comprehensive identification of host modulators of HIV-1 replication using multiple orthologous RNAi reagents. *Cell Reports* **9**, 752-766, doi:10.1016/j.celrep.2014.09.031 (2014).
- 341 Szklarczyk, D. *et al.* The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research* **39**, D561-D568, doi:10.1093/nar/gkq973 (2011).
- 342 König, R. & Stertz, S. Recent strategies and progress in identifying host factors involved in virus replication. *Current Opinion in Microbiology* **26**, 79-88, doi:10.1016/j.mib.2015.06.001 (2015).
- 343 Law, G. L., Tisoncik-Go, J., Korth, M. J. & Katze, M. G. Drug repurposing: a better approach for infectious disease drug discovery? *Current opinion in immunology* **25**, 588-592 (2013).
- 344 Bourdakou, M. M., Athanasiadis, E. I. & Spyrou, G. M. Discovering gene re-ranking efficiency and conserved gene-gene relationships derived from gene co-expression network analysis on breast cancer data. *Scientific Reports* **6**, doi:10.1038/srep20518 (2016).
- 345 Sutejo, R. *et al.* Activation of type I and III interferon signalling pathways occurs in lung epithelial cells infected with low pathogenic avian influenza viruses. *PLoS One* **7**, e33732, doi:10.1371/journal.pone.0033732 (2012).
- 346 Taye, B. *et al.* Systems-based approach to examine the cytokine responses in primary mouse lung macrophages infected with low pathogenic avian Influenza virus circulating in South East Asia. *BMC Genomics* **18**, 420, doi:10.1186/s12864-017-3803-6 (2017).
- 347 Yeo, D. S. Y. *et al.* Molecular characterization of low pathogenic avian influenza viruses, isolated from food products imported into Singapore. *Veterinary Microbiology* **138**, 304-317, doi:10.1016/j.vetmic.2009.04.025 (2009).
- 348 Atkins, C. *et al.* Global Human-Kinase Screening Identifies Therapeutic Host Targets against Influenza. *Journal of biomolecular screening* **19**, 936-946, doi:10.1177/1087057113518068 (2014).

- 349 Bakre, A. *et al.* Identification of Host Kinase Genes Required for Influenza Virus Replication and the Regulatory Role of MicroRNAs. *PLoS One* **8**, e66796, doi:10.1371/journal.pone.0066796 (2013).
- 350 Konig, R. *et al.* Human host factors required for influenza virus replication. *Nature* **463**, 813-817, doi:10.1038/nature08699 (2010).
- 351 Shapira, S. D. *et al.* A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139**, 1255-1267, doi:10.1016/j.cell.2009.12.018 (2009).
- 352 Sugden, L. A., Tackett, M. R., Savva, Y. A., Thompson, W. A. & Lawrence, C. E. Assessing the validity and reproducibility of genome-scale predictions. *Bioinformatics* **29**, 2844-2851, doi:10.1093/bioinformatics/btt508 (2013).
- 353 Tran, A. T. *et al.* Knockdown of specific host factors protects against influenza virus-induced cell death. *Cell death & disease* **4**, e769, doi:10.1038/cddis.2013.296 (2013).
- 354 Watanabe, T. *et al.* Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell host & microbe* **16**, 795-805, doi:10.1016/j.chom.2014.11.002 (2014).
- 355 Bradel-Tretheway, B. G. *et al.* Comprehensive proteomic analysis of influenza virus polymerase complex reveals a novel association with mitochondrial proteins and RNA polymerase accessory factors. *J Virol* **85**, 8569-8581, doi:10.1128/JVI.00496-11 (2011).
- 356 de Chassey, B. *et al.* Structure homology and interaction redundancy for discovering virus-host protein interactions. *EMBO reports* **14**, 938-944, doi:10.1038/embo.2013.130 (2013).
- 357 Munier, S., Rolland, T., Diot, C., Jacob, Y. & Naffakh, N. Exploration of binary virus-host interactions using an infectious protein complementation assay. *Molecular & cellular proteomics : MCP* **12**, 2845-2855, doi:10.1074/mcp.M113.028688 (2013).
- 358 Tafforeau, L. *et al.* Generation and comprehensive analysis of an influenza virus polymerase cellular interaction network. *J Virol* **85**, 13010-13018, doi:10.1128/JVI.02651-10 (2011).
- 359 Mayer, D. *et al.* Identification of cellular interaction partners of the influenza virus ribonucleoprotein complex and polymerase complex using proteomic-based approaches. *Journal of proteome research* **6**, 672-682, doi:10.1021/pr060432u (2007).
- 360 Generous, A. *et al.* Identification of putative interactions between swine and human influenza A virus nucleoprotein and human host proteins. *Virol J* **11**, 2509, doi:10.1186/s12985-014-0228-6 (2014).
- 361 Gorai, T. *et al.* F1Fo-ATPase, F-type proton-translocating ATPase, at the plasma membrane is critical for efficient influenza virus budding. *Proc Natl Acad Sci U S A* **109**, 4615-4620, doi:10.1073/pnas.1114728109 (2012).
- 362 Emig, D. *et al.* AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res* **38**, W755-762, doi:10.1093/nar/gkq405 (2010).
- 363 de Hoon, M. J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453-1454, doi:10.1093/bioinformatics/bth078 (2004).
- 364 Oliveros, J. C. Venny. An interactive tool for comparing lists with Venn's diagrams. Access date Jan 2017. (2007-2015).
- 365 Atkins, C. *et al.* Global human-kinase screening identifies therapeutic host targets against influenza. *Journal of Biomolecular Screening* **19**, 936-946, doi:10.1177/1087057113518068 (2014).

- 366 Watanabe, T. *et al.* Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell Host and Microbe* **16**, 795-805, doi:10.1016/j.chom.2014.11.002 (2014).
- 367 York, A., Hutchinson, E. C. & Fodor, E. Interactome analysis of the influenza A virus transcription/replication machinery identifies protein phosphatase 6 as a cellular factor required for efficient virus replication. *Journal of virology* **88**, 13284-13299, doi:10.1128/JVI.01813-14 (2014).
- 368 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
- 369 Law, V. *et al.* DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Research* **42**, D1091-D1097, doi:10.1093/nar/gkt1068 (2014).
- 370 Cunningham, F. *et al.* Ensembl 2015. *Nucleic Acids Research* **43**, D662-D669, doi:10.1093/nar/gku1010 (2015).
- 371 Warde-Farley, D. *et al.* The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research* **38**, W214-W220 (2010).
- 372 Gray, K. A., Yates, B., Seal, R. L., Wright, M. W. & Bruford, E. A. Genenames.org: The HGNC resources in 2015. *Nucleic Acids Research* **43**, D1079-D1085, doi:10.1093/nar/gku1071 (2015).
- 373 Kramer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-530, doi:10.1093/bioinformatics/btt703 (2014).
- 374 Hayashi, T., MacDonald, L. A. & Takimoto, T. Influenza A virus protein PA-X contributes to viral growth and suppression of the host antiviral and immune responses. *Journal of Virology* **89**, 6442-6452, doi:10.1128/JVI.00319-15 (2015).
- 375 Altenhoff, A. M. *et al.* The OMA orthology database in 2015: Function predictions, better plant support, synteny view and other improvements. *Nucleic Acids Research* **43**, D240-D249, doi:10.1093/nar/gku1158 (2015).
- 376 Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* **43**, D447-452, doi:10.1093/nar/gku1003 (2015).
- 377 Zhou, H., Gao, M. & Skolnick, J. Comprehensive prediction of drug-protein interactions and side effects for the human proteome. *Scientific Reports* **5**, doi:10.1038/srep11090 (2015).
- 378 Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).
- 379 de Hoon, M. J. L., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453-1454, doi:10.1093/bioinformatics/bth078 (2004).
- 380 Camacho, C. *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, doi:10.1186/1471-2105-10-421 (2009).
- 381 Kans, J. Entrez Direct: E-utilities on the UNIX Command Line. Access date Feb 2016. (2013).
- 382 Fox, J., and Bouchet-Valat, M. Rcmdr: R Commander. R package version 2.3-2. (2017).
- 383 Oliveros, J. C. Venny. An interactive tool for comparing lists with Venn's diagrams. ((2007-2015)).

- 384 Itoh, Y. *et al.* Protective Efficacy of Passive Immunization with Monoclonal Antibodies in Animal Models of H5N1 Highly Pathogenic Avian Influenza Virus Infection. *PLoS Pathogens* **10** (2014).
- 385 von Mering, C. *et al.* STRING: Known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Research* **33**, D433-D437, doi:10.1093/nar/gki005 (2005).
- 386 Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C. & Morris, Q. GeneMANIA: A real-time multiple association network integration algorithm for predicting gene function. *Genome Biology* **9**, doi:10.1186/gb-2008-9-s1-s4 (2008).
- 387 Matsuoka, Y. *et al.* A comprehensive map of the influenza A virus replication cycle. *BMC Systems Biology* **7**, doi:10.1186/1752-0509-7-97 (2013).
- 388 Garten, Y., Tatonetti, N. P. & Altman, R. B. Improving the prediction of pharmacogenes using text-derived drug-gene relationships. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 305-314 (2010).
- 389 Liu, Y., Liang, Y. & Wishart, D. PolySearch2: a significantly improved text-mining system for discovering associations between human diseases, genes, drugs, metabolites, toxins and more. *Nucleic acids research* **43**, W535-542, doi:10.1093/nar/gkv383 (2015).
- 390 Tripathi, S. *et al.* Meta- and Orthogonal Integration of Influenza "OMICS" Data Defines a Role for UBR4 in Virus Budding. *Cell host & microbe* **18**, 723-735, doi:10.1016/j.chom.2015.11.002 (2015).
- 391 Geiss, G. K. *et al.* Global impact of influenza virus on cellular pathways is mediated by both replication-dependent and-independent events. *Journal of Virology* **75**, 4321-4331 (2001).
- 392 Chakrabarti, A. K. *et al.* Host gene expression profiling in influenza A virus-infected lung epithelial (A549) cells: a comparative analysis between highly pathogenic and modified H5N1 viruses. *Virol J* **7**, 219, doi:10.1186/1743-422X-7-219 (2010).
- 393 Maughan, M. N. *et al.* Transcriptional analysis of the innate immune response of ducks to different species-of-origin low pathogenic H7 avian influenza viruses. *Virol J* **10**, 94, doi:10.1186/1743-422X-10-94 (2013).
- 394 Kroeker, A. L., Ezzati, P., Halayko, A. J. & Coombs, K. M. Response of primary human airway epithelial cells to influenza infection: A quantitative proteomic study. *Journal of proteome research* **11**, 4132-4146, doi:10.1021/pr300239r (2012).
- 395 Blanc, M. *et al.* Host defense against viral infection involves interferon mediated down-regulation of sterol biosynthesis. *PLoS Biology* **9**, doi:10.1371/journal.pbio.1000598 (2011).
- 396 Terrier, O. *et al.* Cellular transcriptional profiling in human lung epithelial cells infected by different subtypes of influenza A viruses reveals an overall down-regulation of the host p53 pathway. *Virol J* **8**, 285, doi:10.1186/1743-422X-8-285 (2011).
- 397 Kabátková, M. *et al.* Interactive effects of inflammatory cytokine and abundant low-molecular-weight PAHs on inhibition of gap junctional intercellular communication, disruption of cell proliferation control, and the AhR-dependent transcription. *Toxicology Letters* **232**, 113-121, doi:10.1016/j.toxlet.2014.09.023 (2015).
- 398 Strengert, M. *et al.* Mucosal reactive oxygen species are required for antiviral response: role of Duox in influenza a virus infection. *Antioxidants & redox signaling* **20**, 2695-2709, doi:10.1089/ars.2013.5353 (2014).

- 399 Vreede, F. T. & Fodor, E. The role of the influenza virus RNA polymerase in host shut-off. *Virulence* **1**, 436-439, doi:10.4161/viru.1.5.12967 (2010).
- 400 Reich, S. *et al.* Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **516**, 361-366, doi:10.1038/nature14009 (2014).
- 401 Košík, I. *et al.* The ubiquitination of the influenza A virus PB1-F2 protein is crucial for its biological function. *PLoS ONE* **10**, doi:10.1371/journal.pone.0118477 (2015).
- 402 Gao, S. *et al.* Interaction of NS2 with AIMP2 facilitates the switch from ubiquitination to SUMOylation of M1 in influenza A virus-infected cells. *Journal of Virology* **89**, 300-311, doi:10.1128/JVI.02170-14 (2015).
- 403 Rajsbaum, R. *et al.* Species-Specific Inhibition of RIG-I Ubiquitination and IFN Induction by the Influenza A Virus NS1 Protein. *PLoS Pathogens* **8**, doi:10.1371/journal.ppat.1003059 (2012).
- 404 Zhou, Z. *et al.* TRIM14 is a mitochondrial adaptor that facilitates retinoic acid-inducible gene-I-like receptor-mediated innate immune response. *Proc Natl Acad Sci U S A* **111**, E245-254, doi:10.1073/pnas.1316941111 (2014).
- 405 Young, J. A. *et al.* Fas-associated death domain (FADD) and the E3 ubiquitin-protein ligase TRIM21 interact to negatively regulate virus-induced interferon production. *The Journal of biological chemistry* **286**, 6521-6531, doi:10.1074/jbc.M110.172288 (2011).
- 406 Manocha, G. D. *et al.* Regulatory role of TRIM21 in the type-I interferon pathway in Japanese encephalitis virus-infected human microglial cells. *Journal of neuroinflammation* **11**, 24, doi:10.1186/1742-2094-11-24 (2014).
- 407 Liao, T. L., Wu, C. Y., Su, W. C., Jeng, K. S. & Lai, M. M. C. Ubiquitination and deubiquitination of NP protein regulates influenza A virus RNA replication. *EMBO Journal* **29**, 3879-3890, doi:10.1038/emboj.2010.250 (2010).
- 408 Wang, X. *et al.* Stabilization of p53 in influenza A virus-infected cells is associated with compromised MDM2-mediated ubiquitination of p53. *Journal of Biological Chemistry* **287**, 18366-18375, doi:10.1074/jbc.M111.335422 (2012).
- 409 Yount, J. S., Karssemeijer, R. A. & Hang, H. C. S-palmitoylation and ubiquitination differentially regulate interferon-induced transmembrane protein 3 (IFITM3)-mediated resistance to influenza virus. *Journal of Biological Chemistry* **287**, 19631-19641, doi:10.1074/jbc.M112.362095 (2012).
- 410 Chesarino, N. M., McMichael, T. M. & Yount, J. S. E3 Ubiquitin Ligase NEDD4 Promotes Influenza Virus Infection by Decreasing Levels of the Antiviral Protein IFITM3. *PLoS Pathogens* **11**, doi:10.1371/journal.ppat.1005095 (2015).
- 411 Chesarino N, M. T., Hach J, and Yount J. Identification of an E3 ubiquitin ligase that regulates cellular resistance to influenza virus infection by controlling levels of the antiviral protein IFITM3 *The Journal of Immunology* **194**, 127.124 (2015).
- 412 Vreede, F. T., Chan, A. Y., Sharps, J. & Fodor, E. Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology* **396**, 125-134, doi:10.1016/j.virol.2009.10.003 (2010).
- 413 Wang, P., Zhao, W., Zhao, K., Zhang, L. & Gao, C. TRIM26 Negatively Regulates Interferon- β Production and Antiviral Response through Polyubiquitination and Degradation of Nuclear IRF3. *PLoS Pathogens* **11**, 1-22, doi:10.1371/journal.ppat.1004726 (2015).

414 Sze, A., OLAGNIER, D., Lin, R., van Grevenynghe, J. & Hiscott, J. SAMHD1 host
restriction factor: a link with innate immune sensing of retrovirus infection. *Journal of*
415 *molecular biology* **425**, 4981-4994, doi:10.1016/j.jmb.2013.10.022 (2013).

416 Hrecka, K. *et al.* Vpx relieves inhibition of HIV-1 infection of macrophages mediated by
the SAMHD1 protein. *Nature* **474**, 658-661, doi:10.1038/nature10195 (2011).

417 Laguette, N. *et al.* SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction
factor counteracted by Vpx. *Nature* **474**, 654-657, doi:10.1038/nature10117 (2011).

418 Ryoo, J. *et al.* The ribonuclease activity of SAMHD1 is required for HIV-1 restriction.
Nat Med **20**, 936-941, doi:10.1038/nm.3626 (2014).

419 Chen, Z. *et al.* Inhibition of Hepatitis B virus replication by SAMHD1. *Biochemical and*
Biophysical Research Communications **450**, 1462-1468, doi:10.1016/j.bbrc.2014.07.023
(2014).

420 Badia, R. *et al.* Inhibition of herpes simplex virus type 1 by the CDK6 inhibitor PD-
0332991 (palbociclib) through the control of SAMHD1. *Journal of Antimicrobial*
Chemotherapy **71**, 387-394, doi:10.1093/jac/dkv363 (2016).

421 Antonucci, J. M. *et al.* SAMHD1-mediated HIV-1 restriction in cells does not involve
ribonuclease activity. *Nature Medicine* **22**, 1072-1074, doi:10.1038/nm.4163 (2016).

422 Crow, Y. J. & Manel, N. Aicardi-Goutières syndrome and the type I interferonopathies.
Nature Reviews Immunology **15**, 429-440, doi:10.1038/nri3850 (2015).

423 Crow, Y. J. *et al.* Characterization of human disease phenotypes associated with
mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and
IFIH1. *American journal of medical genetics. Part A* **167**, 296-312,
doi:10.1002/ajmg.a.36887 (2015).

424 Riess, M. *et al.* Interferons induce expression of SAMHD1 in monocytes through down-
regulation of miR-181a and miR-30a. *Journal of Biological Chemistry* **292**, 264-277,
doi:10.1074/jbc.M116.752584 (2017).

425 García, M. A. *et al.* Impact of protein kinase PKR in cell biology: From antiviral to
antiproliferative action. *Microbiology and Molecular Biology Reviews* **70**, 1032-1060,
doi:10.1128/MMBR.00027-06 (2006).

426 Dauber, B. & Wolff, T. Activation of the antiviral kinase PKR and viral countermeasures.
Viruses **1**, 523-544, doi:10.3390/v1030523 (2009).

427 Goodman, A. G., Tanner, B. C. W., Chang, S. T., Esteban, M. & Katze, M. G. Virus
infection rapidly activates the P58IPK pathway, delaying peak kinase activation to
enhance viral replication. *Virology* **417**, 27-36, doi:10.1016/j.virol.2011.04.020 (2011).

428 Sharma, K. *et al.* Influenza A virus nucleoprotein exploits Hsp40 to inhibit PKR
activation. *PLoS ONE* **6**, doi:10.1371/journal.pone.0020215 (2011).

429 Goodman, A. G. *et al.* p58IPK: A novel "CIHD" member of the host innate defense
response against pathogenic virus infection. *PLoS Pathogens* **5**,
doi:10.1371/journal.ppat.1000438 (2009).

430 Bloch, D. B. *et al.* Sp110 localizes to the PML-Sp100 nuclear body and may function as a
nuclear hormone receptor transcriptional coactivator. *Molecular and Cellular Biology* **20**,
6138-6146, doi:10.1128/MCB.20.16.6138-6146.2000 (2000).

Regad, T. & Chelbi-Alix, M. K. Role and fate of PML nuclear bodies in response to
interferon and viral infections. *Oncogene* **20**, 7274-7286, doi:10.1038/sj.onc.1204854
(2001).

- 431 Berscheminski, J. *et al.* Sp100 isoform-specific regulation of human adenovirus 5 gene expression. *Journal of Virology* **88**, 6076-6092, doi:10.1128/JVI.00469-14 (2014).
- 432 Chelbi-Alix, M. K., Quignon, F., Pelicano, L., Koken, M. H. M. & De Thé, H. Resistance to virus infection conferred by the interferon-induced promyelocytic leukemia protein. *Journal of Virology* **72**, 1043-1051 (1998).
- 433 Goujon, C. *et al.* Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature* **502**, 559-562, doi:10.1038/nature12542 (2013).
- 434 Jin, H. K., Takada, A., Kon, Y., Haller, O. & Watanabe, T. Identification of the murine Mx2 gene: Interferon-induced expression of the Mx2 protein from the feral mouse gene confers resistance to vesicular stomatitis virus. *Journal of Virology* **73**, 4925-4930 (1999).
- 435 Jin, H. K. *et al.* Mouse Mx2 protein inhibits hantavirus but not influenza virus replication. *Archives of Virology* **146**, 41-49, doi:10.1007/s007050170189 (2001).
- 436 Zürcher, T., Pavlovic, J. & Staeheli, P. Mouse Mx2 protein inhibits vesicular stomatitis virus but not influenza virus. *Virology* **187**, 796-800, doi:10.1016/0042-6822(92)90481-4 (1992).
- 437 Johannes, L., Arnheiter, H. & Meier, E. Switch in antiviral specificity of a GTPase upon translocation from the cytoplasm to the nucleus. *Journal of Virology* **67**, 1653-1657 (1993).
- 438 Gu, M. *et al.* Ctenopharyngodon idella IRF2 plays an antagonistic role to IRF1 in transcriptional regulation of IFN and ISG genes. *Developmental and Comparative Immunology* **49**, 103-112, doi:10.1016/j.dci.2014.11.014 (2015).
- 439 McDermott, J. E. *et al.* Conserved host response to highly pathogenic avian influenza virus infection in human cell culture, mouse and macaque model systems. *BMC Systems Biology* **5**, doi:10.1186/1752-0509-5-190 (2011).
- 440 Ihle, J. N. & Gilliland, D. G. Jak2: normal function and role in hematopoietic disorders. *Current Opinion in Genetics and Development* **17**, 8-14, doi:10.1016/j.gde.2006.12.009 (2007).
- 441 Wang, S. *et al.* Tyrosine 132 phosphorylation of influenza A virus M1 protein is crucial for virus replication by controlling the nuclear import of M1. *Journal of Virology* **87**, 6182-6191, doi:10.1128/JVI.03024-12 (2013).
- 442 Zheng, W. *et al.* Phosphorylation controls the nuclear-cytoplasmic shuttling of influenza A virus nucleoprotein. *Journal of Virology* **89**, 5822-5834, doi:10.1128/JVI.00015-15 (2015).
- 443 Peña, G., Cai, B., Deitch, E. A. & Ulloa, L. JAK2 inhibition prevents innate immune responses and rescues animals from sepsis. *Journal of Molecular Medicine* **88**, 851-859, doi:10.1007/s00109-010-0628-z (2010).
- 444 Sokol, A. M., Sztolsztener, M. E., Wasilewski, M., Heinz, E. & Chacinska, A. Mitochondrial protein translocases for survival and wellbeing. *FEBS Letters* **588**, 2484-2495, doi:10.1016/j.febslet.2014.05.028 (2014).
- 445 Iacovino, M. *et al.* The conserved translocase Tim17 prevents mitochondrial DNA loss. *Human Molecular Genetics* **18**, 65-74, doi:10.1093/hmg/ddn313 (2009).
- 446 Salhab, M., Patani, N., Jiang, W. & Mokbel, K. High TIMM17A expression is associated with adverse pathological and clinical outcomes in human breast cancer. *Breast Cancer* **19**, 153-160, doi:10.1007/s12282-010-0228-3 (2012).

- 447 Ilyinskii, P. O., Gabai, V. L., Sunyaev, S. R., Thoidis, G. & Shneider, A. M. Toxicity of
influenza A virus matrix protein 2 for mammalian cells is associated with its intrinsic
proton-channeling activity. *Cell Cycle* **6**, 2043-2047 (2007).
- 448 Chen, R., Yik, J. H. N., Lew, Q. J. & Chao, S. H. Brd4 and HEXIM1: Multiple roles in P-
TEFb regulation and cancer. *BioMed Research International* **2014**,
doi:10.1155/2014/232870 (2014).
- 449 Contreras, X., Barboric, M., Lenasi, T. & Peterlin, B. M. HMBA releases P-TEFb from
HEXIM1 and 7SK snRNA via PI3K/Akt and activates HIV transcription. *PLoS
Pathogens* **3**, 1459-1469, doi:10.1371/journal.ppat.0030146 (2007).
- 450 Dahlhoff, M., Wolf, E. & Schneider, M. R. The ABC of BTC: Structural properties and
biological roles of betacellulin. *Seminars in Cell and Developmental Biology* **28**, 42-48,
doi:10.1016/j.semcdb.2014.01.002 (2014).
- 451 Riese, D. J., 2nd *et al.* Betacellulin activates the epidermal growth factor receptor and
erbB-4, and induces cellular response patterns distinct from those stimulated by
epidermal growth factor or neuregulin-beta. *Oncogene* **12**, 345-353 (1996).
- 452 Eierhoff, T., Hrncius, E. R., Rescher, U., Ludwig, S. & Ehrhardt, C. The epidermal
growth factor receptor (EGFR) promotes uptake of influenza A viruses (IAV) into host
cells. *PLoS Pathog* **6**, e1001099, doi:10.1371/journal.ppat.1001099 (2010).
- 453 Aoki, J., Nagai, Y., Hosono, H., Inoue, K. & Arai, H. Structure and function of
phosphatidylserine-specific phospholipase A1. *Biochimica et Biophysica Acta -
Molecular and Cell Biology of Lipids* **1582**, 26-32, doi:10.1016/S1388-1981(02)00134-8
(2002).
- 454 Nakano, T. *et al.* A possible role of lysophospholipids produced by calcium-independent
phospholipase A2 in membrane-raft budding and fission. *Biochimica et Biophysica Acta -
Biomembranes* **1788**, 2222-2228, doi:10.1016/j.bbamem.2009.07.015 (2009).
- 455 Veit, M., Serebryakova, M. V. & Kordyukova, L. V. Palmitoylation of influenza virus
proteins. *Biochemical Society Transactions* **41**, 50-55, doi:10.1042/BST20120210 (2013).
- 456 Lambrecht, B. & Schmidt, M. F. G. Membrane fusion induced by influenza virus
hemagglutinin requires protein bound fatty acids. *FEBS Letters* **202**, 127-132,
doi:10.1016/0014-5793(86)80662-7 (1986).
- 457 Naeve, C. W. & Williams, D. Fatty acids on the A/Japan/305/57 influenza virus
hemagglutinin have a role in membrane fusion. *EMBO Journal* **9**, 3857-3866 (1990).
- 458 Sakai, T., Ohuchi, R. & Ohuchi, M. Fatty acids on the A/USSR/77 influenza virus
hemagglutinin facilitate the transition from hemifusion to fusion pore formation. *Journal
of Virology* **76**, 4603-4611, doi:10.1128/JVI.76.9.4603-4611.2002 (2002).
- 459 Zhou, J. *et al.* Recombinant influenza A H3N2 viruses with mutations of HA
transmembrane cysteines exhibited altered virological characteristics. *Virus Genes* **48**,
273-282, doi:10.1007/s11262-013-1011-2 (2014).
- 460 Munday, A. D. & López, J. A. Posttranslational protein palmitoylation: Promoting
platelet purpose. *Arteriosclerosis, Thrombosis, and Vascular Biology* **27**, 1496-1499,
doi:10.1161/ATVBAHA.106.136226 (2007).
- 461 Hayakawa, S. *et al.* ZAPS is a potent stimulator of signaling mediated by the RNA
helicase RIG-I during antiviral responses. *Nature immunology* **12**, 37-44,
doi:10.1038/ni.1963 (2011).
- 462 Katze, M. G., He, Y. & Gale, M., Jr. Viruses and interferon: a fight for supremacy.
Nature reviews. Immunology **2**, 675-687, doi:10.1038/nri888 (2002).

- 463 Iwasaki, A. & Pillai, P. S. Innate immunity to influenza virus infection. *Nature Reviews Immunology* **14**, 315-328, doi:10.1038/nri3665 (2014).
- 464 Verhelst, J., Parthoens, E., Schepens, B., Fiers, W. & Saelens, X. Interferon-inducible protein Mx1 inhibits influenza virus by interfering with functional viral ribonucleoprotein complex assembly. *J Virol* **86**, 13445-13455, doi:10.1128/JVI.01682-12 (2012).
- 465 Uchida, Y. *et al.* Identification of host genes linked with the survivability of chickens infected with recombinant viruses possessing H5N1 surface antigens from a highly pathogenic avian influenza virus. *J Virol* **86**, 2686-2695, doi:10.1128/JVI.06374-11 (2012).
- 466 Morito, D. *et al.* Moyamoya disease-associated protein mysterin/RNF213 is a novel AAA+ ATPase, which dynamically changes its oligomeric state. *Scientific Reports* **4**, doi:10.1038/srep04442 (2014).
- 467 Everitt, A. R. *et al.* IFITM3 restricts the morbidity and mortality associated with influenza. *Nature* **484**, 519-523, doi:10.1038/nature10921 (2012).
- 468 Silverman, R. H. & Weiss, S. R. Viral phosphodiesterases that antagonize double-stranded RNA signaling to RNase L by degrading 2-5A. *Journal of Interferon and Cytokine Research* **34**, 455-463, doi:10.1089/jir.2014.0007 (2014).
- 469 Desai, T. M. *et al.* IFITM3 Restricts Influenza A Virus Entry by Blocking the Formation of Fusion Pores following Virus-Endosome Hemifusion. *PLoS Pathogens* **10**, doi:10.1371/journal.ppat.1004048 (2014).
- 470 Perwitasari, O., Yan, X., O'Donnell, J., Johnson, S. & Tripp, R. A. Repurposing Kinase Inhibitors as Antiviral Agents to Control Influenza A Virus Replication. *Assay and drug development technologies*, doi:10.1089/adt.2015.0003.drrr (2015).
- 471 Peuschel, K. E. Some clinical evidence of the hypothesis of an indirect antiviral effect of propranolol through immunoactivation. *Medical Hypotheses* **76**, 689-691, doi:10.1016/j.mehy.2011.01.034 (2011).
- 472 Geiler, J. *et al.* N-acetyl-l-cysteine (NAC) inhibits virus replication and expression of pro-inflammatory molecules in A549 cells infected with highly pathogenic H5N1 influenza A virus. *Biochemical Pharmacology* **79**, 413-420, doi:10.1016/j.bcp.2009.08.025 (2010).
- 473 Reperant, L. A., Kuiken, T. & Osterhaus, A. D. M. E. Adaptive pathways of zoonotic influenza viruses: From exposure to establishment in humans. *Vaccine* **30**, 4419-4434, doi:10.1016/j.vaccine.2012.04.049 (2012).
- 474 Kuiken, T. *et al.* Host species barriers to influenza virus infections. *Science* **312**, 394-397, doi:10.1126/science.1122818 (2006).
- 475 Ince, W. L., Gueye-Mbaye, A., Bennink, J. R. & Yewdell, J. W. Reassortment complements spontaneous mutation in influenza A virus NP and M1 genes to accelerate adaptation to a new host. *Journal of Virology* **87**, 4330-4338, doi:10.1128/JVI.02749-12 (2013).
- 476 Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18590-18595, doi:10.1073/pnas.0507415102 (2005).
- 477 Zeng, H. *et al.* Tropism and infectivity of influenza virus, including highly pathogenic avian H5N1 virus, in ferret tracheal differentiated primary epithelial cell cultures. *J Virol* **87**, 2597-2607, doi:10.1128/JVI.02885-12 (2013).

- 478 Thangavel, R. R. & Bouvier, N. M. Animal models for influenza virus pathogenesis,
transmission, and immunology. *Journal of immunological methods* **410**, 60-79,
doi:10.1016/j.jim.2014.03.023 (2014).
- 479 Ballester, B., Johnson, N., Proctor, G. & Flicek, P. Consistent annotation of gene
expression arrays. *BMC Genomics* **11**, doi:10.1186/1471-2164-11-294 (2010).
- 480 Buza, T. J., Kumar, R., Gresham, C. R., Burgess, S. C. & McCarthy, F. M. Facilitating
functional annotation of chicken microarray data. *BMC Bioinformatics* **10** (2009).
- 481 Lu, Y., Huggins, P. & Bar-Joseph, Z. Cross species analysis of microarray expression
data. *Bioinformatics* **25**, 1476-1483, doi:10.1093/bioinformatics/btp247 (2009).
- 482 Dowell, R. D. The similarity of gene expression between human and mouse tissues.
Genome Biology **12**, doi:10.1186/gb-2011-12-1-101 (2011).
- 483 Briggs, J. *et al.* A compendium of canine normal tissue gene expression. *PLoS ONE* **6**,
doi:10.1371/journal.pone.0017107 (2011).
- 484 Shay, T. *et al.* Conservation and divergence in the transcriptional programs of the human
and mouse immune systems. *Proceedings of the National Academy of Sciences of the
United States of America* **110**, 2946-2951, doi:10.1073/pnas.1222738110 (2013).
- 485 Lin, S. *et al.* Comparison of the transcriptional landscapes between human and mouse
tissues. *Proceedings of the National Academy of Sciences of the United States of America*
111, 17224-17229, doi:10.1073/pnas.1413624111 (2014).
- 486 Lugovtsev, V. Y., Melnyk, D. & Weir, J. P. Heterogeneity of the MDCK Cell Line and
Its Applicability for Influenza Virus Research. *PLoS ONE* **8**,
doi:10.1371/journal.pone.0075014 (2013).
- 487 Barber, M. R. W. *et al.* Identification of avian RIG-I responsive genes during influenza
infection. *Molecular Immunology* **54**, 89-97, doi:10.1016/j.molimm.2012.10.038 (2013).
- 488 Ranaware, P. B. *et al.* Genome wide host gene expression analysis in chicken lungs
infected with avian influenza viruses. *PLoS ONE* **11**, doi:10.1371/journal.pone.0153671
(2016).
- 489 Cheng, Y. *et al.* Chicken STING mediates activation of the IFN gene independently of
the RIG-I gene. *Journal of Immunology* **195**, 3922-3936, doi:10.4049/jimmunol.1500638
(2015).
- 490 Johansson, M. *et al.* The two variants of oxysterol binding protein-related protein-1
display different tissue expression patterns, have different intracellular localization, and
are functionally distinct. *Molecular Biology of the Cell* **14**, 903-915,
doi:10.1091/mbc.E02-08-0459 (2003).
- 491 Lee, S. *et al.* Sterol-dependent nuclear import of ORP1S promotes LXR regulated trans-
activation of apoE. *Experimental Cell Research* **318**, 2128-2142,
doi:10.1016/j.yexcr.2012.06.012 (2012).
- 492 Jansen, M. *et al.* Role of ORPs in Sterol Transport from Plasma Membrane to ER and
Lipid Droplets in Mammalian Cells. *Traffic* **12**, 218-231, doi:10.1111/j.1600-
0854.2010.01142.x (2011).
- 493 Sun, X. & Whittaker, G. R. Role for Influenza Virus Envelope Cholesterol in Virus Entry
and Infection. *Journal of Virology* **77**, 12543-12551, doi:10.1128/JVI.77.23.12543-
12551.2003 (2003).
- 494 Barman, S. & Nayak, D. P. Lipid raft disruption by cholesterol depletion enhances
influenza A virus budding from MDCK cells. *Journal of Virology* **81**, 12169-12178,
doi:10.1128/JVI.00835-07 (2007).

- 495 Kawaguchi, A., Hirohama, M., Harada, Y., Osari, S. & Nagata, K. Influenza Virus
Induces Cholesterol-Enriched Endocytic Recycling Compartments for Budozone
Formation via Cell Cycle-Independent Centrosome Maturation. *PLoS Pathogens* **11**,
doi:10.1371/journal.ppat.1005284 (2015).
- 496 Rocha, N. *et al.* Cholesterol sensor ORP1L contacts the ER protein VAP to control Rab7-
RILP-p150Glued and late endosome positioning. *Journal of Cell Biology* **185**, 1209-
1225, doi:10.1083/jcb.200811005 (2009).
- 497 Johansson, M. *et al.* Activation of endosomal dynein motors by stepwise assembly of
Rab7-RILP-p150Glued, ORP1L, and the receptor β III spectrin. *Journal of Cell Biology*
176, 459-471, doi:10.1083/jcb.200606077 (2007).
- 498 Shah, A. H., Cianciola, N. L., Mills, J. L., Sönnichsen, F. D. & Carlin, C. Adenovirus
RiD α regulates endosome maturation by mimicking GTP-Rab7. *Journal of Cell Biology*
179, 965-980, doi:10.1083/jcb.200702187 (2007).
- 499 Liu, S. L. *et al.* Three-dimensional tracking of rab5-and rab7-associated infection process
of influenza virus. *Small* **10**, 4746-4753, doi:10.1002/sml.201400944 (2014).
- 500 Feeley, E. M. *et al.* IFITM3 inhibits influenza a virus infection by preventing cytosolic
entry. *PLoS Pathogens* **7**, doi:10.1371/journal.ppat.1002337 (2011).
- 501 Amini-Bavil-Olyaei, S. *et al.* The antiviral effector IFITM3 disrupts intracellular
cholesterol homeostasis to block viral entry. *Cell Host and Microbe* **13**, 452-464,
doi:10.1016/j.chom.2013.03.006 (2013).
- 502 Smith, J. *et al.* A comparative analysis of host responses to avian influenza infection in
ducks and chickens highlights a role for the interferon-induced transmembrane proteins
in viral resistance. *BMC Genomics* **16**, doi:10.1186/s12864-015-1778-8 (2015).
- 503 Blyth, G. A. D., Chan, W. F., Webster, R. G. & Magor, K. E. Duck interferon-inducible
transmembrane protein 3 mediates restriction of influenza viruses. *Journal of Virology*
90, 103-116, doi:10.1128/JVI.01593-15 (2016).
- 504 Chi, X., Wang, S., Huang, Y., Stamnes, M. & Chen, J. L. Roles of Rho GTPases in
intracellular transport and cellular transformation. *International Journal of Molecular
Sciences* **14**, 7089-7108, doi:10.3390/ijms14047089 (2013).
- 505 Davies, S. M. K. *et al.* MRPS27 is a pentatricopeptide repeat domain protein required for
the translation of mitochondrially encoded proteins. *FEBS Letters* **586**, 3555-3561,
doi:10.1016/j.febslet.2012.07.043 (2012).
- 506 Richter, U., Lahtinen, T., Marttinen, P., Suomi, F. & Battersby, B. J. Quality control of
mitochondrial protein synthesis is required for membrane integrity and cell fitness.
Journal of Cell Biology **211**, 373-389, doi:10.1083/jcb.201504062 (2015).
- 507 Anand, S. K. & Tikoo, S. K. Viruses as modulators of mitochondrial functions. *Advances
in Virology* **2013**, doi:10.1155/2013/738794 (2013).
- 508 Sánchez-Alcázar, J. A., Ault, J. G., Khodjakov, A. & Schneider, E. Increased
mitochondrial cytochrome c levels and mitochondrial hyperpolarization precede
camptothecin-induced apoptosis in Jurkat cells. *Cell Death and Differentiation* **7**, 1090-
1100 (2000).
- 509 Kroemer, G., Galluzzi, L. & Brenner, C. Mitochondrial membrane permeabilization in
cell death. *Physiological Reviews* **87**, 99-163, doi:10.1152/physrev.00013.2006 (2007).
- 510 Öhman, T., Rintahaka, J., Kalkkinen, N., Matikainen, S. & Nyman, T. A. Actin and RIG-
I/MAVS signaling components translocate to mitochondria upon influenza a virus

- infection of human primary macrophages. *Journal of Immunology* **182**, 5682-5692, doi:10.4049/jimmunol.0803093 (2009).
- 511 Henkel, M. *et al.* Proapoptotic influenza A virus protein PB1-F2 forms a nonselective ion channel. *PLoS ONE* **5**, doi:10.1371/journal.pone.0011112 (2010).
- 512 Pasricha, G., Mishra, A. C. & Chakrabarti, A. K. Comprehensive global amino acid sequence analysis of PB1F2 protein of influenza A H5N1 viruses and the influenza A virus subtypes responsible for the 20th-century pandemics. *Influenza and other Respiratory Viruses* **7**, 497-505, doi:10.1111/j.1750-2659.2012.00400.x (2013).
- 513 Schaefer, L., Beermann, M. L. & Miller, J. B. Coding sequence, genomic organization, chromosomal localization, and expression pattern of the signalosome component Cops2: The mouse homologue of *Drosophila* alien. *Genomics* **56**, 310-316, doi:10.1006/geno.1998.5728 (1999).
- 514 Zhang, W. *et al.* Cops2 promotes pluripotency maintenance by Stabilizing Nanog Protein and Repressing Transcription. *Scientific Reports* **6**, doi:10.1038/srep26804 (2016).
- 515 Koyuncu, O. O., Song, R., Greco, T. M., Cristea, I. M. & Enquist, L. W. The number of alphaherpesvirus particles infecting axons and the axonal protein repertoire determines the outcome of neuronal infection. *mBio* **6**, doi:10.1128/mBio.00276-15 (2015).
- 516 Marques, R. E., Guabiraba, R., Russo, R. C. & Teixeira, M. M. Targeting CCL5 in inflammation. *Expert Opinion on Therapeutic Targets* **17**, 1439-1460, doi:10.1517/14728222.2013.837886 (2013).
- 517 Thomas, L. H., Friedland, J. S., Sharland, M. & Becker, S. Respiratory syncytial virus-induced RANTES production from human bronchial epithelial cells is dependent on nuclear factor- κ B nuclear binding and is inhibited by adenovirus-mediated expression of inhibitor of κ B α . *Journal of Immunology* **161**, 1007-1016 (1998).
- 518 Phung, T. T. *et al.* Key role of regulated upon activation normal T-cell expressed and secreted, nonstructural protein1 and myeloperoxidase in cytokine storm induced by influenza virus PR-8 (A/H1N1) infection in A549 bronchial epithelial cells. *Microbiology and immunology* **55**, 874-884, doi:10.1111/j.1348-0421.2011.00396.x (2011).
- 519 Wang, J. *et al.* Innate immune response of human alveolar macrophages during influenza a infection. *PLoS ONE* **7**, doi:10.1371/journal.pone.0029879 (2012).
- 520 Korteweg, C. & Gu, J. Pathology, molecular biology, and pathogenesis of avian influenza A (H5N1) infection in humans. *American Journal of Pathology* **172**, 1155-1170, doi:10.2353/ajpath.2008.070791 (2008).
- 521 Murooka, T. T. *et al.* CCL5-CCR5-mediated apoptosis in T cells: Requirement for glycosaminoglycan binding and CCL5 aggregation. *Journal of Biological Chemistry* **281**, 25184-25194, doi:10.1074/jbc.M603912200 (2006).
- 522 Tyner, J. W. *et al.* CCL5-CCR5 interaction provides antiapoptotic signals for macrophage survival during viral infection. *Nature Medicine* **11**, 1180-1187, doi:10.1038/nm1303 (2005).
- 523 Lu, Z. & Xu, S. ERK1/2 MAP kinases in cell survival and apoptosis. *IUBMB Life* **58**, 621-631, doi:10.1080/15216540600957438 (2006).
- 524 Mebratu, Y. & Tesfagzi, Y. How ERK1/2 activation controls cell proliferation and cell death is subcellular localization the answer? *Cell Cycle* **8**, 1168-1175 (2009).
- 525 Ritter, J. B., Wahl, A. S., Freund, S., Genzel, Y. & Reichl, U. Metabolic effects of influenza virus infection in cultured animal cells: Intra- and extracellular metabolite profiling. *BMC Systems Biology* **4**, doi:10.1186/1752-0509-4-61 (2010).

- 526 Janke, R., Genzel, Y., Wetzel, M. & Reichl, U. Effect of influenza virus infection on key metabolic enzyme activities in MDCK cells. *BMC proceedings* **5 Suppl 8**, P129, doi:10.1186/1753-6561-5-S8-P129 (2011).
- 527 Smallwood, H. S. *et al.* Targeting Metabolic Reprogramming by Influenza Infection for Therapeutic Intervention. *Cell Reports* **19**, 1640-1653, doi:10.1016/j.celrep.2017.04.039 (2017).
- 528 Meyer, L. *et al.* Transcriptomic profiling of a chicken lung epithelial cell line (CLEC213) reveals a mitochondrial respiratory chain activity boost during influenza virus infection. *PLoS ONE* **12**, doi:10.1371/journal.pone.0176355 (2017).
- 529 Ilyushina, N. A. *et al.* Adaptation of pandemic H1N1 influenza viruses in mice. *Journal of Virology* **84**, 8607-8616, doi:10.1128/JVI.00159-10 (2010).
- 530 Sakabe, S., Ozawa, M., Takano, R., Iwastuki-Horimoto, K. & Kawaoka, Y. Mutations in PA, NP, and HA of a pandemic (H1N1) 2009 influenza virus contribute to its adaptation to mice. *Virus Research* **158**, 124-129, doi:10.1016/j.virusres.2011.03.022 (2011).
- 531 Murray, P. J. *et al.* Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity* **41**, 14-20, doi:10.1016/j.immuni.2014.06.008 (2014).
- 532 Pearce, E. L. & Pearce, E. J. Metabolic pathways in immune cell activation and quiescence. *Immunity* **38**, 633-643, doi:10.1016/j.immuni.2013.04.005 (2013).
- 533 N, A. G. & Castrillo, A. Liver X receptors as regulators of macrophage inflammatory and metabolic pathways. *Biochimica et biophysica acta* **1812**, 982-994, doi:10.1016/j.bbadis.2010.12.015 (2011).
- 534 Nicol, M. Q. & Dutia, B. M. The role of macrophages in influenza A virus infection. *Future Virology* **9**, 847-862, doi:10.2217/fvl.14.65 (2014).
- 535 Lee, S. M. Y., Dutry, I. & Malik Peiris, J. S. Editorial: Macrophage heterogeneity and responses to influenza virus infection. *Journal of Leukocyte Biology* **92**, 1-4, doi:10.1189/jlb.0312130 (2012).
- 536 Pawelek, K. A., Dor, D., Jr., Salmeron, C. & Handel, A. Within-host models of high and low pathogenic influenza virus infections: The role of macrophages. *PLoS ONE* **11**, doi:10.1371/journal.pone.0150568 (2016).
- 537 Andres-Terre, M. *et al.* Integrated, Multi-cohort Analysis Identifies Conserved Transcriptional Signatures across Multiple Respiratory Viruses. *Immunity* **43**, 1199-1211, doi:10.1016/j.immuni.2015.11.003 (2015).
- 538 Iwasaki, A. & Pillai, P. S. Innate immunity to influenza virus infection. *Nature reviews. Immunology* **14**, 315-328, doi:10.1038/nri3665 (2014).
- 539 Stifter, S. A. *et al.* Functional Interplay between Type I and II Interferons Is Essential to Limit Influenza A Virus-Induced Tissue Inflammation. *PLoS Pathog* **12**, e1005378, doi:10.1371/journal.ppat.1005378 (2016).
- 540 Lee, S. M., Dutry, I. & Peiris, J. S. Editorial: macrophage heterogeneity and responses to influenza virus infection. *J Leukoc Biol* **92**, 1-4, doi:10.1189/jlb.0312130 (2012).
- 541 Blackwell, T. S. *et al.* NF-kappaB signaling in fetal lung macrophages disrupts airway morphogenesis. *J Immunol* **187**, 2740-2747, doi:10.4049/jimmunol.1101495 (2011).
- 542 Ensan, S. *et al.* Self-renewing resident arterial macrophages arise from embryonic CX3CR1 + precursors and circulating monocytes immediately after birth. *Nature immunology* **17**, 159-168, doi:10.1038/ni.3343 (2016).
- 543 Schoggins, J. W. & Rice, C. M. Interferon-stimulated genes and their antiviral effector functions. *Curr Opin Virol* **1**, 519-525, doi:10.1016/j.coviro.2011.10.008 (2011).

- 544 Chawla-Sarkar, M. *et al.* Apoptosis and interferons: role of interferon-stimulated genes as mediators of apoptosis. *Apoptosis : an international journal on programmed cell death* **8**, 237-249 (2003).
- 545 Nordmann, A., Wixler, L., Boergeling, Y., Wixler, V. & Ludwig, S. A new splice variant of the human guanylate-binding protein 3 mediates anti-influenza activity through inhibition of viral transcription and replication. *FASEB Journal* **26**, 1290-1300, doi:10.1096/fj.11-189886 (2012).
- 546 Lubeseder-Martellato, C. *et al.* Guanylate-binding protein-1 expression is selectively induced by inflammatory cytokines and is an activation marker of endothelial cells during inflammatory diseases. *American Journal of Pathology* **161**, 1749-1759 (2002).
- 547 Pichlmair, A. *et al.* IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nature immunology* **12**, 624-630, doi:10.1038/ni.2048 (2011).
- 548 Pinto, A. K. *et al.* Human and murine IFIT1 proteins do not restrict infection of negative-sense RNA viruses of the orthomyxoviridae, Bunyaviridae, and filoviridae families. *Journal of Virology* **89**, 9465-9476, doi:10.1128/JVI.00996-15 (2015).
- 549 Li, H. T. *et al.* The interaction between interferon-induced protein with tetratricopeptide repeats-1 and eukaryotic elongation factor-1A. *Molecular and Cellular Biochemistry* **337**, 101-110, doi:10.1007/s11010-009-0289-9 (2010).
- 550 Sedger, L. M. & McDermott, M. F. TNF and TNF-receptors: From mediators of cell death and inflammation to therapeutic giants - past, present and future. *Cytokine and Growth Factor Reviews*, doi:10.1016/j.cytogfr.2014.07.016 (2014).
- 551 McNelis, J. C. & Olefsky, J. M. Macrophages, immunity, and metabolic disease. *Immunity* **41**, 36-48, doi:10.1016/j.immuni.2014.05.010 (2014).
- 552 Biswas, S. K. & Mantovani, A. Orchestration of metabolism by macrophages. *Cell metabolism* **15**, 432-437, doi:10.1016/j.cmet.2011.11.013 (2012).
- 553 Zhu, L., Zhao, Q., Yang, T., Ding, W. & Zhao, Y. Cellular metabolism and macrophage functional polarization. *International reviews of immunology* **34**, 82-100, doi:10.3109/08830185.2014.969421 (2015).
- 554 Kelly, B. & O'Neill, L. A. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. *Cell Res* **25**, 771-784, doi:10.1038/cr.2015.68 (2015).
- 555 Galvan-Pena, S. & O'Neill, L. A. Metabolic reprogramming in macrophage polarization. *Frontiers in immunology* **5**, 420, doi:10.3389/fimmu.2014.00420 (2014).
- 556 Haschemi, A. *et al.* The sedoheptulose kinase CARKL directs macrophage polarization through control of glucose metabolism. *Cell metabolism* **15**, 813-826, doi:10.1016/j.cmet.2012.04.023 (2012).
- 557 Stegemann-Koniszewski, S. *et al.* Alveolar Type II Epithelial Cells Contribute to the Anti-Influenza A Virus Response in the Lung by Integrating Pathogen- and Microenvironment-Derived Signals. *MBio* **7**, doi:10.1128/mBio.00276-16 (2016).
- 558 Waithman, J. & Mintern, J. D. Dendritic cells and influenza A virus infection. *Virulence* **3**, 603-608, doi:10.4161/viru.21864 (2012).
- 559 Pulendran, B. & Maddur, M. S. Innate immune sensing and response to influenza. *Current topics in microbiology and immunology* **386**, 23-71, doi:10.1007/82_2014_405 (2015).
- 560 Graham, A. C., Temple, R. M. & Obar, J. J. Mast cells and influenza a virus: association with allergic responses and beyond. *Frontiers in immunology* **6**, 238, doi:10.3389/fimmu.2015.00238 (2015).

- 561 Ivan, F. X. *et al.* Differential pulmonary transcriptomic profiles in murine lungs infected with low and highly virulent influenza H3N2 viruses reveal dysregulation of TREM1 signaling, cytokines, and chemokines. *Functional & integrative genomics* **12**, 105-117, doi:10.1007/s10142-011-0247-y (2012).
- 562 Rosseau, S. *et al.* Comparative transcriptional profiling of the lung reveals shared and distinct features of Streptococcus pneumoniae and influenza A virus infection. *Immunology* **120**, 380-391, doi:10.1111/j.1365-2567.2006.02514.x (2007).
- 563 Qiu, X. *et al.* Diversity in compartmental dynamics of gene regulatory networks: The immune response in primary influenza a infection in mice. *PLoS ONE* **10**, doi:10.1371/journal.pone.0138110 (2015).
- 564 Boon, A. C. M. *et al.* H5n1 influenza virus pathogenesis in genetically diverse mice is mediated at the level of viral load. *mBio* **2**, doi:10.1128/mBio.00171-11 (2011).
- 565 Boon, A. C. M. *et al.* Host genetic variation affects resistance to infection with a highly pathogenic H5N1 influenza A virus in mice. *Journal of Virology* **83**, 10417-10426, doi:10.1128/JVI.00514-09 (2009).
- 566 Morrison, J. *et al.* H7N9 and other pathogenic avian influenza viruses elicit a three-pronged transcriptomic signature that is reminiscent of 1918 influenza virus and is associated with lethal outcome in mice. *Journal of virology* **88**, 10556-10568, doi:10.1128/JVI.00570-14 (2014).
- 567 van Riel, D. *et al.* Highly pathogenic avian influenza virus H5N1 infects alveolar macrophages without virus production or excessive TNF-alpha induction. *PLoS Pathogens* **7**, doi:10.1371/journal.ppat.1002099 (2011).
- 568 Cheung, C. Y. *et al.* Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: A mechanism for the unusual severity of human disease? *Lancet* **360**, 1831-1837, doi:10.1016/S0140-6736(02)11772-7 (2002).
- 569 Tate, M. D., Schilter, H. C., Brooks, A. G. & Reading, P. C. Responses of mouse airway epithelial cells and alveolar macrophages to virulent and avirulent strains of influenza a virus. *Viral Immunology* **24**, 77-88, doi:10.1089/vim.2010.0118 (2011).
- 570 Lin, K. L., Suzuki, Y., Nakano, H., Ramsburg, E. & Gunn, M. D. CCR2+ monocyte-derived dendritic cells and exudate macrophages produce influenza-induced pulmonary immune pathology and mortality. *Journal of Immunology* **180**, 2562-2572 (2008).
- 571 Herold, S. *et al.* Lung epithelial apoptosis in influenza virus pneumonia: The role of macrophage-expressed TNF-related apoptosis-inducing ligand. *Journal of Experimental Medicine* **205**, 3065-3077, doi:10.1084/jem.20080201 (2008).
- 572 Lee, S. M. Y. *et al.* Systems-level comparison of host-responses elicited by avian H5N1 and seasonal H1N1 influenza viruses in primary human macrophages. *PLoS ONE* **4**, doi:10.1371/journal.pone.0008072 (2009).
- 573 Yu, W. C. L. *et al.* Viral replication and innate host responses in primary human alveolar epithelial cells and alveolar macrophages infected with influenza H5N1 and H1N1 viruses. *Journal of Virology* **85**, 6844-6855, doi:10.1128/JVI.02200-10 (2011).
- 574 Hui, K. P. Y. *et al.* Induction of proinflammatory cytokines in primary human macrophages by influenza A virus (H5N1) is selectively regulated by IFN regulatory factor 3 and p38 MAPK. *Journal of Immunology* **182**, 1088-1098 (2009).
- 575 Anikina, A. G., Shkurupii, V. A., Potapova, O. V., Kovner, A. V. & Shestopalov, A. M. Expression of profibrotic growth factors and their receptors by mouse lung macrophages and fibroblasts under conditions of acute viral inflammation in influenza A/H5N1 virus.

- Bulletin of Experimental Biology and Medicine* **156**, 833-837, doi:10.1007/s10517-014-2463-7 (2014).
- 576 Perrone, L. A., Plowden, J. K., García-Sastre, A., Katz, J. M. & Tumpey, T. M. H5N1 and 1918 pandemic influenza virus infection results in early and excessive infiltration of macrophages and neutrophils in the lungs of mice. *PLoS Pathogens* **4**, doi:10.1371/journal.ppat.1000115 (2008).
- 577 Us, D. Cytokine storm in avian influenza. *Mikrobiyoloji Bulteni* **42**, 365-380 (2008).
- 578 Szretter, K. J. *et al.* Role of host cytokine responses in the pathogenesis of avian H5N1 influenza viruses in mice. *Journal of Virology* **81**, 2736-2744, doi:10.1128/JVI.02336-06 (2007).
- 579 Chan, R. W. Y., Leung, C. Y. H., Nicholls, J. M., Peiris, J. S. M. & Chan, M. C. W. Proinflammatory Cytokine Response and Viral Replication in Mouse Bone Marrow Derived Macrophages Infected with Influenza H1N1 and H5N1 Viruses. *PLoS ONE* **7**, doi:10.1371/journal.pone.0051057 (2012).
- 580 Hang, D. T. T., Song, J. Y., Kim, M. Y., Park, J. W. & Shin, Y. K. Involvement of NF- κ B in changes of IFN- γ -induced CIITA/MHC-II and iNOS expression by influenza virus in macrophages. *Molecular Immunology* **48**, 1253-1262, doi:10.1016/j.molimm.2011.03.010 (2011).
- 581 Cannon, G., Callahan, M. A., Gronemus, J. Q. & Lowy, R. J. Early activation of MAP kinases by influenza A virus X-31 in murine macrophage cell lines. *PLoS ONE* **9**, doi:10.1371/journal.pone.0105385 (2014).
- 582 McKinney, L. C., Galliger, S. J. & Lowy, R. J. Active and inactive influenza virus induction of tumor necrosis factor- α and nitric oxide in J774.1 murine macrophages: Modulation by interferon- γ and failure to induce apoptosis. *Virus Research* **97**, 117-126, doi:10.1016/j.virusres.2003.08.007 (2003).
- 583 To, E. E., Broughton, B. R. S., Hendricks, K. S., Vlahos, R. & Selemidis, S. Influenza A virus and TLR7 activation potentiate NOX2 oxidase-dependent ROS production in macrophages. *Free Radical Research* **48**, 940-947, doi:10.3109/10715762.2014.927579 (2014).
- 584 Hoeve, M. A., Nash, A. A., Jackson, D., Randall, R. E. & Dransfield, I. Influenza virus A infection of human monocyte and macrophage subpopulations reveals increased susceptibility associated with cell differentiation. *PLoS ONE* **7**, doi:10.1371/journal.pone.0029443 (2012).
- 585 Monteerarat, Y. *et al.* Induction of TNF- α in human macrophages by avian and human influenza viruses. *Archives of Virology* **155**, 1273-1279, doi:10.1007/s00705-010-0716-y (2010).
- 586 Österlund, P. *et al.* Pandemic H1N1 2009 influenza a virus induces weak cytokine responses in human macrophages and dendritic cells and is highly sensitive to the antiviral actions of interferons. *Journal of Virology* **84**, 1414-1422, doi:10.1128/JVI.01619-09 (2010).
- 587 Zhao, C. *et al.* Pro-inflammatory cytokine dysregulation is associated with novel avian influenza a (H7N9) virus in primary human macrophages. *Journal of General Virology* **97**, 299-305, doi:10.1099/jgv.0.000357 (2016).
- 588 Zhou, J. *et al.* Differential expression of chemokines and their receptors in adult and neonatal macrophages infected with human or avian influenza viruses. *Journal of Infectious Diseases* **194**, 61-70, doi:10.1086/504690 (2006).

- 589 Ravi, L. I. *et al.* A systems-based approach to analyse the host response in murine lung
macrophages challenged with respiratory syncytial virus. *BMC Genomics* **14**, 190,
doi:10.1186/1471-2164-14-190 (2013).
- 590 Lay, M. K., Bueno, S. M., Gálvez, N., Riedel, C. A. & Kalergis, A. M. New insights on
the viral and host factors contributing to the airway pathogenesis caused by the
respiratory syncytial virus. *Critical Reviews in Microbiology* **42**, 800-812,
doi:10.3109/1040841X.2015.1055711 (2016).
- 591 Van Drunen Littel-Van Den Hurk, S. & Watkiss, E. R. Pathogenesis of respiratory
syncytial virus. *Current Opinion in Virology* **2**, 300-305,
doi:10.1016/j.coviro.2012.01.008 (2012).
- 592 Short, K. R. *et al.* One health, multiple challenges: The inter-species transmission of
influenza A virus. *One health* **1**, 1-13, doi:10.1016/j.onehlt.2015.03.001 (2015).
- 593 Wang, W. *et al.* Glycosylation at 158N of the hemagglutinin protein and receptor binding
specificity synergistically affect the antigenicity and immunogenicity of a live attenuated
H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets. *Journal of Virology* **84**, 6570-6577,
doi:10.1128/JVI.00221-10 (2010).
- 594 Zhao, X. *et al.* PI3K/Akt signaling pathway modulates influenza virus induced mouse
alveolar macrophage polarization to M1/M2b. *PLoS ONE* **9**,
doi:10.1371/journal.pone.0104506 (2014).
- 595 Raman, K. Construction and analysis of protein-protein interaction networks. *Automated
experimentation* **2**, 2, doi:10.1186/1759-4499-2-2 (2010).
- 596 Kemmeren, P. *et al.* Protein interaction verification and functional annotation by
integrated analysis of genome-scale data. *Molecular cell* **9**, 1133-1143 (2002).
- 597 Fievet, B. T. *et al.* Systematic genetic interaction screens uncover cell polarity regulators
and functional redundancy. *Nature Cell Biology* **15**, 103-112, doi:10.1038/ncb2639
(2013).
- 598 Shoemaker, B. A. & Panchenko, A. R. Deciphering protein-protein interactions. Part I.
Experimental techniques and databases. *PLoS Computational Biology* **3**, 0337-0344,
doi:10.1371/journal.pcbi.0030042 (2007).
- 599 Syafrizayanti, Betzen, C., Hoheisel, J. D. & Kastelic, D. Methods for analyzing and
quantifying protein-protein interaction. *Expert Review of Proteomics* **11**, 107-120 (2014).
- 600 Keskin, O., Tuncbag, N. & Gursoy, A. Predicting Protein-Protein Interactions from the
Molecular to the Proteome Level. *Chemical reviews* **116**, 4884-4909,
doi:10.1021/acs.chemrev.5b00683 (2016).
- 601 Pattin, K. A. & Moore, J. H. Role for protein-protein interaction databases in human
genetics. *Expert Review of Proteomics* **6**, 647-659 (2009).
- 602 Alcaraz, N. *et al.* KeyPathwayMiner 4.0: Condition-specific pathway analysis by
combining multiple omics studies and networks with Cytoscape. *BMC Systems Biology* **8**,
doi:10.1186/s12918-014-0099-x (2014).
- 603 Ulitsky, I., Krishnamurthy, A., Karp, R. M. & Shamir, R. DEGAS: de novo discovery of
dysregulated pathways in human diseases. *PLoS One* **5**, e13367,
doi:10.1371/journal.pone.0013367 (2010).
- 604 Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A. F. Discovering regulatory and
signalling circuits in molecular interaction networks. *Bioinformatics* **18 Suppl 1**, S233-
240 (2002).

- 605 Gaballa, A. *et al.* Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in Bacilli. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6482-6486, doi:10.1073/pnas.1000928107 (2010).
- 606 Vlasblom, J. *et al.* Novel function discovery with GeneMANIA: A new integrated resource for gene function prediction in Escherichia coli. *Bioinformatics* **31**, 306-310, doi:10.1093/bioinformatics/btu671 (2014).
- 607 Lee, T. L., Raygada, M. J. & Rennert, O. M. Integrative gene network analysis provides novel regulatory relationships, genetic contributions and susceptible targets in autism spectrum disorders. *Gene* **496**, 88-96, doi:10.1016/j.gene.2012.01.020 (2012).
- 608 Yao, L., Evans, J. A. & Rzhetsky, A. Novel opportunities for computational biology and sociology in drug discovery. *Trends in Biotechnology* **27**, 531-540, doi:10.1016/j.tibtech.2009.06.003 (2009).
- 609 Yang, P. *et al.* Angiotensin-converting enzyme 2 (ACE2) mediates influenza H7N9 virus-induced acute lung injury. *Scientific Reports* **4**, 7027, doi:10.1038/srep07027 (2014).
- 610 Yan, Y. W. *et al.* Angiotensin II receptor blocker as a novel therapy in acute lung injury induced by avian influenza A H5N1 virus infection in mouse. *Science China Life Sciences* **58**, 208-211, doi:10.1007/s11427-015-4814-7 (2015).
- 611 Eng, E. O., Chew, J. S. W., Jin, P. L. & Chua, R. C. S. In vitro inhibition of human influenza A virus replication by chloroquine. *Virology Journal* **3** (2006).
- 612 Yan, Y. *et al.* Anti-malaria drug chloroquine is highly effective in treating avian influenza A H5N1 virus infection in an animal model. *Cell Research* **23**, 300-302 (2013).
- 613 Wu, L. *et al.* Chloroquine enhances replication of influenza A virus A/WSN/33 (H1N1) in dose-, time-, and MOI-dependent manners in human lung epithelial cells A549. *Journal of Medical Virology* **87**, 1096-1103, doi:10.1002/jmv.24135 (2015).

Appendix

Appendix A:

- List of studies and genes from the siRNA screening studies and interactome studies
- List of IHFs mapped to the three network growing tools
- List of randomized human genes and IHFs for growing networks
- List of DEGs in at least 3 viruses in A549 cell infections
- Customized Perl scripts

Appendix B:

- List of predicted FDA approved drugs and their host targets
- List of drugs predicated or tested against IAVs
- List of ranked FDA approved drugs by our method

Appendix C:

- List of non-annotated probe sets from canine and chicken affymetrix arrays
- Annotated probe sets by our method
- List of chicken probe sets/genes orthologs to human
- List of canine probe sets/genes orthologs to human
- List of probe sets/genes orthologs to the three species

Appendix D:

- Expression signals of IFNs in A549, CEF, MDCK and Macrophages

Appendix E:

- comparison of expression changes in IAV and RSV infections in pulmonary mouse macrophages

Appendix F:

- Search space comparison of the three gene network growing tools
- Network topology comparison of GeneMANIA and STRING
- Topological characteristics of intersecting and non-intersecting genes
- % of one siRNA screening study to detect genes from other siRNA screen experiments

Appendix G:

- Pathways and GO BP from known and new candidate IHFs

SCIENTIFIC REPORTS

OPEN

Benchmarking selected computational gene network growing tools in context of virus-host interactions

Received: 7 September 2016
Accepted: 7 June 2017
Published online: 19 July 2017

Biruhalem Taye^{1,2,3,4}, Candida Vaz¹, Vivek Tanavde^{1,2}, Vladimir A. Kuznetsov^{1,4},
Frank Eisenhaber^{1,4,5}, Richard J. Sugrue⁶ & Sebastian Maurer-Stroh^{1,4,6}

Several available online tools provide network growing functions where an algorithm utilizing different data sources suggests additional genes/proteins that should connect an input gene set into functionally meaningful networks. Using the well-studied system of influenza host interactions, we compare the network growing function of two free tools GeneMANIA and STRING and the commercial IPA for their performance of recovering known influenza A virus host factors previously identified from siRNA screens. The result showed that given small (~30 genes) or medium (~150 genes) input sets all three network growing tools detect significantly more known host factors than random human genes with STRING overall performing strongest. Extending the networks with all the three tools significantly improved the detection of GO biological processes of known host factors compared to not growing networks. Interestingly, the rate of identification of true host factors using computational network growing is equal or better to doing another experimental siRNA screening study which could also be true and applied to other biological pathways/processes.

Gene and/or protein networks are used as a cardinal representation of various types of biological processes and help in the prediction of molecular and cellular function¹. It has been indicated that interacting genes/proteins may be part of the same pathway or biological process and, in larger scale, work together in similar cellular functions^{2,3}. Reliable prediction and precise treatment of complex cellular functions require knowing the network of as many as possible genes and their products (e.g., proteins) connected in functional pathways. Pairwise interactions at the gene and protein levels have been analyzed extensively, often from low or high-throughput experimental screens, with the links (annotated as functional interactions) available in online databases (reviewed in ref. 4). However, these methods do not fully investigate the highly complex interaction patterns in cellular systems and variability in terms of their accuracy and reproducibility⁵ is encountered. Additionally, interactions seen *in vitro*, especially in large-scale screens, may not occur *in vivo* due to spatial and temporal constraints. At the same time, the exact number of protein-protein interaction (PPI) in human is not known, and the available data is estimated to represent approximately 10% of the total PPIs in human⁶. Hence, computational PPI predictions have become increasingly important for detection of new interactions and protein networks. There are several computational gene and/or protein network databases that combine different experimental and computationally predicted interactions through the integration of PPI information that are obtained from different sources (reviewed in refs 7–9).

Network growing functions allow extension of networks with additional related genes and are important to improve the understanding of the greater functional context and organization of genes and/or proteins, as well as

¹Bioinformatics Institute, A*STAR, 30 Biopolis Street #07-01 Matrix, Singapore, 138671, Singapore. ²School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore, 637551, Singapore. ³Aklilu Lemma Institute of Pathobiology, Addis Ababa University, P.O. BOX 1176, Addis Ababa, Ethiopia. ⁴Department of Biological Sciences, National University of Singapore, 8 Medical Drive, Singapore, 117597, Singapore. ⁵School of Computer Engineering, Nanyang Technological University, 50 Nanyang Drive, Singapore, 637553, Singapore. ⁶National Public Health Laboratory, Ministry of Health, 3 Biopolis Drive, Synapse #05-14/16, Singapore, 138623, Singapore. ⁷Institute of Medical Biology, A*STAR, 8A Biomedical Grove, #06-06 Immunos, Singapore, 138648, Singapore. Correspondence and requests for materials should be addressed to B.T. (email: biruhalem@bi.a-star.edu.sg)

RESEARCH ARTICLE

Open Access

Systems-based approach to examine the cytokine responses in primary mouse lung macrophages infected with low pathogenic avian Influenza virus circulating in South East Asia

Biruhalem Taye^{1,2,3}, Hui Chen^{2,8}, Myint Zu Myaing², Boon Huan Tan^{4,5}, Sebastian Maurer-Stroh^{1,6,7} and Richard J. Sugrue^{2*}

Abstract

Background: Influenza A virus (IAV) is a major public health concern, being responsible for the death of approximately half a million people each year. Zoonotic transmissions of the virus from swine and avian origin have occurred in the past, and can potentially lead to the emergence of new IAV strains in future pandemics. Pulmonary macrophages have been implicated in disease severity in the lower airway, and understanding the host response of macrophages infected with avian influenza viruses should provide new therapeutic strategies.

Results: We used a systems-based approach to investigate the transcriptome response of primary murine lung macrophages (PMΦ) infected with the mouse-adapted H1N1/WSN virus and low pathogenic avian influenza (LPAI) viruses H5N2 and H5N3. The results showed that the LPAI viruses H5N2 and H5N3 can infect PMΦ with similar efficiency to the H1N1/WSN virus. While all viruses induced antiviral responses, the H5N3 virus infection resulted in higher expression levels of cytokines and chemokines associated with inflammatory responses.

Conclusions: The LPAI H5N2 and H5N3 viruses are able to infect murine lung macrophages. However, the H5N3 virus was associated with increased expression of pro-inflammatory mediators. Although the H5N3 virus is capable of inducing high levels of cytokines that are associated with inflammation, this property is distinct from its inability to efficiently replicate in a mammalian host.

Keywords: Macrophages, Avian influenza virus, H5N2, H5N3, Inflammatory response

Background

Influenza A virus (IAV) represents one of the most important biological threats to human health and the global economy [1–3]. It causes yearly seasonal epidemics and periodic pandemics [4]. Seasonal influenza is estimated to be the cause up to 5 million severe cases and approximately 500,000 deaths each year. There have been four previous influenza virus pandemics, which have been collectively responsible for the deaths of more than 50

million people [5–7]. The combination of drug resistance, adverse side effects and financial costs of drug and vaccine development are significant challenges in the prevention of the disease due to newly emerging influenza virus strains [8, 9].

IAV normally circulates in wild birds, but it can also infect domestic poultry, pigs, cats, dogs, horses, bats and humans. Although direct zoonotic transmission to humans is not common, interspecies transmission from swine and avian to humans can occur, and zoonotic transmission can contribute to the emergence of new pandemic viruses [10–12]. Based on disease severity in birds, avian IAV can be classified as either being low

* Correspondence: RSUGRUE@ntu.edu.sg

²School of Biological Science, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Republic of Singapore
Full list of author information is available at the end of the article

© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

Article

Inter-Species Host Gene Expression Differences in Response to Human and Avian Influenza A Virus Strains

Biruhalem Taye ^{1,2,3} , Dawn Yeo ¹, Raphael Tze Chorn Lee ² , Boon Huan Tan ^{4,5}, Richard J. Sng ^{1,6} and Sebastian Maurer-Stroh ^{2,6,*}

¹ School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore; biruhalem@biol.sc.nyu.edu (B.T.); yaoym@biol.sc.nyu.edu (D.Y.)

² Bioinformatics Institute (BI), 30 Biopolis Street, #07-01 Matrix Building, Singapore 136671, Singapore; leeet@biol.sc.nyu.edu

³ Addis Lemma Institute of Pathobiology, Addis Ababa University, Addis Ababa P.O. Box 1176, Ethiopia

⁴ Detection and Diagnostics Laboratory, DSO National Laboratories, 27 Medical Drive, Singapore 117510, Singapore; bhoonhuan@biol.sc.nyu.edu

⁵ LKC School of Medicine, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore

⁶ Department of Biological Sciences (DBS), National University Singapore (NUS), 14 Science Drive 4, Singapore 117543, Singapore

* Correspondence: rjng@nus.edu.sg (R.J.S.); sebastianm@biol.sc.nyu.edu (S.M.-S.); Tel.: +65-6316-2889 (R.J.S.); +65-6678-8377 (S.M.-S.)

Received: 28 September 2017; Accepted: 25 October 2017; Published: 1 November 2017

Abstract: Low pathogenic avian influenza (LPAI) viruses are a source of sporadic human infections and could also contribute to future pandemic outbreaks but little is known about inter-species differences in the host responses to these viruses. Here, we studied host gene expression signatures of cell lines from three species (human, chicken, and canine) in response to six different viruses (H1N1/WSN, H5N2/F59, H5N2/F118, H5N2/F189, H5N3 and H9N2). Comprehensive microarray probe set re-annotation and ortholog mapping of the host genes was necessary to allow comparison over extended functionally annotated gene sets and orthologous pathways. The annotations are made available to the community for commonly used microarray chips. We observe a strong tendency of the response being cell type- rather than virus-specific. In chicken cells, we found up-regulation of host factors inducing virus infectivity (e.g., ovosteryl binding protein like 1A (OSBP1A) and Rho GTPase activating protein 21 (ARHGAP21)) while reducing apoptosis (e.g., mitochondrial ribosomal protein S27 (MRPS27)) and increasing cell proliferation (e.g., COP9 signalosome subunit 2 (COPS2)). On the other hand, increased antiviral, pro-apoptotic and inflammatory signatures have been identified in human cells while cell cycle and metabolic pathways were down-regulated. This signature describes how low pathogenic avian influenza (LPAI) viruses are being tolerated and shed from chicken but potentially causing cellular disruption in mammalian cells.

Keywords: influenza A virus; inter-species transcriptome; apoptosis; metabolic pathways

1. Introduction

Influenza A virus (IAV) is a zoonotic virus that is present in several animal reservoirs (e.g., avian, pigs) [1]. These animal viruses play an important role in the evolution of influenza viruses that circulate with the human population. However, several barriers to infection must be overcome when viruses are transmitted from an avian to a human host. Furthermore, several additional obstacles must be overcome for the virus to be maintained in the new host [2]. In essence, the virus must adapt

Posters, Awards and invited talks

- **Biruhalem Taye**, Chen Hui, Myint Zu Myaing, Boon Huan Tan, Sebastian Maurer-Stroh, and Richard J Sugrue: Systems-based approach to examine the cytokine responses in primary mouse lung macrophages infected with low pathogenic avian Influenza virus circulating in South East Asia. 17th international congress of virology (17 - 21 July 2017, Sands Expo and Convention Center, Singapore).
- **Biruhalem Taye**, Richard J Sugrue, and Sebastian Maurer-Stroh: Benchmarking gene network growing algorithms for detection of known and potentially new influenza A virus therapeutic host target candidates. Option IX for the control of Influenza (24-28 August 2016, Sheraton Grand, Chicago, USA).
- **Biruhalem Taye**, Chen Hui, Myint Zu Myaing, Boon Huan Tan, Sebastian Maurer-Stroh, and Richard J Sugrue: Pathway analysis of host gene expression during infection with a variety of low pathogenic avian and human strains of influenza A virus highlights potential new host targets. The 4th ISIRV-AVG conference (2-4 June 2015, University of Texas at Austin, USA).
- **Biruhalem Taye**, Richard J Sugrue, and Sebastian Maurer-Stroh: Benchmarking gene network growing algorithms for detection of known and potentially new influenza A virus therapeutic host target candidates. International Conference on Bioinformatics (INCOB2016) (21-23 September 2016, Singapore).
- **Biruhalem Taye**, Chen Hui, Myint Zu Myaing, Boon Huan Tan, Sebastian Maurer-Stroh, and Richard J Sugrue: 2015. Consistent up-regulation of TRIM21, SAMHD1, and SP100 upon infection with different influenza A viruses suggest a novel innate immune pathway. International meeting on respiratory pathogens (IMRP)-Singapore (2-4 September 2015, Singapore).