SYSTEMS-LEVEL HOST RESPONSES
TO HIGHLY AND LESS VIRULENT INFLUENZA A (H3N2)
VIRUS INFECTIONS

FRANSISKUS XAVERIUS IVAN

NANYANG TECHNOLOGICAL UNIVERSITY

2013
SYSTEMS-LEVEL HOST RESPONSES
TO HIGHLY AND LESS VIRULENT INFLUENZA A (H3N2) VIRUS INFECTIONS

FRANSISKUS XAVERIUS IVAN
(MCompIT, University of New South Wales)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHYLLOSOPHY
IN COMPUTATION AND SYSTEMS BIOLOGY (CSB)
SINGAPORE-MIT ALLIANCE
NANYANG TECHNOLOGICAL UNIVERSITY
2013
DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis. This thesis has also not been submitted for any degree in any university previously.

________________________________________
Fransiskus Xaverius Ivan
15 November 2013
Acknowledgement

I would like to express my gratitude to:

Asst. Prof. Zheng Jie and Prof. Jagath C. Rajapakse, my advisors from Nanyang Technological University (NTU), Singapore. I heartily thank them for their advice and guidance during my PhD study. I am also thankful for the opportunities to serve as Teaching Assistant for Statistical Learning and Data Mining (course code: SMA5303), Journal Club Coordinator and Research Associate at Bioinformatics Research Centre (BIRC).

Assoc. Prof. Vincent T. K. Chow, my advisor from National University of Singapore (NUS), Singapore. I am very grateful that he has provided a great opportunity for me to carry out biology research related to influenza and neutrophils in his lab, namely Human Genome Laboratory (HGL). I have always been inspired by his vision about influenza research, and I am thankful for his unfailing advice and guidance.

Prof. Roy E. Welsch, my advisor from Massachusetts Institute of Technology (MIT), USA. I sincerely thank him for his advice, feedback and support, which always come in no time. His gentle approach to his students is always memorable, and the way he delivers statistical theory and application into my research is undoubted.

Assoc. Prof. Steve G. Rozen, a collaborator from Duke-NUS Graduate Medical School, Singapore. His broad knowledge in microarray data preprocessing and analysis have been a great help to initiate my genomic research. His advice in the beginning of my PhD study has always been in my heart and overcome my struggles to accomplish what I want to achieve in research.

Assoc. Prof. Bevin Engelward, a collaborator from MIT, USA. It is a great opportunity to collaborate with her and I am thankful for her valuable comments to improve my research manuscripts.

Dr. Teluguakula Narasaraju, Dr. Ramar Perumal, Gordon Xiong, Kai Sen, Huey Hian, Audrey-Ann, Anandi Narayana Moorthy, Prashant Rai, Li Na, Edwin Yang, Hui Ann, Jung Pu, Teo Youjin, Wee Peng, Xie Meilan, Wu Yan, Wai Chii,
Fiona Setiawan, Fabian Hia, great companions in HGL lab, who kindly shared experience in carrying out lab experiments. I especially thank Dr. Narasaraju, who has always been my role model for being a research scientist in biology, and Gordon, whose thesis has been the main resource for my initial works during my PhD study.

Mrs. Phoon Meng Chee, Ms. Lau Suk Hiang, Mrs. Thong Kar Tiang, Ms. Lee Shu Ying, Mr. Chan Yue Ng, Ms. Nalini Srinivasan, and other lab/admin staffs in Yong Loo Lin Medical School, especially in Department of Microbiology. I am very thankful for the great technical and administrative assistance.

Dr. Merlin V. A. James, Dr. Piyushkumar A. Mundra, Dr. Cheng Jierong, Iti Chaturvedi, Alvin Ng, Su Tran To Chinh, and many other graduate students from different research groups at BIRC, NTU, who have been very friendly and helpful. It is a great pleasure to share knowledge in computing and bioinformatics, especially during BIRC Journal Club.

Tan Sing Yau and Kesavan Asaithambi, who always provide unfailing administrative support at BIRC, NTU.

Dr. Dahai Zeng, Min Zin Oo, Yin Lu, and many other colleagues and researchers who are part of Singapore-MIT Alliance (SMA) and Singapore-MIT Alliance for Research and Technology (SMART). It has been a wonderful journey to be together in the research hubs.

Tan Yingrou and Fitri Juniwati, who have been wonderful friends in my spiritual journey as a Catholic.

Warrick Tan, Lee Jun Yan, Benjamin Tan, Jörg Kanonenberg, Angela Shen, Sophia Chew, Lee Shermaine, and many other who I cannot name them all here, my great teammates and coaches in NTU Tennis Team. Playing tennis with you all has been my most joyful and refreshing moments throughout my PhD study. Special to Shermaine, the time that I mentored you in HGL lab is unforgettable.

Singapore-MIT Alliance (SMA), which provides financial and administrative support throughout my PhD study. I thank all SMA staffs, especially Carol Cheng, Jason Chong, Mei Ling, and Jocelyn Sales. I also thank Prof. Gong Zhi Yuan and Prof. Hew Choy Leong, the CSB Program Chairs, who always assured a good care and service for the students in the program.

Mom, who always gives her best prays for my search of wisdom and success; Dad, whom I believe watches over me from His Great Kingdom; Brothers and Sisters, who are always there in my journey and provides unfailing support to their youngest brother; and the last but not the least, Dewi Novita Fauzie, who have been a faithful friend for the whole years since we left high school.
Author’s publications

Journal papers


Conference publications


Contents

Acknowledgement ............................................................. i
Author’s publications ...................................................... iii
Abstract ........................................................................... ix
Summary ........................................................................... x
List of tables ....................................................................... xiii
List of figures ....................................................................... xiv
List of genes/proteins .......................................................... xix
Abbreviations ...................................................................... xxvi

1 Introduction ..................................................................... 1
  1.1 Motivations and objectives .............................................. 2
  1.2 Contributions of the thesis .............................................. 4
  1.3 Background and recent work .......................................... 7
    1.3.1 Influenza A virus (IAV) ........................................... 7
    1.3.2 Host innate immune defenses against IAV infection ....... 12
    1.3.3 Systems biology of host response to IAV infection ..... 16
    1.3.4 Neutrophils ......................................................... 26
    1.3.5 IAV infection in neutrophils .................................... 36
    1.3.6 Induced differentiated myeloid cell lines as an alternative
source for neutrophils .............................................. 37

1.3.7 Computational and statistical approaches ............... 40

1.4 Outline of the thesis ............................................. 49

2 Transcriptional profiling of lungs infected with highly and less virulent IAV (H3N2) viruses 50

2.1 Motivations ......................................................... 50

2.2 Materials and methods ........................................... 51

2.3 Results .............................................................. 55

  2.3.1 Murine lungs infected with LVI and HVI display distinct pathologic features ......................... 55

  2.3.2 Transcriptomic data analyses and kinetics of LVI versus HVI infections ................................. 57

  2.3.3 Early inflammatory and immune responses of LVI versus HVI infections ........................... 60

  2.3.4 HVI infection is characterized by hypercytokinemia and dysregulated TREM1 signaling ............ 60

  2.3.5 Trends of type I IFN-related and apoptotic gene expressions ............................................. 63

  2.3.6 Under-expressed genes in HVI infected lungs are associated with metabolism and developmental processes .................. 67

2.4 Discussion ............................................................ 68

3 High content image approach for estimating the purity of neutrophils in giemsa-stained NPROs 76

3.1 Motivations ............................................................ 77
3.2 Materials and methods ......................................................... 79
3.3 Results .................................................................................. 83
   3.3.1 Diversity of cell stages in NPROs ................................. 83
   3.3.2 Iterative global image thresholding to isolate a single cell
          from cell clusters in high content images of giemsa-stained
          NPROs ............................................................................. 86
   3.3.3 Training the \( k \)-nearest neighbor (\( k \)-NN) classifier and
          selecting features to classify cell images into neutrophils
          and non-neutrophils ......................................................... 88
   3.3.4 Giemsa-staining approach for estimating neutrophil purity
          in NPROs: automated versus manual .......................... 90
   3.3.5 Comparison between giemsa-staining and flow cytometry
          approaches for estimating neutrophil purity in NPROs .... 92
3.4 Discussion ................................................................. 95

4 Transcriptional profiling of neutrophils infected with highly and less
virulent IAV (H3N2) viruses ............................................. 100
   4.1 Motivations ................................................................. 101
   4.2 Materials and methods ................................................ 103
   4.3 Results ............................................................................. 109
      4.3.1 Neutrophils from induced differentiated MPRO cells
             (NPROs) are infectable by IAV ................................. 109
      4.3.2 NPROs do not support active IAV replication ............. 111
      4.3.3 IAV infection of NPROs augments early and late apoptosis 111
      4.3.4 Transcriptomic data analyses of NPROs infected with MPI,
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVI and HVI</td>
<td>113</td>
</tr>
<tr>
<td>4.3.5 Early and stronger type I IFN response characterizes HVI infection of NPROs</td>
<td>114</td>
</tr>
<tr>
<td>4.3.6 Expression of apoptosis-related genes during MPI, LVI and HVI infections of NPROs</td>
<td>119</td>
</tr>
<tr>
<td>4.4 Discussion</td>
<td>120</td>
</tr>
<tr>
<td>5 Conclusion and future prospects</td>
<td>126</td>
</tr>
<tr>
<td>5.1 Conclusion</td>
<td>126</td>
</tr>
<tr>
<td>5.2 Future prospects</td>
<td>127</td>
</tr>
<tr>
<td>5.2.1 Large scale investigations on systems-level host response to IAV infection</td>
<td>127</td>
</tr>
<tr>
<td>5.2.2 Using NPROs for modeling secondary bacterial infection in neutrophils</td>
<td>128</td>
</tr>
<tr>
<td>5.2.3 Obtaining NPROs with high neutrophil yields</td>
<td>129</td>
</tr>
<tr>
<td>5.2.4 Host-oriented drug discovery for influenza</td>
<td>129</td>
</tr>
<tr>
<td>5.3 Final remarks</td>
<td>130</td>
</tr>
<tr>
<td>References</td>
<td>131</td>
</tr>
<tr>
<td>Appendix A</td>
<td>170</td>
</tr>
<tr>
<td>Appendix B</td>
<td>175</td>
</tr>
</tbody>
</table>
Abstract

Severe influenza infections have been associated with dysregulated innate immunity that involves macrophages and neutrophils. While the contributions of macrophages to the dysregulation have been broadly investigated, the contributions of neutrophils remain unclear. Hence, in this thesis, we uncovered systems-level neutrophil response to highly virulent influenza infection by employing MPRO neutrophils and highly virulent, mouse adapted H3N2 influenza virus (HVI). Firstly, we showed that HVI induced hypercytokinemia and increased antiviral (interferon) response in the infected lungs. Moreover, increased apoptotic activity and under-expression of genes associated with metabolic and developmental processes mirrored severe pathological changes in HVI-infected lungs. Following pathway analysis, we highlighted the significant roles of the TREM1 signaling pathway in enhancing cytokine expression, and linked the hypercytokinemia to metabolic defect through the activation of LPS/IL1-mediated inhibition of retinoid X receptor (RXR) function pathway. With regards to infection of MPRO neutrophils (optimally containing 20%-30% mature neutrophils, as inspected with differential counting of giemsa-stained cells and flow cytometry based on neutrophil markers), influenza virus could induce apoptosis even though its infection was abortive. Finally, we revealed that HVI mainly activated a rapid induction of type I interferon-inducible genes in MPRO neutrophils, an event that potentially contributes to the dysregulated innate immunity observed in vivo.
Summary

Systems biology approaches have provided remarkable understanding on the host response to influenza A virus infection (IAV) over the last decade. This includes the identification of early and prolonged innate immune response in lungs infected with highly virulent IAVs. In this thesis, we have utilized genomic approach to reveal novel aspects of host response to highly and less virulent influenza H3N2 viruses (designated as HVI and LVI, respectively) in a murine model, in vivo (in lungs) and in vitro (in neutrophils).

In Chapter 2, it is shown that lungs infected with HVI, but not LVI, elicited hyperinduction of cytokine/chemokine and type I interferon (IFN)-inducible genes. Prominent over-expression of neutrophil chemoattractant CXCL1, macrophage chemoattractant CCL7, and antiviral type I IFN-inducible ISG15 were observed even at 12 h postinfection. Pathway analysis suggests that dysregulation of TREM1 signaling might play a role in causing the aberrant innate immune response. Furthermore, HVI infection culminated in the repression of genes associated with the metabolism and developmental processes at 96 h postinfection. The activation of LPS/IL-1 mediated inhibition of RXR function pathway has been proposed to cause the repression of metabolic genes. Finally, over-expression of apoptosis-related genes was more intense in HVI-infected lungs than in LVI-infected lungs.
In Chapter 3, we report the outcomes of examining the purity of the neutrophils in induced differentiated mouse promyelocytic MPRO cells (NPROs) that were used for modeling IAV infection in neutrophils. The purity of neutrophils in NPROs was determined by using giemsa and flow cytometry approaches. In addition, to determine the optimal purity, NPROs were collected within 6 days of differentiation. For giemsa approach, we have proposed and assessed an automated image analysis system to determine neutrophil purity from high content images of giemsa-stained NPROs. We demonstrated that the automated system gave relatively comparable estimates to manual estimates. For flow cytometry, the estimation was performed by using specific and non-specific neutrophil marker antibodies, i.e. anti-Ly-6G and anti Gr-1, respectively. Notably, flow cytometry based on anti-Ly-6G generally gave estimates of neutrophil purity that were lower than those given by giemsa and Gr-1-based flow cytometry approaches. On the day when neutrophil yield was optimum (i.e., day 5 after differentiation), the quantity of neutrophil-like cells in giemsa-stained cells was about 30%, while Ly-6G$^+$ cells was only about 20%. Interestingly, the population of Gr-1$^+$ cells on that day was about 32%, close to the percentage of neutrophils estimated from giemsa-stained NPROs.

Next, we present the outcomes of modeling IAV infection of neutrophils using NPROs as a neutrophil model in Chapter 4. We show that the characteristics of IAV infection in NPROs were similar to IAV infection in normal neutrophils, i.e. the infection was abortive (viral proteins were synthesized but not new progeny), and it augmented early and late apoptosis as indicated by annexin V and TUNEL assays. Furthermore, the analysis on transcriptomic data of NPRO responses to LVI and HVI infections revealed that IFN regulatory factor and IFN signaling pathways were the top canonical pathways significantly enriched by differentially expressed genes in
both infections, with activation of related genes in HVI as early as 3 h postinfection. Relatively consistent results were obtained by real-time RT-PCR of selected genes associated with type I IFN pathway. Early after HVI infection, comparatively enhanced expression of apoptosis-related genes was also elicited.

Finally, in Chapter 5, we conclude our thesis and present future prospects of our studies. In short, we have confirmed that lungs infected with a highly virulent H3N2 virus elicit early over-expression of cytokine/chemokine and type I IFN-inducible genes. Moreover, we have succeeded in employing NPROs for modeling IAV infection in neutrophils and showed that neutrophils may contribute to early activation of type I IFN response during severe influenza. Altogether, our studies have provided new directions in the influenza and neutrophil research.
List of tables

3.1 List of features extracted from nucleus and cytoplasm regions of a single cell. .................................................. 82

3.2 Selected features for classification and their ranking for each scale. Mean test error for classification at corresponding scale using the selected features are shown at the last row. ......................... 90

4.1 Changes in expression of genes associated with apoptotic activity in NPROs infected with IAVs. The gene expression changes for MPI, LVI and HVI infections at the time-points are expressed as log ratios. ................................................................. 120
List of figures

1.1 The structure of IAV virion. .................................................. 8

1.2 Type I IFN and type II IFN signaling pathways. ..................... 33

1.3 TREM1 signaling pathway in neutrophils (black), monocytes (white) or both (gray); hatched lines illustrate protein degradation. . . 35

1.4 A DNA microarray workflow. The gray boxes represent the main steps in the workflow, whereas the white boxes represent work associated with each of the steps. ................................. 40

1.5 A scheme for automated image analysis system. The gray boxes represent the main steps in automating the analysis of image data, whereas the white boxes represent a prior experiment that needs to be carried out to determine the set of features and classifier’s parameters for image object classification. ................................. 46

2.1 Histopathology of mouse lungs infected with LVI or HVI at 12, 48, and 96 h postinfection (p.i.). Sections of lung tissues were stained with H&E. At 12 h p.i., both LVI- and HVI-infected lungs exhibited normal morphology. At 48 h p.i., LVI-infected lungs revealed alveolitis with mild infiltration of inflammatory cells, whereas HVI-infected lungs displayed a more severe pathology with bronchopneumonia and emphysema. At 96 h p.i., LVI-infected lungs exhibited recovery, while HVI-infected lungs displayed severe bronchopneumonia, massive leukocyte infiltration, necrotizing bronchitis, hemorrhage and accumulation of proteinaceous materials. ................................................................. 56

2.2 Summary of transcriptional response of murine lungs infected with LVI and HVI. (A) Venn diagram showing the overlap between significant groups of genes as detected by two-way ANOVA analysis with Storey correction for multiple tests (q-value <0.05). (B) Heatmap of 1,195 genes that had a significant infection effect and an absolute log ratio >0.6 at any time-point of any treatment. (C) Venn diagram showing the summary of the overlap of differentially expressed genes in LVI (dark-gray circles) and HVI (light-gray circles) that had an absolute log ratio >0.6 at 12, 48 and
2.3 Overview of cytokine gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.

2.4 Differential regulation of TREM1 signaling in LVI and HVI. (A) Dynamic expression profiles of selected genes involved in the TREM1 signaling pathway in LVI and HVI. (B) A proposed gene sub-network implicated in the development of hypercytokinemia. Genes in the sub-network had a significant infection effect (q-value <0.05), and exhibited an absolute log ratio of >0.6 in at least one time-point after LVI or HVI infection. Direct interactions and indirect interactions are represented by solid lines and dashed lines, respectively. Gray nodes represent transiently under-expressed genes in LVI, while white nodes represent over-expressed genes at all time-points in both LVI and HVI.

2.5 Overview of type I IFN-related gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.

2.6 Overview of apoptotic gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.

2.7 A gene sub-network highlighting the relationships between under-expressed genes in HVI. Genes in the sub-network had a significant infection effect (q-value <0.05) and showed a log ratio <−0.6 at 96 h post-infection. Gray nodes represent genes associated with metabolism of small molecules, while white nodes represent genes associated with developmental processes.

3.1 Giemsa-stained NPROs. (A) The diversity of induced differentiated MPRO population observed on day 0, 3 and 5 after ATRA-induced differentiation. (B) Representatives of NPROs at each stage of neutrophil maturation: promyelocyte, myelocyte, metamyelocyte, band cell, and mature neutrophil.

3.2 The viabilities of NPROs over a six day period of maintenance with excess and limited supply of new medium. The viabilities were determined based on the trypan blue exclusion test. The viabilities are shown as means ± SE of four independent determinants. The volumes added during the maintenance are shown in the table above the horizontal axis (qV = q × initial volume for differentiation at day 0). The differences between the viabilities of NPROs maintained with excess and limited supply of new medium were tested using the two-tail t-test (* p-value <0.05).

3.3 (A) An example of cell segmentation results output by the iterative global image thresholding. For each iteration, the objects in a blue bounding box are the images of a single cell successfully detected by the iterative thresholding algorithm, whereas the objects in a red

xv
Bounding box are processed in the next iteration. (B) Fifty fields of giemsa-stained NPRO cells for testing the performance of iterative global image thresholding to isolate a single cell from cell clusters.

3.4 Samples of training image data for neutrophils (A) and non-neutrophils (B).

3.5 (A) Test error of $k$-NN for various $k$, the number of neighbors (all features were used for the classification). (B) Test error of $k$-NN ($k=10$) using the first $k$ features selected by mRMR criterion. (C) The recall and precision of each class given by $k$-NN ($k=10$) based on mRMR’s first five features. Error bars in (A), (B) and (C) were based on 100 experiments.

3.6 Comparison between automated (at scale 40×, 20× and 10×) and manual estimates for neutrophil purity in cytophinned giemsa-stained NPROs collected on day 0 to 6 after differentiation. NPROs were maintained with excess or limited medium (left and right panels, respectively). The estimate at each day of differentiation are shown as means ± SE of four independent biological replicates. For each magnification, the differences between the automated and manual estimates at each day of differentiation were tested using the paired two-tail t-test with Bonferroni correction (* adjusted p-value <0.05).

3.7 Anti-Ly-6G-based flow cytometry estimates (average of four replicated samples) for neutrophil purity in NPROs collected on day 0 to 6 after differentiation, in comparison with manual estimates for neutrophil purity in cytophinned giemsa-stained NPROs. NPROs were maintained with excess or limited medium (left and right panels, respectively). The estimate at each day of differentiation are shown as means ± SE of four independent biological replicates. For each magnification, the differences between the flow and manual estimates at each day of differentiation were tested using the unpaired two-tail t-test with Bonferroni correction (* adjusted p-value <0.05).

3.8 A representative of flow cytometry analysis of NPROs which were unstained, stained with PerCP-Cy5.5 anti-Ly-6G or FITC anti-Gr-1 only, and stained with both PerCP-Cy5.5 anti-Ly-6G and FITC anti-Gr-1.

4.1 Comparison between viral replication kinetics of LVI, HVI and MPI in MDCK cells. The data represent mean ± SE of three independent measurements of log virus titers. The difference between log virus titers for each pair of viruses at each time point was tested using two-tail unpaired t-test (* p-value <0.05).

4.2 A representative of flow cytometry analysis of NPROs infected either with unlabelled or DiD-labelled MPI, and unstained or stained.
with FITC Ly-6G.

4.3 Confocal microscopy on NPROs examined 1 h after incubation with DiD-labelled MPI. Two representative examples are depicted: the left panels show nuclei stained with DAPI (blue), middle panels exhibit the cytoplasm of infected cells containing DiD label (red), while right panels are the merged images.

4.4 (A) Comparison of virus titers of supernatants from infected MDCK cells and NPROs at 6, 12 and 24 h postinfection with MPI at MOI of 0.1. The data represent means ± SE of three independent measurements of log virus titers. The means of log virus titers vs detection limit was tested using one-tail unpaired t-test (* indicates p-value < 0.05). (B) Detection of viral protein synthesis in infected NPROs via immunofluorescence staining (IF) using polyclonal antibody against H3N2 virus. Viral proteins could be observed in the cytoplasm and/or nucleus of neutrophils at 9 h after incubation with MPI. (C) Kinetics of viral protein synthesis based on the percentage of cells positive for viral staining. The percentages are shown as means ± SE of three independent IF stained samples of NPROs that were initially infected with MPI at MOI of 1.

4.5 Induction of enhanced apoptosis in NPROs infected by IAV. About $1.5 \times 10^6$ cells were initially infected with MPI at MOI of 1, and compared with uninfected cells. (A) Flow cytometry based on A5 and PI staining indicated that the proportions of infected cells undergoing early apoptosis (A5-positive and PI-negative) were significantly higher than uninfected cells at multiple time-points. (B) TUNEL assay also revealed augmented late apoptosis of infected NPROs at 12 h postinfection, as illustrated by representative images, and quantification of the proportion of TUNEL-positive cells. The proportions are shown as means ± SE of five independent determinants for (A) and four determinants for (B). The differences between the proportions of apoptotic cells in infected and uninfected samples were tested using the one-tail paired t-test (* p-value < 0.05).

4.6 Analyses of gene microarray data of IAV infection of NPROs. (A) Venn diagram showing the overlap between significant groups of genes as analyzed by two-way ANOVA with Benjamini-Hochberg correction for multiple tests (adjusted p-value < 0.05). (B) Heatmap of 129 genes with a significant infection effect, and an absolute log ratio of >0.6 at any time-point of infection with LVI, HVI and MPI.

4.7 Overview of IFN-related genes in response to infection with LVI, HVI and MPI at 3 and 9 h postinfection (p.i.).

4.8 A subnetwork of canonical pathways associated with IFN-β regulation and signaling. Genes that were over-expressed in LVI, HVI and MPI infections by microarray data analysis are shaded
grey; genes over-expressed in LVI and MPI but not in HVI shaded black. ................................................................. 117

4.9 Validation of microarray data for several IFN-related genes by real-time RT-PCR. Log fold changes (relative to mock control) shown in the graphs reflect the average of two biological samples (each biological sample was an average of two technical measurements).  118
List of genes/proteins

Note that:
1. Hyphen mark (–) indicates a family of genes or a group of genes that form a complex. The hyphen mark can be substituted by nothing, or numbers and/or alphabets that indicate the member of the gene family or the complex.
2. Gene name is listed on the left and its description on the right.

Viral genes/proteins

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA</td>
<td>hemagglutinin</td>
</tr>
<tr>
<td>M–</td>
<td>2matrix protein (i.e. M1 and M2)</td>
</tr>
<tr>
<td>NA</td>
<td>neuraminidase</td>
</tr>
<tr>
<td>NP</td>
<td>nucleoprotein</td>
</tr>
<tr>
<td>NS–</td>
<td>non-structural protein (i.e. NS1 and NS2)</td>
</tr>
<tr>
<td>PA</td>
<td>polymerase acidic</td>
</tr>
<tr>
<td>PB–</td>
<td>polymerase basic protein (i.e. PB1 and PB2)</td>
</tr>
</tbody>
</table>

Host genes/proteins

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALDH3A1</td>
<td>aldehyde dehydrogenase 3 family, member A1</td>
</tr>
<tr>
<td>B2M</td>
<td>beta-2-microglobulin</td>
</tr>
<tr>
<td>BCL10</td>
<td>B-cell CLL/lymphoma 10</td>
</tr>
<tr>
<td>BCL-XL</td>
<td>B-cell lymphoma-extra large</td>
</tr>
<tr>
<td>C–</td>
<td>complement factor (e.g. C5a and C7)</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>CARD9</td>
<td>caspase-recruitment domain 9</td>
</tr>
<tr>
<td>CASP-</td>
<td>caspase (e.g. CASP1, CASP4 and CASP8)</td>
</tr>
<tr>
<td>C-CBL</td>
<td>casitas B-lineage lymphoma</td>
</tr>
<tr>
<td>CCL-</td>
<td>CC chemokine (e.g. CCL2 (MCP1), CCL3 (MIP1-α), CCL4 (MIP1-β), CCL5 (RANTES), CCL7 (MCP3), CCL12 and CCL21)</td>
</tr>
<tr>
<td>CCR-</td>
<td>CC chemokine receptor (e.g. CCR1, CCR2, CCR3, CCR10)</td>
</tr>
<tr>
<td>CD-</td>
<td>cluster of differentiation (e.g. CD4, CD14, CD32 (FCGR2), CD40, CD43, CD45, CD54, CD79A, CD83 and CD86)</td>
</tr>
<tr>
<td>CORO1A</td>
<td>coronin, actin binding protein, 1A</td>
</tr>
<tr>
<td>CSF-</td>
<td>colony stimulating factor (i.e. CSF1 (M-CSF), CSF2 (GM-CSF), and CSF3 (G-CSF))</td>
</tr>
<tr>
<td>CSRNP1</td>
<td>cysteine-serine-rich nuclear protein 1</td>
</tr>
<tr>
<td>CXCL-</td>
<td>CXC chemokine (e.g. CXCL1 (Gro-α; KC) and CXCL10 (IP10))</td>
</tr>
<tr>
<td>CXCR-</td>
<td>CXC chemokine receptor (e.g. CXCR1 (IL8RA), CXCR2 (IL8RB), CXCR3 and CXCR4)</td>
</tr>
<tr>
<td>CYP</td>
<td>cytochrome p450</td>
</tr>
<tr>
<td>DAXX</td>
<td>death domain-associated protein</td>
</tr>
<tr>
<td>DBP</td>
<td>albumin D box-binding protein</td>
</tr>
<tr>
<td>DDX58</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 58</td>
</tr>
<tr>
<td>EAR2</td>
<td>ERBA-related gene 2</td>
</tr>
<tr>
<td>EIF2AK2</td>
<td>eukaryotic translation initiation factor 2-alpha kinase 2</td>
</tr>
<tr>
<td>ELA2</td>
<td>elastase 2</td>
</tr>
<tr>
<td>ERK-</td>
<td>extracellular regulated kinase (e.g. ERK1 and ERK2)</td>
</tr>
<tr>
<td>FADD</td>
<td>FAS-associated via death domain</td>
</tr>
<tr>
<td>FAS</td>
<td>TNF receptor superfamily, member 6</td>
</tr>
</tbody>
</table>

xx
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCAR</td>
<td>Fc fragment of Ig-α receptor</td>
</tr>
<tr>
<td>FCER1G</td>
<td>Fc fragment of Ig-ε receptor (e.g. FCER1G)</td>
</tr>
<tr>
<td>FCGR–</td>
<td>Fc fragment of Ig-γ receptor (e.g. FCGR2A, FCGR2B, FCGR2C, FCGR3A and FCGR3B)</td>
</tr>
<tr>
<td>FGF1</td>
<td>fibroblast growth factor 1</td>
</tr>
<tr>
<td>GADD45G</td>
<td>growth arrest and DNA-damage-inducible, gamma</td>
</tr>
<tr>
<td>GAPDH</td>
<td>glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GBP–</td>
<td>guanylate binding protein (e.g. GBP1, GBP2, GBP3 and GBP4)</td>
</tr>
<tr>
<td>G-CSF</td>
<td>granulocyte colony stimulating factor</td>
</tr>
<tr>
<td>GCH1</td>
<td>GTP cyclohydrolase 1</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>granulocyte/macrophage colony stimulating factor</td>
</tr>
<tr>
<td>GPX3</td>
<td>glutathione peroxidise 3</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione s-transferase</td>
</tr>
<tr>
<td>H2-D1</td>
<td>histocompatibility 2, D region locus 1</td>
</tr>
<tr>
<td>H2-L</td>
<td>histocompatibility 2, D region locus L</td>
</tr>
<tr>
<td>HMOX1</td>
<td>heme oxygenase (decycling) 1</td>
</tr>
<tr>
<td>IFI–</td>
<td>IFN-inducible protein (e.g. IFI16, IFI35 and IFI204)</td>
</tr>
<tr>
<td>IFIH1</td>
<td>IFN-induced with helicase C domain 1</td>
</tr>
<tr>
<td>IFIT–</td>
<td>IFN-induced protein with tetratricopeptide (e.g. IFIT1, IFIT2 and IFIT3)</td>
</tr>
<tr>
<td>IFN</td>
<td>interferon</td>
</tr>
<tr>
<td>IFN-α</td>
<td>interferon alpha (e.g. IFN-α1, IFN-α4, IFN-α5 and IFN-α8)</td>
</tr>
<tr>
<td>IFNAR–</td>
<td>IFN-α/β receptor (i.e. IFNAR1 and IFNAR2)</td>
</tr>
<tr>
<td>IFN-β</td>
<td>interferon beta (i.e. IFNB1)</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>interferon gamma</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>IFNGR—</td>
<td>IFN-γ receptor (i.e. IFNGR1 and IFNGR2)</td>
</tr>
<tr>
<td>Ig-α</td>
<td>immunoglobulin alpha</td>
</tr>
<tr>
<td>Ig-ε</td>
<td>immunoglobulin epsilon</td>
</tr>
<tr>
<td>Ig-γ</td>
<td>immunoglobulin gamma</td>
</tr>
<tr>
<td>IL—</td>
<td>interleukin (e.g. IL1B, IL3, IL6, IL8 and IL12)</td>
</tr>
<tr>
<td>IL–R</td>
<td>interleukin receptor (e.g. IL1R, IL6R, IL10R, IL17R and IL2RG)</td>
</tr>
<tr>
<td>IRAK</td>
<td>IL1R-associated kinase</td>
</tr>
<tr>
<td>IRF–</td>
<td>IFN regulatory factory (e.g. IRF1, IRF3, IRF7 and IRF9)</td>
</tr>
<tr>
<td>ISG–</td>
<td>IFN-stimulated gene (e.g. ISG15 and ISG20)</td>
</tr>
<tr>
<td>ISGF3</td>
<td>IFN-stimulated gene factor 3</td>
</tr>
<tr>
<td>JAK–</td>
<td>Janus kinase (e.g. JAK1 and JAK2)</td>
</tr>
<tr>
<td>JNK</td>
<td>Jun N-terminal kinase</td>
</tr>
<tr>
<td>LIR—</td>
<td>leukocyte immunoglobulin-like receptor (e.g. LIR1 and LIR9)</td>
</tr>
<tr>
<td>LYN</td>
<td>LCK/YES-related novel kinase</td>
</tr>
<tr>
<td>MAC-1</td>
<td>macrophage antigen 1</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>M-CSF</td>
<td>macrophage colony stimulating factor</td>
</tr>
<tr>
<td>MMP9</td>
<td>matrix metalloproteinase 9</td>
</tr>
<tr>
<td>MX–</td>
<td>myxovirus (influenza virus) resistance protein (i.e. MX1 and MX2)</td>
</tr>
<tr>
<td>NAD</td>
<td>aldehyde dehydrogenase</td>
</tr>
<tr>
<td>NCR–</td>
<td>NK cell activating receptor (e.g. NCR1 (NKp46) and NCR2 (NKp44))</td>
</tr>
<tr>
<td>NF-κB–</td>
<td>nuclear factor-kappa B</td>
</tr>
<tr>
<td>NFKBIA</td>
<td>nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NLR(s)</td>
<td>NAIP, CIITA, HET-E, TP-I-leucine-rich repeat receptor(s) (NATCH-LRR(s))</td>
</tr>
<tr>
<td>NLRP3</td>
<td>NLR family, pyrin domain containing 3</td>
</tr>
<tr>
<td>NOD–</td>
<td>nucleotide-binding oligomerization domain (e.g. NOD1 and NOD2)</td>
</tr>
<tr>
<td>NOTCH–</td>
<td>neurogenic locus notch homolog protein (e.g. NOTCH3 and NOTCH4)</td>
</tr>
<tr>
<td>NR1D1</td>
<td>nuclear receptor subfamily 1 group D member 1</td>
</tr>
<tr>
<td>NTAL</td>
<td>non-T cell activation linker</td>
</tr>
<tr>
<td>OAS–</td>
<td>2′-5′ oligoadenylate synthetase (e.g. OAS1 and OAS3)</td>
</tr>
<tr>
<td>OASL1</td>
<td>2′-5′ oligoadenylate synthetase-like 1</td>
</tr>
<tr>
<td>OASL2</td>
<td>2′-5′ oligoadenylate synthetase-like 2</td>
</tr>
<tr>
<td>ORM2</td>
<td>orosomucoid 2</td>
</tr>
<tr>
<td>PARP–</td>
<td>poly (ADP-ribose) polymerase family, member (e.g. PARP9, PARP10 and PARP14)</td>
</tr>
<tr>
<td>PI3K</td>
<td>phosphoinositide-3-kinase</td>
</tr>
<tr>
<td>PKB</td>
<td>protein kinase B</td>
</tr>
<tr>
<td>PKC-δ</td>
<td>protein kinase C-δ</td>
</tr>
<tr>
<td>PLC-γ</td>
<td>phospholipase C-γ</td>
</tr>
<tr>
<td>PRDX5</td>
<td>peroxiredoxin 5</td>
</tr>
<tr>
<td>RAG2</td>
<td>recombination activating gene 2</td>
</tr>
<tr>
<td>RAC2</td>
<td>ras-related C3 botulinum toxin substrate 2</td>
</tr>
<tr>
<td>RAR</td>
<td>retinoic acid receptor</td>
</tr>
<tr>
<td>RELA</td>
<td>v-rel reticuloendotheliosis viral oncogene homolog A</td>
</tr>
<tr>
<td>RETNLG</td>
<td>resistin like gamma</td>
</tr>
<tr>
<td>RSAD2</td>
<td>radical S-adenosyl methionine domain containing 2; viperin</td>
</tr>
</tbody>
</table>
RXR                        retinoid X receptor
S100–                      S100 calcium binding protein (S100A8, S100A9)
SAA3                       serum amyloid A3
SELP                       selectin P
SHP1                       SH2-containing tyrosine phosphatase-1
SOCS–                      suppressor of cytokine signaling (e.g. SOCS1 and SOCS3)
SOD2                       superoxide dismutase 2
SOS1                       guanine nucleotide exchange factor son of sevenless 1
SP11                       serine protease 11
STAB1                      stabilin 1
STAT–                      signal transducer and activator of transcription (e.g. STAT1, STAT2
                           and STAT3)
TGFBR2                     transforming growth factor, beta receptor 2
THY1                       Thy-1 T-cell antigen
TLR–                       toll-like receptor (e.g. TLR2, TLR7 and TLR8)
TNF-α                      tumor necrosis factor alpha
TNFAIP3                    TNF alpha-induced protein 3
TNF-β                      tumor necrosis factor beta
TNFR–                      tumor necrosis factor receptor (i.e. TNFR1 and TNFR2)
TRAIL                      TNF-related apoptosis-inducing ligand
TRAILR                     TRAIL receptor
TREM1                      triggering receptor expressed by myeloid cells (e.g. TREM1)
TRIM25                     tripartite motif containing 25
TYK2                       tyrosine kinase 2
TYROBP                     TYRO protein tyrosine kinase binding protein
<table>
<thead>
<tr>
<th>VEGFA</th>
<th>vascular endothelial growth factor A</th>
</tr>
</thead>
<tbody>
<tr>
<td>XAF1</td>
<td>XIAP associated factor 1</td>
</tr>
</tbody>
</table>
Abbreviations

A5         annexin V
AM         alveolar macrophage
AML        acute myeloblastic leukemia
ANOVA      analysis of variance
ARDS       acute respiratory distress syndrome
ATCC       American Type Culture Collection
ATRA       all-trans retinoic acid
Ca^{2+}    calcium cation
CA04       2009 pandemic influenza A/California/04/2009 (H1N1) virus
cDNA       complementary DNA
CO_{2}      carbondioxide
cRNA       complementary RNA
DAPI       4′,6-diamidino-2-phenylindole
DiD        1,1′-dioctadecyl-3,3,3′,3′-tetramethyl-indodicarbocyanine
DNA        deoxyribonucleic acid
dsDNA      double stranded DNA
dsRNA      double-stranded RNA
EDTA       ethylenediaminetetraacetic acid
EMEM       Eagle’s Minimum Essential Medium
EML cell lines with erythroid, myeloid, and lymphoid potential
EPRO EML promyelocyte cell lines
FBS fetal bovine serum
FDR false discovery rate
FJ02 influenza A/Fujian/411/2002 (H3N2) virus
GAS IFN-\(\gamma\) activated-site
GO gene ontology
HCS high content screening
HK98 influenza A/Hong Kong/54/1998 (H1N1) virus
HKx31 influenza A/HKx31 (H3N2) virus
HVI highly virulent, mouse adapted influenza H3N2 virus
IAV influenza A virus
IMDM Iscove’s Modified Dulbecco’s Medium
IPA Ingenuity Pathway Analysis
ISRE IFN-stimulated response element
K173 influenza A/Kawasaki/173/2001 (H1N1) virus
\(k\)-NN \(k\)-nearest neighbors
LPS lipopolysaccharide
LVI less virulent, mouse non-adapted influenza H3N2 virus
mAb monoclonal antibody
MDCK(s) Madine-Darby Canine Kidney cell line(s)
MOI multiplicity of infection
MPI HVI propagated in MDCK cells
MPRO(s) mouse promyelocyte cell line(s)
mRMR minimal-redundancy-maximal-relevance
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NADPH</td>
<td>nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NC99</td>
<td>influenza A/New Caledonia/20/1999 (H1N1) virus</td>
</tr>
<tr>
<td>NET(s)</td>
<td>neutrophil extracellular trap(s)</td>
</tr>
<tr>
<td>NK</td>
<td>natural killer</td>
</tr>
<tr>
<td>NL602</td>
<td>influenza A/Netherlands/602/2009 (H1N1) virus</td>
</tr>
<tr>
<td>NPRO(s)</td>
<td>neutrophils from induced differentiated MPRO cell(s)</td>
</tr>
<tr>
<td>p.i.</td>
<td>postinfection</td>
</tr>
<tr>
<td>PAMP(s)</td>
<td>pathogen-associated molecular pattern(s)</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffer saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming units</td>
</tr>
<tr>
<td>PI</td>
<td>propidium iodide</td>
</tr>
<tr>
<td>PN99</td>
<td>influenza A/Panama/2007/1999 (H3N2) virus</td>
</tr>
<tr>
<td>PR8</td>
<td>influenza A/Puerto Rico/8/1934 (H1N1) virus</td>
</tr>
<tr>
<td>PRR(s)</td>
<td>pattern recognition receptor(s)</td>
</tr>
<tr>
<td>r1918</td>
<td>reconstructed influenza 1918 pandemic (H1N1) virus</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase - polymerase chain reaction</td>
</tr>
<tr>
<td>ssRNA</td>
<td>single-stranded RNA</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>TCID&lt;sub&gt;50&lt;/sub&gt;</td>
<td>tissue culture infectious dose 50</td>
</tr>
</tbody>
</table>
Th1  T helper 1
TLR  toll-like receptor
TNF  tumor necrosis factor
TUNEL  terminal deoxynucleotidyl transferase dUTP nick end labelling
TX91 influenza A/Texas/36/1991 (H1N1) virus
VN1203 influenza A/Vietnam/1203/2004 (H5N1) virus
VN3212 influenza A/Vietnam/3212/2004 (H5N1) virus
vRNP viral ribonucleoprotein
VST  variance-stabilizing transform
WSN33 influenza A/WSN/1933 (H1N1) virus
Chapter 1

Introduction

The advent of high-throughput technologies, such as DNA microarrays and high-content imaging, has revolutionized our ways of advancing both biomedical and biological research. These technologies have driven the emergence of systems biology that models complex interactions within biological systems. In the area of infectious diseases, systems biology approaches provide tools to dissect interactions between pathogens and their hosts. The use of these tools have improved our understanding of the molecular aspects of pathogenesis following an infection, enabled the identification of more accurate disease signatures, and driven the development of effective therapeutic treatments. In this thesis, we employed systems biology approaches to explore systems-level analyses of host response to influenza A virus (IAV) infection.

For the systems-level studies, we employed a mouse model. Mouse models have provided conceptual frameworks for various human diseases, including influenza. Mice serve as an attractive model for many reasons, e.g. they are easy and inexpensive to maintain in the laboratory compared to other animals, have a shorter life span, and have provided numerous knock-out models for many genes. Nevertheless, it has recently been shown that genomic responses in mouse vs human during various inflammatory diseases (trauma, burns and endotoxemia) are poorly
correlated [209]. Moreover, in relation to IAV, Go et al [208] has shown that genomic response to 2009 H1N1 pandemic virus in mouse model is different from that in macaque and swine models; hence, it is also likely different from that in a human model. These findings challenge the modern biomedical approaches for studying and curing diseases using mouse models, and thus, the extension of the outcomes in our studies to human models must be done with precautions.

In the next sections of this chapter, we firstly discuss the motivation, objectives, and contributions of our works on systems-level host response to IAV infection. Then, we provide background and related work that underlie our research. Finally, the organization of the thesis is outlined at the end of the chapter.

1.1 Motivations and objectives

Genomic approaches based on DNA microarrays have been very popular for investigating systems-level host response to IAV infections. In vivo, genomic studies have revealed that lungs infected with highly virulent IAVs, in comparison with the lungs infected with less virulent ones, are characterized by early and prolonged over-expression of genes associated with innate immunity [1-6, 195]. A very recent study, which integrates transcriptomes of lungs and sorted cells (including nonhematopoietic cells, neutrophils, lymphocytes, monocytes and alveolar macrophages) following non-lethal to lethal IAV infections, has pointed out the potential role of a feedforward inflammatory circuit involving neutrophils in causing the dysregulated innate immunity [204]. In vitro, genomic studies have mainly been carried out on epithelial cells and monocyte-derived/alveolar macrophages [7-12, 197], two types of cells that have been appreciated as the major target of IAV infection. While most of the results showing that both cells could regulate expression of various genes – including
interferon and innate immune related genes, one report interestingly showed that alveolar macrophages infected with highly virulent IAV only exerted insignificant expression changes [10]. Surprisingly, no genomic in vitro studies have been carried out in neutrophils although together with macrophages, neutrophils have been shown to excessively infiltrate the lungs immediately after infection with highly virulent IAV [14]. This fact has become a key motivation for studying genome-wide neutrophil response to IAV infection proposed in this thesis, and the recent results in [204] has further justified its importance.

In particular, we were motivated to develop an in vitro model of IAV infection in neutrophils and to uncover transcriptomic responses of neutrophils to less and highly virulent IAV infections, and to link the in vitro results to responses observed in vivo. For this, we used less and highly virulent IAVs of H3N2 subtype – designated as LVI and HVI, respectively – that have been described in [13]. Since HVI was derived from LVI through serial lung-to-lung passaging in mice, our infection model may minimize the variability due to strain effects. Overall, the objectives and motivations of our studies are as follows:

(i) Firstly, we aim to compare the early transcriptomic responses of lungs infected with LVI and HVI. The main focus is to identify pathways that can be associated with dysregulation of innate immune response observed during highly virulent IAV infection, especially in relation to cytokine storm or hypercytokinemia. In addition, the identification of pathways or gene sets that enhance lung damage is also of our interest.

(ii) Secondly, we aim to compare the transcriptomic responses of neutrophils infected with LVI and HVI. We are specifically interested in uncovering how neutrophils might contribute to the severity of HVI infection. For this study, we
explore the use of neutrophils from induced differentiated mouse promyelocytic MPRO cells (designated as NPROs, which stands for Neutrophils from induced differentiated MPRO cells) that have been shown to behave like normal neutrophils [15]. This is because the traditional approach that isolates neutrophils from murine blood is expensive, posing technical and ethical issues, making it unsuitable for large scale experiments. Objectives in part (iii) and (iv) are set to assess and validate the use of NPROs for our model.

(iii) Next, we aim to quantify the purity of neutrophils in NPROs. Thus far, only cells undergoing a granulocytic differentiation have been quantified [16]. Moreover, some NPRO associated experiments were done without knowing the purity of neutrophils [15, 17]. Nevertheless, we consider that neutrophil quantification is important to show the reliability of NPROs as a neutrophil source. For the quantification, giemsa staining and flow cytometry approaches will be employed and compared.

(iv) Finally, we aim to verify that IAV infection in NPROs has similar characteristics to IAV infection in normal neutrophils. In particular, we would like to show that the infection is abortive (viral proteins are synthesized but new progeny is not formed) [18] and induces apoptosis [19, 20]. To guide the selection of time-points for microarray experiments, time series data showing the kinetics of viral protein synthesis and apoptosis will be collected within 24 h after infection.

1.2 Contributions of the thesis

Since the early 20th century, IAV has continuously posed a pandemic threat. The ability of IAV to develop a resistance to antiviral drugs and no vaccine effectively
giving protection from all strains of viruses are the major problems to prevent the next
influenza outbreaks or pandemics. Hence, alternative treatments and preventive cares
for influenza are desired. One promising alternative is a pharmacological approach
that exploits host antiviral response or targets cellular components involved in viral
life cycle (e.g. surface receptors and protein kinases). In the light of the above, our
studies provide further understanding on systems-level host response to IAV that can
be useful to specify host drug targets. Specifically, the contributions of our research in
this thesis are as follows:

(i) Identification of differentially regulated pathways in lungs infected with LVI
and HVI. In general, our results have provided more evidence that the
characteristics of highly virulent IAV infections causes early and strong
induction of genes relevant to innate immune response, including type I
interferon (IFN)-related and cytokine/chemokine genes. One novel outcome of
our transcriptomic model is the identification of dysregulated TREM1 signaling,
which potentially contributes to the development of hypercytokinemia in lungs
infected with HVI. In addition, we have linked the hypercytokinemia to
progression of lung damage, represented by under-expressed pathways
associated with metabolic and developmental processes, via a mechanism
referred to as LPS/IL1-mediated inhibition of RXR function.

(ii) Identification of differentially regulated pathways in neutrophils infected with
LVI and HVI. The most novel result in our studies is that for the first time, we
uncover the transcriptomic responses of neutrophils to IAV infections. In
particular, we have demonstrated that neutrophils infected with IAVs are able to
activate type I IFN and apoptotic responses, and these responses are earlier in
HVI- than in LVI-infected neutrophils.
(iii) **Encouraging the use of NPROs for modeling the interaction between neutrophils and pathogens.** Here, we have demonstrated that NPROs can be reliably used as a neutrophil source for investigating neutrophil response to IAV infection. Specifically, we have shown that a significant proportion of relatively pure neutrophils can be obtained from NPROs, as indicated by neutrophil-like cell count in giemsa-stained NPROs and flow cytometric measurements of cells that are positive for two common neutrophil surface markers, i.e. Ly-6G and Gr-1. To some extent, we have also demonstrated that IAV infection in NPROs is a valid model for IAV infection in neutrophils. Moreover, transcriptomic response elicited by infected NPROs is parallel to recent outcomes of neutrophil studies. Thus, the success of our infection model may encourage further use of NPROs for large scale neutrophil studies related to infectious diseases or other areas, which are often limited by the difficulty in providing a sufficient number of neutrophils isolated from bone marrow or blood.

(iv) **Application of high content image analysis approach for estimating the purity of neutrophils in giemsa stained NPROs.** Examining the quantity of neutrophils is a frequent task when exploiting NPROs as a neutrophil source. For this task, two approaches are available: differential counting of giemsa-stained NPROs and flow cytometry based on neutrophil marker antibodies. The first approach is more attractive since it is cheaper in term of facility and reagent cost, and sample can be prepared quickly. However, examining giemsa-stained samples is laborious and only about 100-200 cells are reasonably counted if it is done manually; whereas flow cytometry examines thousands of cells in a few minutes. Hence, we proposed an automated image analysis system for quantifying neutrophils from high content images of giemsa-stained NPROs.
Chapter 1

The proposed automated system tends to give estimates for neutrophil proportion in NPROs that are slightly higher than those given by a manual approach; nevertheless, those estimates are highly correlated. An evaluation of the automated system is also carried out by comparing its estimates to the percent of cells positive for specific neutrophil marker (i.e. Ly-6G). As a result, even though the automated system often gives higher estimates, the trend of its estimates over NPROs collected within six days of differentiation is relatively similar to that given by Ly-6G-based estimates. Overall, the automated system may complement and potentially substitute the manual approach and it can be useful for future quantification of neutrophils, especially for large scale experiments that try to optimize neutrophil yield in NPROs.

1.3 Background and recent work

1.3.1 Influenza A virus (IAV)

IAV is a highly contagious virus that causes an acute respiratory disease known as influenza or "the flu". IAV has a broad range of reservoirs, including humans, birds, pigs, horses and sea mammals. Here, we briefly highlight the structure of IAV virion, IAV subtypes associated with pandemic and seasonal influenza, and the clinical picture of influenza disease. In addition, various techniques to generate IAV virus for experimental studies are also presented.

IAV virion and evolution

IAV is a negative strand RNA virus that belongs to the Orthomyxoviridae family. The structure of the IAV virion is illustrated in FIG. 1.1 [21]. Its genome consists of eight
viral RNA segments that encode surface glycoproteins hemagglutinin (HA) and neuraminidase (NA), nucleoproteins (NP), matrix proteins (M1 and M2), non-structural proteins (NS1 and NS2), and RNA polymerase components (PA, PB1 and PB2).

![The structure of IAV virion.](image)

**FIG 1.1.** The structure of IAV virion.

The HA, NA and M2 proteins are respectively responsible for the virion entry, virion release, and uncoating of virion contents in host cells. These three proteins are embedded in the viral envelope derived from the lipid bilayer of the host plasma membrane. Beneath the envelope, the structural M1 proteins create a second layer encasing the viral genome. Each RNA segment of IAV forms a viral ribonucleoprotein (vRNP) complex with NP proteins and the three RNA polymerase components. The NP proteins encapsidate the vRNP structure, whereas the RNA polymerase is responsible for replication and transcription of the virus in host cells.

Due to its genomic structure, IAV can evolve in two different ways, i.e. via antigenic drift and antigenic shift. Antigenic drift involves gradual, minor point mutations that occur due to the lack of a proofreading mechanism during viral
genome replication. The main factor that stimulates this mutation is the recognition of IAV by antibodies or drugs. By undergoing antigenic drift, a new strain of IAV that is not recognized by antibodies or drugs to earlier IAV can be produced. On the other hand, antigenic shift is a reassortment of RNA segments from two or more viral subtypes. Antigenic shift often occurs when animal IAV transmitted to human and mixed with human IAV, and it may result in a new human IAV subtype.

**Pandemic and seasonal IAVs**

IAV subtypes are determined by the antigenicity of HA and NA. Hitherto, sixteen HA (H1-H16) and nine NA (N1-N9) have been identified. Among IAV subtypes, the ones that have caused influenza pandemics include H1N1, H2N2, and H3N2. The most devastating influenza pandemic was the “Spanish flu H1N1” in 1918, which claimed more than 50 million lives [22]. The global death toll of the following pandemics decreased, i.e. about 2 millions during “Asian flu H2N2” in 1957, 1 million during “Hong Kong flu H3N2” in 1968, and 18,000 during “Swine flu H1N1” in 2009.

Most pandemic viruses emerged from a combination of antigenic drift and antigenic shift. The “Asian flu” was caused by an H2N2 reassortant that contains the circulating H1N1 segments and avian HA, NA and PB1 [23, 24]; the “Hong Kong flu” by an H3N2 subtype that contained RNA segments from “Asian flu” subtypes and avian HA and PB1 [24, 25]; and the “Swine flu” by a quadruple H1N1 reassortant of North American avian virus origin (contributed the PB2 and PA genes), circulating human H3N2 (PB1), classical swine virus origin (HA, NP and NS), and Eurasian avian-like swine virus origin (NA and M) [26-28]. Strikingly, however, there is increasing evidence that the 1918 pandemic virus emerged from antigenic drift that followed a direct transmission of avian H1N1 to human [29, 30]. With regard to the
next influenza pandemic, the greatest fear is the direct transmission of avian H5N1 to humans, which has already caused outbreaks in 1997 and 2003–2004 [21, 31, 32].

In addition to the death tolls during influenza pandemics and outbreaks, IAVs are also responsible for about half a million deaths during seasonal influenza epidemics every year [33]. The most predominant subtypes of IAVs causing seasonal influenza in recent decades are the H1N1 and H3N2. The circulation of seasonal H1N1 in humans has in fact occurred after the 1918 H1N1 pandemic. However, the H1N1 virus disappeared following 1957 H2N2 pandemic but resurfaced in 1977 and has since co-circulated with the seasonal H3N2. The H1N1 evolution has been associated with multiple intra-subtype reassortment events [34], while the seasonal H3N2 evolves mainly through a rapid antigenic drift [35]. The H3N2 subtype is more dominant than H1N1 in causing seasonal influenza that peaks during winter time in temperate regions, but H1N1 have also temporarily predominated during milder epidemic seasons [36].

**Clinical picture of influenza disease**

The flu symptoms include fever, cough, sore throat, runny or stuffy nose, muscle or body aches, headaches, fatigue, and even vomiting and diarrhoea. The severity of influenza is determined by the virulence of the virus and host factors (pregnancy, morbid obesity, underlying chronic illness such as pulmonary and cardiovascular disease and diabetes mellitus). In a mild case, IAV infection mainly attacks the upper respiratory tract (the nose, throat, trachea and bronchi) and the disease lasts for about a week. In a severe case, patients develop pneumonia that can lead to acute respiratory distress syndrome (ARDS) – a severe form of acute lung injury that is characterized by acute onset, bilateral pulmonary infiltrates, and hypoxemia resistant to oxygen
therapy [37]. Respiratory failures caused by ARDS have been the major cause of deaths associated with severe influenza, where infected lungs display massive pulmonary oedema, haemorrhage, infiltrating leukocytes and lymphocytes, and diffuse alveolar damage with hyaline membrane formation [38]. The spectrums of pathologic changes observed in fatal cases due to pandemic and seasonal IAVs interestingly do not show significant differences [39].

**Generation of IAVs for experimental studies**

For experimental studies, IAVs are usually propagated in pathogen-free embryonated eggs or in continuous cell lines, e.g. MDCK and Vero cells. Cell lines provide a more attractive alternative since they are easily obtained and allow the conservation of HA sequence [40]. Nevertheless, virus culture in eggs or cell lines may attenuate the virulence of the virus [41]. Hence, results from experimental studies using such viruses need to be interpreted with some caution.

Another approach to generate IAVs is by serial lung-to-lung passages, where for each passage, lungs from infected animals are homogenized and the homogenate is used as virus stocks or for subsequent passage. This technique is useful for modeling IAV infection in animals (including mice) that are not naturally infected with IAV virus, as well as making an IAV strain from human or other animals more adaptive or virulent in a new host used to model the infection. Several studies have been carried out to reveal the basis of IAV virulence during its adaptation in mice: in [42], the HA of 2009 pandemic H1N1 acquired an enhanced binding to \( \alpha_{2,3} \) and simultaneously reduced binding to \( \alpha_{2,6} \) sialic acid-linked receptors; in association with the viruses used in our studies, non-conservative mutations in HA (Gly218Glu) and NS1 (Asp125Gly) was associated with the increased virulence of HVI compared
to LVI [13]; moreover, amino acid substitutions in other viral genes have also been associated with the increased virulence, e.g. PB2 of chicken and duck H5N1 isolates [43, 44].

Finally, the reverse-genetics techniques have become a very powerful technique to generate IAV with a specific genome from cloned cDNAs, where a recombinant of IAV from different strains or subtypes can be generated and specific point mutations can be introduced. In its early development, the reverse genetics for IAV was done by using a helper influenza virus [45, 46]. However, this approach has a drawback in selecting the target virus from a vast background of helper viruses. This problem was solved by the advent of plasmid-based systems that rescue IAV entirely from cloned cDNAs [47-49]. These systems have been successfully used to resurrect the 1918 pandemic virus [50, 51] and create a reassortant H5/H1N1 virus that is transmissible among ferrets [52]. In addition, the application of the systems for developing live, attenuated influenza vaccines has also been explored [40].

1.3.2 Host innate immune defenses against IAV infection

Upon IAV infection, host innate and adaptive immune systems are mobilized and aim at clearing the infection (reviewed in [53, 54]). The innate immune systems comprise non-specific host defenses that can be immediately mobilized to fight against invading pathogens, whereas the adaptive immune systems consist of highly specialized defense mechanisms with the ability to recognize and remember a specific pathogen, providing a rapid clearance of the pathogen in the persisting or next encounter. The ability to fully recover from IAV infection and to protect the host from the next infection mainly relies on the adaptive immunity that includes the production of virus-neutralizing antibodies by B-cells and efficient elimination of infected cells
by cytotoxic T cells [55]. Nevertheless, the innate immunity is indispensable for early protection against IAV and it will be the focus for the rest of this section.

Non-cellular innate immunity against IAV infection

The first barrier of innate immune defenses against IAV include the mucociliary blanket of surface lung epithelium and non-specific inhibitors such as collectins (e.g. surfactant protein A, surfactant protein D, mannose-binding lectin, and conglutinin) and mucins. The most interesting inhibitor is the surfactant protein D, a potent inhibitor of HA activity that causes IAV aggregation and enhances phagocyte activation [56].

Intracellular innate sensing of IAV infection

In infected cells, IAV is sensed via pattern recognition receptors (PRRs) that recognize ssRNA and/or dsRNA. These include toll-like receptors (TLR3 and TLR7), DEAD-box polypeptide 58 (DDX58; also known as retinoic acid-inducible gene I (RIG-I)), and NLRs (i.e. NOD-like receptor family pryin domain containing 3 (NLRP3) and nucleotide-binding oligomerization domain containing 2 (NOD-2)). The recognition of IAV by these PRRs leads to the induction of various cytokines, including inflammatory cytokines (e.g. IL1B, IL6, IL12 and TNF), chemotactic cytokines (chemokines; e.g. IL8 and CCL5), and antiviral cytokines (IFN-α and IFN-β). TLR3 recognizes influenza dsRNA in alveolar and bronchial epithelial cells and it may stimulate the secretion of IL6, IL8, CCL5 and IFN-β [57]. In spite of its importance for immune response, mice deficient in TLR3 interestingly have an unexpected survival advantage following IAV infection [58]. In contrast, TLR7 may
promote the production of IL6, IL12 and IFN-α upon sensing influenza ssRNA, as demonstrated in plasmacytoid dendritic cells (pDCs) [59]. Similarly, DDX58 is able to recognize influenza ssRNA and induce IFN-α in pDCs [60]. DDX58 is also shown to be activated by IAV in epithelial cells but unlike TLR3, it mediates both type I IFN and proinflammatory responses [61]. Finally, NLRP3 recognition to influenza ssRNA or dsRNA leads to the formation of NLRP3 inflammasomes that regulate IL1B maturation (conversion of pro-IL1B into IL1B) in macrophages and dendritic cells [62, 63], while NOD2 recognition to ssRNA trigger the activation of interferon regulatory factor 3 (IRF3) and production of IFN-β, leading to type I IFN antiviral responses [64].

**Cellular innate immunity against IAV infection**

In addition to innate defenses at the molecular level, innate immunity against IAV also includes the mobilization of innate immune cells to contain the virus and phagocytose the infected cells. Among the innate immune cells, neutrophils and macrophages are the most dominant innate immune cells at the site of infection. Though both cells are known as a professional phagocyte, they do not phagocytose IAV directly. Some evidence demonstrated that neutrophils and macrophages are able to phagocytize IAV infected cells that undergo apoptosis *in vivo* [65]. The importance of these cells in host defense during IAV infection is shown by direct depletion of the cells (using anti-Gr1 (mAb RB6-8C5) or anti-Ly-6G (mAb 1A8) for neutrophils, clodronate for macrophages) before virus inoculation in animal models, which results in uncontrolled virus growth and increased severity and mortality [66-70].

Despite the crucial role of neutrophils and macrophages to limit viral spread, a high influx of both cells into the lungs has a potential to cause a severe lung injury
and enhanced production of cytokines such as IFN-γ and TNF-α that can be detrimental to the host [66]. Recently, a comparison of histopathologic features of infected lungs in control animals and animals depleted of either macrophages or neutrophils demonstrated that a more severe lung injury – characterized by ARDS-like pathologic signs – could be observed in macrophage depleted animals, but not in control and neutrophil depleted animals [71]. This result highlights that a severe lung injury during IAV infection is more closely related to neutrophil activities than macrophage activities.

Besides neutrophils and macrophages, the roles of dendritic cells (DCs) and natural killer (NK) cells have also been investigated. The main role of DCs is to function as a professional antigen-presenting cell, linking the innate immunity to adaptive immunity by degrading influenza viral proteins and subsequently displaying a fragment of the viral proteins on cell membrane via the major histocompatibility complex (MHC) class I to activate CD8+ cytotoxic T-cells or class II to activate CD4+ helper T-cells (Th cells). The role of DCs in stimulating proliferative and cytolytic responses of CD8+ and CD4+ T-cells had been demonstrated in vitro [72]. On the other hand, NK cells lyse IAV-infected cells expressing the viral HA after binding of this antigen with either cytotoxicity receptors NCR1 (NKp46) [73] or NCR2 (NKp44) [74]. Furthermore, once adaptive immunity is activated, NK cells may also clear the antibody-bound IAV infected cells through a mechanism called antibody-dependent cell cytotoxicity [75]. Importantly, depletion of NK cells before IAV infection increases morbidity and mortality, which signifies the crucial role of NK cells in early innate immune defenses during IAV infection [76-78].
1.3.3 Systems biology of host response to IAV infection

Recent studies that employ systems biology approaches have unravelled aspects of host response to IAV infection, in vivo and in vitro. Specifically, in vivo studies have uncovered the global pulmonary response to IAV infection in various animal models (including mice, macaques, ferrets and pigs), as well as the extent of the contribution of various cell types in infected lungs to the response; whereas in vitro studies have mainly revealed the genomic transcriptional response of epithelial cells and macrophages infected with the virus. Here we present some interesting results for both in vivo and in vitro studies and conclude the patterns of the genomic response. In addition, we also briefly highlight the use of genomic approaches for influenza studies in clinical settings.

Global pulmonary response to IAV infection

The earliest study on global pulmonary response to IAV infection was carried out by Kash et al. [1] using a mouse model. This study revealed marked alterations in gene expression changes in murine lungs collected on day 1 postinfection with highly virulent WSN33 virus or its recombinant (that contains HA and NA from 1918 pandemic virus), but not less virulent NC99 virus. These include the early regulation of genes involved in inflammation (high TNF, CSF1 (M-CSF), IL2RG, B2M, SELP), cell death (high TNF, GADD45G, NFKBIA), and oxidative stress (high HMOX1, low GPX3 and PRDX5). In the following study [2], the same group of researchers provided insights on global response of murine lungs infected with fully reconstructed 1918 pandemic virus (r1918), less virulent TX91, and their recombinants. This study highlights a significant activation of proinflammatory, type I IFN and cell death responses that is concomitant with severe pathology and higher weight loss, mortality
rate, and levels of virus titers in lungs of the r1918-infected group. Most interestingly, the expression of cytokines and genes associated with the activation of neutrophils, macrophages, NK and Th1 cells in lungs infected with the r1918 virus was higher and prolonged until day 5.

Mouse models have also been used to reveal differences in global pulmonary responses early after infection with r1918 and highly pathogenic avian H5N1 virus VN1203 [195]. Interestingly, VN1203 was more pathogenic than the r1918 in mice, and this was associated with early over-expression of inflammatory related genes, including the key inflammasome components CASP1, IL1B and NLRP3, observed since day 1 postinfection. In addition, VN1203 disseminated to multiple organs, and this was correlated with prolonged inflammatory response and the impairment of lipoxin-mediated anti-inflammatory responses. Another interesting observation was that both r1918 and VN1203 could induce type I IFN response even in mice lacking the corresponding receptor.

Next, Pommerenke et al [196] used a non-lethal PR8 (H1N1) infection model to produce transcriptomic data representing distinct phases of murine lung responses that span over a period of 60 days after infection. Overall, their analysis revealed the transition from innate immune response that peaked after 2 days postinfection to adaptive immune response of T and B cells that peaked after 8 and 14 days postinfection, respectively, followed by repair processes in the late phase of infection. By using mice deficient for RAG2, a gene that is essential for V(D)J rearrangement during B and T cell development, the study also demonstrated that the failure of infected RAG2-/- mice to elicit adaptive immune response could be captured by analyzing their transcriptome.
Very recently, Brandes et al [204] has used murine PR8 infection model and employed a top-down systems approach to dissect the contribution of direct pathogen damage versus damage caused by over-activation of immune response. The top-down approach involved a deconstruction of lung transcriptomes using transcriptomes of sorted cells (including nonhematopoietic cells, neutrophils, lymphocytes, monocytes and alveolar macrophages) that were collected following lethal and non-lethal infections. Data-driven gene set analysis was performed and the results revealed that antiviral and IFN responses were shared by all cell types, whereas inflammatory innate immune response was mainly driven by neutrophils. In addition to the greater spread of the virus, only the inflammatory but not antiviral signatures were associated with the lethality. Moreover, the study has also provided evidence for the role of the CXCL2-feedforward inflammatory circuit in regulating the magnitude of neutrophil infiltration.

In addition to comparing global pulmonary response to varying strains and subtypes of IAVs, mainly of H1N1 and H5N1 subtypes, mouse model has also been used to compare transcriptomic host response to H3N2 virus infection (i.e. HKx31 virus, which is classified to be of intermediate virulence in mice [69]) and pneumococcal infection by Streptococcus pneumonia [79]. This study showed that both viral and bacterial infections triggered the expression of TNF-α, IL6, and type I and II IFN-inducible genes (e.g. GBP2, GBP4, IFI16, IFI35 and IFI204). Nevertheless, more pronounced expression of myeloid cell-associated genes (including some neutrophil and monocyte/macrophage chemoattractant CCL12, CXCL5, CCL3 and CCL4) were observed in pneumococcal infection. Furthermore, while the expression of T-cell-associated genes was conspicuous in IAV infection, the quantity of pulmonary B-cells and the expression of associated genes dropped in
pneumococcal infection. Overall, these results highlight that the clearance of pneumococcal infection mainly relies on the innate immune mechanism, whereas the clearance of IAV infection strongly depends on adaptive immunity.

While the mouse model was being investigated, global pulmonary response to IAV infection in pigtailed macaques (*Macaca nemestrina*) was also being assessed by using both genomic and proteomic approaches [80, 81]. The virus that was used for the experiments was the less virulent TX91 of H1N1 subtype that causes a mild influenza in the animals (less virulent), with characteristics similar to those seen in human patients previously naive for the disease. Transcriptomic analyses were performed on lungs, tracheobronchial lymph nodes, as well as peripheral white blood cells collected within 7 days postinfection. The most interesting outcome of the analyses was the identification of a considerable expression of IFN-inducible genes, including IFIT2, IFI6, ISG15, ISG20, GBP1, GBP3, IRF7, MX1, MX2, OASL, OAS1 and OAS3, in white blood cells. Concomitantly, proteomic analyses of infected lungs indicated apparent regulation of IFN-inducible MX1, IFIT2, IFIT3, and GBP2, and complement C7 during the infection. Overall, these reports demonstrated that nonhuman primate is an excellent model to investigate IAV infection.

Further studies on global pulmonary response to IAV infection in macaques have focused on modeling highly virulent IAV infections [3, 5, 6]. The viruses being investigated were of H1N1 (less virulent: K173, TX91; highly virulent: r1918 and its recombinants with TX91) and H5N1 (highly pathogenic: VN1203) subtypes. For gene expression analyses, bronchi or whole lungs were collected on day 3, 6 and 8 postinfection in [3]; on day 1, 2, 4 and 7 postinfection in [5]; and at 12, 24 and 48 h postinfection in [6]. Overall, these transcriptional studies demonstrated that early and prolonged innate immune response – including high over-expression of
proinflammatory, cytokine, and IFN-inducible genes – is the major hallmark of a highly virulent IAV infection. This response was shown to be rapidly developed as early as 12-24 h postinfection, and still unabated until day 7-8 postinfection. In contrast, a less virulent IAV infection only led to a moderate over-expression of genes associated with innate immune response that peaks on day 3-5 postinfection and the levels of the expression gradually decreased on the subsequent days. Of interest, the comparison between responses to r1918 and VN1203 in macaques showed intriguing differences and similarities to those observed in mouse model discussed previously. In macaques, VN1203 infection only caused less injury, lower levels of viral loads, and lower number of apoptotic cells than those caused by r1918 early after infection [6]. These results were corroborated with global gene expression data from infected animals that demonstrated differential regulation of inflammation, cell death and type I IFN-related genes in response to the viruses. Most interestingly, it was the r1918 infection that caused over-expression of key inflammasome components (NLRP3 and IL1B); while during VN1203 infection, these two genes were under-expressed

In addition to mice and macaques, IAV infection has also been modelled in vivo in pigs and ferrets. In pigs [82], animals were infected with a swine H1N1 virus (i.e. A/swine/Hubei/101/2009) that caused a mild influenza. Differentially expressed genes associated with inflammatory and type I IFN responses were observed in lungs on day 3 postinfection, concomitant with lung histopathology that showed edema, large amount of infiltrates, alveolar wall thickening and lesions. However, less immune related genes were observed on day 7 postinfection and the majority of the genes were associated with signal transduction, transcription, metabolism, development and transport. Lung response to 2009 pandemic H1N1 has also been investigated using swine in comparison with mouse and macaque infection models
The comparative study has pointed out the variation in gene expression patterns among the three species, including differences in inflammatory response and lipid metabolism. In ferrets [4], the hyperinduction and persistent expression of IFN-related genes and suppression of T- and B-cell signaling were observed in lungs infected with the highly pathogenic VN1203 virus throughout the course of study (up to 6 days postinfection), but they were less notable in lungs infected with the less virulent PN99 of H3N2 subtype. Of particular interest was the robust expression of CXCL10 (also known as IFN-γ-inducible protein 10 (IP10)) in H5N1-infected group, which may play a role in recruiting immune cells, including Th1 lymphocyte, NK and neutrophils. Blocking the activity of CXCR3, the cognate receptor of CXCL10, strikingly led to reduced viral loads in the lungs and delayed mortality in H5N1-infected ferrets.

In summary, we conclude that lungs infected with highly virulent IAVs elicit an early and prolonged innate immune response, whereas lungs infected with less virulent ones elicit a moderate response at a later time-point. In particular, highly virulent IAV infections induce an early and persistent hyperinduction of cytokine/chemokine and type I IFN-related transcripts. Genes of particular interest include IL1B, CCL3, CCL4, CCL5, CXCL10, TNF, IL6, ISG15, and MX1. Concomitantly, a highly virulent infection in lungs is also accompanied with severe cell death response.

**Global cellular response to IAV infection**

Global host response to IAV infection has been investigated *in vitro* in specific cells earlier than *in vivo* in lungs. The earliest *in vitro* study investigated global response of epithelial cell line A549 at 8 h postinfection, with focus on the role of viral NS1 in
suppressing IFN response [7]. The study has demonstrated that mutant H1N1 viruses (i.e. PR8 lacking the NS gene and PR8 containing truncated NS gene) were able to induce a higher expression of IFN and NF-κβ related genes compared to a wild type PR8. In addition, the study has also revealed that in comparison to wt WSN33, a recombinant WSN33 containing r1918 NS gene elicits a lower level of antiviral ISG-genes.

The impact of NS1 on global lung epithelial A549 was further investigated by Billharz et al [8], which used an improved microarray technology, a larger array of genes, and more observation times (2, 6 and 24 h postinfection). The viruses used for the experiments were also of H1N1 subtypes, i.e. the r1918, TX91, and their recombinants containing other’s NS gene. The study suggest that r1918 NS requires the complete r1918 genome for its maximal functionalities since a pronounced difference was only observed when comparing the global response of the cells to r1918 and its recombinant that contains TX91 NS gene, but not TX91 and its recombinant that contains r1918 NS gene. Strikingly, although r1918 virus infection induced over-expression of many immune related genes associated with cytokine/chemokine signaling (e.g. CCL2, CCL19, CCL21 and CCR10) and activation of neutrophils and lymphocytes (e.g. ELA2, MMP9, CD4, THY1, CD79A and SP11), it led to under-expression of IFN-related genes (e.g. DDX58, IFIT1, IFIT3, ISG15, OAS1, OAS2, and STAT1). Another finding was the stronger suppression of genes associated with lipid metabolism in cells infected with viruses containing r1918 NS in comparison with viruses containing TX91 NS.

A more recent study by Chengjun et al [11] engaged a computational approach that is independent from existing gene interactions to discover temporal transcriptional subnetworks regulated in human epithelial Calu-3 cells infected with
highly virulent H5N1 and less virulent H1N1 viruses, i.e. VN1203 and NL602, respectively. This study recovered over-expressed modules associated with immune and keratin signaling networks and under-expressed modules associated with cellular homeostatic processes and cell cycle in VN1203-infected Calu3. In contrast, NL602-infected cells showed an attenuated host response; in particular, the number of differentially expressed genes was much smaller than those observed in VN1203 infection, some cytokines (e.g. IFN-α4, IFN-α5, and IFN-α8) and keratin molecules were not over-expressed, and the expression of IL6, TNF and CSF3 (G-CSF) was delayed.

In addition to lung epithelial, transcriptome changes in monocyte-derived and alveolar macrophages after infection with IAVs have also been investigated. In [9], cultures of primary human monocyte-derived macrophages were infected with highly pathogenic H5N1 and less virulent seasonal H1N1 viruses isolated from human, i.e. VN3212 and HK54, respectively. Transcriptomic analyses revealed over-expression of type I IFN-inducible and cytokine/chemokine genes in macrophages infected with both viruses at 6 h postinfection, but not at previous time-points (1 and 3 h postinfection). However, although differentially expressed genes in both infections were highly overlapped (109 in H5N1-infected group, 64 in H1N1-infected group, and 60 overlapping genes), a much stronger response was observed in H5N1 infection. Furthermore, measurements of TNF-α, CXCL10, IFN-β, IFN-α1, SOCS1 and SOCS3 gene levels with qPCR showed a good correlation with microarray data. As a conclusion, the report suggested that the synergy between type I IFN and TNF-β signaling pathways potentially contributes to enhance proinflammatory response in H5N1 infection.
In [10], large scale analyses of global responses of alveolar macrophages (AMs) to IAVs and bacteria were carried out. For this, AMs of mice and macaques were infected with either influenza FJ02 of H3N2 subtype, PR8 of H1N1 subtype, *Mycobacterium tuberculosis* or *Francisella tularensis*, and then total RNAs of AMs were collected at 0, 1, 2, 6, 12 and 24 h postinfection for microarray experiments. One astonishing outcome of this study was that mouse AMs failed to produce cytokines in response to PR8 that are virulent in mice. On the other hand, when mouse AMs were infected with FJ02, or when macaque AMs were infected with any virus, cytokine expression changes could be observed. Nevertheless, the induction of cytokine and chemokine responses, in addition to type I IFN response, has recently been observed in human AMs infected with PR8 virus [197].

In summary, both lung epithelial cells and macrophages were able to elicit immune response against IAV infection by expressing type I IFN-inducible and cytokine/chemokine genes. In addition to this response, genes related to various processes such as lipid metabolism, cell cycle and keratin signaling were also found to be differentially regulated in infected epithelial cells. Of particular interest, the IFN response in epithelial cells could be suppressed by the viral NS1, and the suppression level was dependent on the context of the viral genome. For macrophages, although the highly pathogenic H5N1 virus infection in human macrophages induced high expression levels of genes related to immune and interferon response, the highly virulent PR8 infection interestingly failed to induce similar response in mouse macrophages. Hence, the highly virulent IAV infection in macrophages may have no implication to the dysregulation of innate immune response observed *in vivo*; infection in other types of immune cells may play a role.
Global response to IAV infection in clinical settings

Genomic approach has also been used in clinical settings to develop a peripheral blood gene expression signature that can classify individuals with symptomatic IAV infection. In the early study, Ramilo et al [198] used peripheral blood leukocytes to obtain a transcriptional signature that distinguished patients with IAV infection from those with bacterial infection. The signature genes identified in this study are dominated by IFN-inducible genes, including OAS1, ISG15 and MX1. Then, using human viral challenge studies, Zaas et al [163] identified a signature from whole peripheral blood that could distinguish individuals with peak symptom of acute viral respiratory infections (including influenza H3N2 infection) from healthy individuals. The signature was able to accurately classify patients with IAV infection in the previous study.

Further study on the full temporal spectrum of gene expression dynamics in whole peripheral blood of human subjects challenged with influenza H3N2 infection has been carried out by Huang et al [199]. In this study, gene expression profiles were measured within 108 hours post-inoculation at an interval of ~8 hours. An accurate genomic signature that discriminates asymptomatic subjects and symptomatic subjects before onset and after onset of acute infection was obtained by using an unsupervised factor analysis. Using a supervised analysis, important host factors whose temporal expression patterns are significantly different in symptomatic and asymptomatic subjects could be determined. One interesting outcome of the analysis is the regulation of IL1B gene, which was under-expressed in the asymptomatic subjects, but over-expressed in the symptomatic subjects.

More recently, gene signatures based on whole peripheral blood transcriptome from human subjects challenged with influenza H3N2 and H1N1 have also been
developed [200]. The gene signatures at the time of maximal symptoms were obtained independently for H3N2 and H1N1 studies, and it appeared that the signature from one study could be applied to another study. However, the signatures were not sufficient to discriminate H3N2 and H1N1 as they are highly similar. Further investigation revealed that the signature could be detected as early as 29 hours post-inoculation and can be applied to accurately identify patients with pandemic 2009 H1N1 infections in a clinical cohort.

Shortly, the results obtained in these studies shows promise for the development of diagnostic tools for IAV infection, especially for early detection. In addition, these studies also provide biological insight into the global response to IAV infections, which could be invaluable for the development of influenza drugs.

1.3.4 Neutrophils

Here, we highlight some aspects of neutrophils that become the object of our studies. These include the hematopoiesis and defense mechanisms of neutrophils, as well as surface receptors expressed by the cells and their signal transductions.

Hematopoiesis of neutrophils

Neutrophils, also known as polymorphonuclear leukocytes (PMN or PMNL), constitute the majority of white blood cells (leukocytes), i.e. about 50%-70% of total leukocytes. Like other leukocytes, neutrophils arise from pluripotent hematopoietic stem cells (HSCs) residing in fetal liver or adult bone marrow. Specifically, neutrophil maturation begins when myeloblasts differentiate into committed myeloid progenitors known as promyelocytes. Neutrophil maturation is terminal (the cells do not further replicate) and it is a complex process that requires specific growth factors, including
Chapter 1

the granulocyte-colony stimulating factor (G-CSF), granulocyte/macrophage-colony stimulating factor (GM-CSF; or CSF2), and interleukin 3 (IL3). In addition to growth factors, retinoic acid that activates retinoid signaling pathways also plays a crucial role for neutrophil maturation. In [84], the binding of all-trans retinoic acid (ATRA) to RAR-RXR heterodimers has been shown to initiate transcriptions of genes that coordinate neutrophil maturation.

The characteristics of neutrophil maturation are the sequential changes of nucleus structure and the formation of primary (azurophil), secondary (specific) and tertiary (gelatinase) granules that contain antimicrobial proteins. According to these, five stages of neutrophil maturation are defined: promyelocytes, myelocytes, metamyelocytes, band cells, and (mature) neutrophils. Promyelocytes and myelocytes have spherical nuclei, but their granule compositions are different. Promyelocytes mainly contains primary granules and in myelocytes, secondary granules appear. Metamyelocytes elicit the early stage of nucleus segmentation, where the nucleus is indented, forming a broad-bean-like shape. In the next stage, the nucleus transforms to a band shape, hence it is called a band cell. Tertiary granules start to appear during this stage. Finally, the mature neutrophils can be recognized from its multi-lobed (2-5 lobes) nucleus and more specific and tertiary granules are formed. Under giemsa staining, the cytoplasm of promyelocyte is dark blue-purple, and at the later stages, the cytoplasm becomes transparent.

Recently, systems biology approaches – that include genomic and proteomic approaches – have been applied to uncover transcriptional and translational programs that underlie the process of neutrophil maturation from promyelocytes in vivo [85, 86] and in vitro [87, 88]. Overall, these studies have demonstrated that neutrophil maturation is a complex process governed by highly coordinated molecular programs.
Chapter 1

In addition to the trends of under-expression of genes associated with cell cycle towards mature neutrophils, other outcomes of these studies also include the identification of a broad spectrum of mature neutrophil receptors and candidates for granule proteins. Hence, systems biology approaches have taken us to a new level of understanding on neutrophil differentiation as well as its potential responsiveness toward various inflammatory stimuli.

**Innate immune defenses of neutrophils**

The life of neutrophils is short. Neutrophils naturally die within 1-2 days of being released from bone marrow or they will die after fulfilling their ultimate function as the first line of defense against pathogens. Neutrophils migrate rapidly to the sites of infection, transmigrate across capillary endothelium, become activated and initiate their defense mechanisms to kill pathogens through the release of reactive oxygen species (ROS; e.g. superoxide, hydrogen peroxide and hypochlorous acid) and antimicrobial granule proteins. Classically, these defense mechanisms are activated once neutrophils engulf the pathogens by a process known as phagocytosis, internalize them into phagocytic vacuoles (phagosomes), and fuse the phagosomes with intracellular granules to form the phagolysosomes. Whereas antimicrobial proteins are released from granules synthesized during neutrophil maturation, microbicidal oxidants are mainly generated through the membrane-bound NADPH oxidase enzyme complex. The role of NADPH oxidase is to transfer electrons from NADPH to oxygen in order to form superoxide which subsequently can be converted into other ROS by a spontaneous or enzymatic reaction.

In addition to the phagocytic defense mechanism, a new mechanism of clearing pathogens has been described recently by Brinkman et al [89]. This
mechanism involves the release of extracellular nuclear DNA containing antimicrobial peptides that can entrap and kill pathogens; hence, it is called as neutrophil extracellular traps (NETs). In addition to nuclear DNA, mitochondrial DNA has also been shown to form NETs in response to GM-CSF priming and subsequent stimulation with LPS or complement C5a [90]. Both types of NETs are dependent on the production of ROS; but while mitochondrial NETs are released by live neutrophils, nuclear NETs are considered as a form of cell death (hence, dubbed as NETosis) [90, 91]. NETs have been associated with infectious diseases caused by bacteria, fungi and parasites, as well as non-infectious diseases such as preeclampsia and systemic lupus erythematosus [92]. Interestingly, the formation of NETs has also recently been reported in a murine IAV infection model [71].

**Neutrophil receptors and their signal transductions**

The biological activities of activated neutrophils are modulated by surface receptors that are over-expressed markedly at the end stage of neutrophil maturation. Studies that employ genomic approaches have revealed a large array of neutrophil surface receptors (reviewed in [93]), which include:

- cytokine receptors:
  - interleukin receptors: IL1R, IL4R, IL6R, IL10R, IL13R, IL17R and IL18R
  - Chemokine receptors:
    - CC chemokine receptors: CCR1, CCR2 and CCR3
    - CXC chemokine receptors: CXCR1 (IL8RA), CXCR2 (IL8RB) and CXCR4
  - tumor necrosis factor receptors: TNFR1 and TNFR2
transforming growth factor receptor: TGFBR2

IFN receptors: IFNAR1, IFNAR2, IFNGR1 and IFNGR2

leukocyte immunoglobulin-like receptors (LIRs): LIR1 to LIR7 and LIR9

PRRs: TLR1, TLR2, TLR4, TLR6, and TLR8

triggering receptor expressed by myeloid cells: TREM1

fragment, cryztallizable (Fc) receptors: FCAR, FCGR2A, FCGR2B, FCGR2C, FCGR3A, FCGR3B, and FCER1G

cell death receptors: TRAILR and TNFR1

Note that not all neutrophil surface receptors were successfully recovered by genomic approaches, e.g. Fas death receptor [94] and the rest of TLR1 to TLR10, except TLR3 [95], which is likely because the expression occurs earlier. Here, we briefly discuss the importance of each type of receptor and highlight IFN and TREM1 signaling pathways at the end of this section due to their relevance in our experimental results discussed in Chapter 2 and 3.

Cytokine receptors play a role in intercellular communication and specifically, chemokine (chemotactic cytokine) receptors deliver a signal for cell migration. Some interesting cytokine receptor-ligand bindings are IL1R that binds to interleukin 1 (IL1), CXCR1 that binds to interleukin 8 (IL8), and CXCR2 that binds to IL8 and CXCL1 (Gro-α or KC). The former binding plays a role in the regulation of inflammatory processes, whereas the last two are essential in neutrophil chemotaxis and activation.

The roles of LIRs in neutrophils have not yet been explored, but studies in other immune cells showed their importance in regulating the immune response, playing inhibitory and stimulatory roles in cellular responses [96].
Next receptors, i.e. the TLRs, are a type of PRRs. PRRs are essential for recognition of pathogen-associated molecular patterns (PAMPs) that are broadly shared among pathogens but distinguishable from host molecules. In general, signaling by different TLRs stimulates distinct patterns of gene expression associated with innate and adaptive immunity [97, 98]. In neutrophils, TLRs activation has been shown to induce cytokine production, superoxide generation, reduced chemotaxis, increased in phagocytosis and prolonged neutrophil survival [95, 99]. As described later, some TLRs work in conjunction with the triggering receptor expressed by myeloid cells 1 (TREM1) to amplify their signals to trigger associated cellular processes.

Next, the Fc receptors bind to the immunoglobulins (antibodies) produced by plasma cells: FCAR for immunoglobulin-α (Ig-α), FCGR for immunoglobulin-γ (Ig-γ), and FCER for immunoglobulin-ε (Ig-ε). The most common Fc receptor is the FCGR, which mediates phagocytosis of opsonised pathogen captured by Ig-γ [100, 101]. Among the FCGRs, FCGR3B is exclusively expressed in neutrophils; but interestingly, this receptor may have no role in increasing phagocytosis of neutrophils although it induces intracellular Ca$^{2+}$ release and receptor capping (receptor aggregation) [102].

Finally, the cell death receptors TRAILR and TNFR1 play a role in programmed cell death (apoptosis). The ligand for TRAILR and TNFR1 are the TNF-related apoptosis-inducing ligand (TRAIL) and the tumor necrosis factor-α (TNF-α), respectively. Unlike TRAIL that only promotes apoptosis, TNF-α uniquely promotes apoptosis at high doses and cell survival at low doses [94]. TNF-α-induced apoptosis is mediated by mitochondrial ROS production and activation of calpain, Jun N-terminal kinase & mitogen-activated protein kinase (JNK MAPK), and Src homology
domain 2 (SH2)-containing tyrosine phosphatase-1 (SHP1), whereas the survival is mediated by activation of phosphoinositide-3-kinase (PI3K), protein kinase C-δ (PKC-δ) and extracellular regulated kinase 1/2 (ERK1/2), release of IL8, and expression of pro-survival factors AI and B-cell lymphoma-extra large (BCL-XL) [103].

As mentioned previously, we highlight IFN and TREM1 signaling pathways in more details as follows:

- **IFN Receptors and Their Signaling Pathways.** Two complexes of IFN receptors exist: IFNAR (a heterodimer of IFNAR1 and IFNAR2) and IFNGR (a heterodimer of IFNGR1 and IFNGR2). Accordingly, IFNs are classified into type I and type II, which bind to IFNAR and IFNGR respectively. Type I IFNs consist of about 20 IFNs, including IFN-α and IFN-β that are predominantly synthesized during viral and bacterial infections. On the other hand, type II IFN has only one member, i.e. IFN-γ. The current understanding views type I IFNs as a potent antiviral immunoregulator, while type II IFN as an enhancer of antibacterial immunity. Despite these differences, the pathways utilized by type I and type II IFNs to transduce their signals are overlapped (FIG.1.2 [104]).

  Both type I and type 2 IFNs depend on the activation of Janus kinase (JAK)-signal transducer and activator of transcription (STAT) signaling pathway to transcribe a large set of antiviral gene products associated with IFN-stimulated response element (ISRE) and IFN-γ activated-site (GAS) promoter sequences. Specifically, the binding of IFNs activates the JAKs that are permanently bound through their respective receptors (JAK1 and tyrosine kinase 2 (TYK2) for IFNAR, and JAK1 and JAK2 for IFNGR). Activation of JAKs subsequently leads to STAT phosphorylation on tyrosine residues.
(STAT1 and STAT2 for IFNAR and STAT1 for IFNGR) and the phosphorylated STATs form dimers (STAT1 homodimers and STAT1-STAT2 heterodimers downstream of IFNAR, and STAT1 homodimers downstream of IFNGR). While STAT1 homodimers move to nucleus immediately to bind GAS promoter sequence, STAT1-STAT2 heterodimers form a complex called as IFN-stimulated gene factor 3 (ISGF3) with a third protein, IFN regulatory factory 9 (IRF9), to bind ISRE promoter sequence to transcribe IFN-inducible genes.

FIG. 1.2. Type I IFN and type II IFN signaling pathways.

- **TREM1 and Its Signaling Pathway.** TREM1 is a transmembrane glycoprotein of Ig-superfamily selectively expressed in myeloid cells, including neutrophils and CD14\textsuperscript{high} monocytes/macrophages. TREM1 signaling is mediated through a
transmembrane immunoreceptor tyrosine-based activation motif (ITAM)-
bearing adapter molecule, tyrosine kinase-binding protein (TYROBP) also
known as DNAX-activation protein 12 (DAP12) [105]. Although the natural
TREM1 ligands are so far unknown, the roles of TREM1 in the innate immune
response and pathways associated with these roles have been uncovered to some
extent. TREM1 signaling in neutrophils results in the degranulation and the
production of IL8, myeloperoxidase, lactoferrin [105, 106], and respiratory
burst [106, 107]; while in monocytes/macrophages, it results in increased
intracellular Ca\(^{2+}\), the production of cytokines and chemokines (e.g. CCL2,
CCL3, CCL7, IL8 and TNF-\(\alpha\)) [105, 108], and over-expression of co-
stimulatory molecules (e.g. CD32 (FCGR2), CD40, CD54, CD83 and CD86)
[105, 109]. In general, TREM1 has been proposed to play a role in the
amplification of cytokines and chemokines and other proinflammatory
responses induced by TLRs [106, 108] and NAIP, CIITA, HET-E, TP-I-leucine-
rich repeat receptors (NATCH-LRRs ; NLRs) [110].

Signal transduction pathways triggered by TREM1 have also been
investigated in both neutrophils and monocytes/macrophages and they are
nicely summarized by Tessarz and Cerwenka [111] as illustrated in FIG.1.3.
Briefly, TREM1 engagement in neutrophils leads to the phosphorylation of
JAK2, LCK/YES-related novel kinase (LYN), protein kinase B (PKB; AKT),
ERK1/2, phospholipase C-\(\gamma\) (PLC-\(\gamma\)), and IL1R-associated kinase (IRAK).
Downstream of these are the activation of STAT3, STAT5, and RELA (also
known as p65; a subunit of the nuclear factor-kappa B (NF-\(\kappa\)B) family).
Moreover, TREM1 in neutrophils can synergize with TLR2, TLR4 and TLR7/8
to induce respiratory burst, degranulation and phagocytosis, as well as accelerate neutrophil apoptosis.

![TREM1 signaling pathway](image)

**FIG. 1.3.** TREM1 signaling pathway in neutrophils (black), monocytes (white) or both (gray); hatched lines illustrate protein degradation (adapted from [111]).

On the other hand, engagement of TREM1 in monocytes/macrophages elicits activation of ERK1/2 and PLC-γ, and also increased level of NF-κB subunits comprising NF-κB1 (p50) and RELA. The complex formation of caspase-recruitment domain 9 (CARD9) and B-cell CLL/lymphoma 10 (BCL10) have been suggested to be essential for transcriptional activation of NF-κB. Furthermore, TREM1 in monocytes/ macrophages can synergize with and TLR2 to 4 and TLR9 to amplify proinflammatory cytokine and chemokine production. In addition to these, TREM1 engagement in both neutrophils and monocytes/macrophages also leads to the phosphorylation of the adaptor protein...
non-T cell activation linker (NTAL) that forms a complex with growth factor receptor-bound 2 (GRB2), ubiquitin-protein ligase casitas B-lineage lymphoma (C-CBL), and guanine nucleotide exchange factor son of sevenless 1 (SOS1). NTAL activation is subsequently involved in Ca\textsuperscript{2+} mobilization and negative regulation of ERK1/2 phosphorylation and the production of IL8 and TNF-\(\alpha\); thus, balancing the signals mediated by TYROBP.

1.3.5 IAV infection in neutrophils

Neutrophils, together with macrophages, are the first and most dominant innate immune cells recruited to the site of injury during IAV infection. However, neutrophils are relatively more essential in the early containment of the virus than macrophages [112, 113]. In particular, neutrophils are shown to be important in controlling virus replication during an intermediate to highly virulent IAV infection, but not a less virulent one [69]. In addition to phagocytizing infected cells, the containment of IAV by neutrophils and macrophages may occur by direct uptake; in other words, both neutrophils and macrophages are infected. IAV infection in neutrophils likely occurs through the virus HA binding to various highly sialated membrane components such as CD43 and CD45 [114, 115]. Although some experiments indicated that specific strains of IAV can replicate in macrophages [116-118], no reports claimed that IAV productively replicates in neutrophils. Cassidy et al [18] has demonstrated that the nature of IAV infection in neutrophils is abortive in that some viral proteins can be synthesized but new virus progenies are not formed. Viral proteins found to be synthesized included HA, NP, NS1 and NS2 [18, 119].

Despite its virulence, several effects of IAV binding to surface sialic acid residues and IAV infection in neutrophils have been uncovered. These include the
ability of the viral HA to activate an atypical respiratory burst, where hydrogen peroxide is produced in the absence of detectable superoxide, and depress or deactivate neutrophil functions in terms of chemotaxis, degranulation, and bacterial killing [115, 120]. The IAV-induced respiratory burst activation, but not the IAV-induced functional depression, was shown to be preceded by activation of phospholipase C, phosphatidic acid production, intracellular calcium ion mobilization and pH changes [121-122]. Furthermore, the neutrophil respiratory burst and increased expression of FAS antigen and surface and extracellular FAS ligand have been implicated as a cause of IAV-induced neutrophil apoptosis [19, 20]. The rate and extent of apoptosis induction are more pronounced when neutrophils were treated with IAV and bacteria such as *Escherichia coli*, *Streptococcus pneumonia*, or *Staphylococcus aureus* [19, 20, 124]. Finally, IAV was also shown to activate or increase the surface expression of toll-like receptors (TLRs), including TLR2, TLR7 and TLR8. Increased surface expression of TLR2 on neutrophil upon IAV exposure occurs rapidly and it modulates functional response to ligands that bind to the receptor, e.g. increased uptake of zymosan or *S. aureus*, and increased neutrophil respiratory burst response to peptidoglycan (PGN) [124]. On the other hand, the activation of TLR7/8 potentially contributes to cytokine and chemokine induction (e.g. IL8 and CCL4), and this induction requires endosomal acidification and viral uncoating [119].

1.3.6 Induced differentiated myeloid cell lines as an alternative source for neutrophils

Traditionally, *in vitro* study of neutrophils is done by using neutrophils isolated from whole blood or bone marrow. About $3.5 \times 10^5$ and up to $6 \times 10^6$ neutrophils can be
drawn from whole blood and bone marrow from one mouse, respectively [201]. However, this approach is technically difficult in term of obtaining highly pure and viable neutrophils, laborious and expensive, and may provide limited number of neutrophils due to ethical issue for killing many mice. Hence, possessing an alternative source of neutrophils that provides an easy-to-use in vitro system is helpful and for this, cell lines are the alternative. Unfortunately, no valid neutrophil cell lines exist since neutrophils are terminally differentiated (do not further replicate) and they naturally die within 1-2 days. However, there are several myeloid (neutrophil progenitor) cell lines that can be differentiated into neutrophils. These include HL-60 and NB4 for human, and 32Dcl3, MPRO and EPRO for mouse.

Historically, HL-60 and NB4 were isolated from leukemic patients: HL-60 is isolated from a 36-year-old female with an acute myeloblastic leukemia (AML) with maturation, an AML type M2 according to French-American-British (FAB) classification [125]; whereas NB4 is isolated from a 20-year-old woman with a typical acute promyelocytic leukemia, an AML type M3, carrying a t(15;17) (q22;q11-12) translocation [126]. On the other hand, murine 32Dcl3 cells are established from nonadherent cell populations harvested from continuous mouse bone marrow cultures [127]; MPRO cells were generated from FDCP mix A4 (long term bone marrow cultures from male BDF1 mouse that spontaneously differentiate into neutrophils and monocytes) transduced with a retroviral vector containing dominant-negative RARα403 construct, a truncated RARα harbouring a truncation in the ligand binding domain and activation function domain-1 (AF-1) [128]; and EPRO cells were derived from cells with erythroid, myeloid, and lymphoid potential, designated as EML, which also harbour RARα403 construct [129].
Protocols for maintaining these cell lines and differentiating them into neutrophils have been described in [130]. Briefly, both HL-60 and NB4 can be maintained in Roswell Park Memorial Institute (RPMI) medium containing serum only and no additional growth factors. On the other hand, 32Dcl3, MPRO and EPRO, which can be maintained in Iscove’s Modified Dulbecco’s Medium (IMDM) containing serum, requires additional growth factors. Specifically, 32Dcl3 requires IL3, whereas MPRO and EPRO require GM-CSF (to obtain EPRO from EML, IL3, stem cell factor and ATRA are required). In general, most of these cell lines can be differentiated into neutrophils by supplementing ATRA into their complete growth medium, except 32Dcl3 which requires G-CSF instead of ATRA and IL3.

Furthermore, functional responses of neutrophils derived from these four cell lines – including respiratory burst, chemotaxis toward multiple chemoattractants, phagocytosis and adhesion – have also been broadly discussed and assessed in [15]. Briefly, only neutrophils from induced differentiated MPRO and EML cells demonstrate a complete functional neutrophil activation. Induced differentiated HL-60 and NB4 cells have defective chemotaxis to multiple chemoattractants, while induced differentiated 32Dcl3 cells lack a respiratory burst. In addition, HL-60, NB4 and 32Dcl3 cells fail to properly over-express MAC-1 (a neutrophil adhesion molecule) after PMA stimulation. Hence, MPRO and EML are the best choice of myeloid cell lines used for providing an alternative source of neutrophils. Perhaps, the ability of neutrophils derived from MPRO and EML cells to release NETs, which has been demonstrated recently [17], indicates the complete functionalities of MPRO and EML neutrophils.
1.3.7 Computational and statistical approaches

A scheme for DNA microarray experiments, data preprocessing and data analyses

DNA microarrays, or simply microarrays, allow the measurements of expression levels (mRNA levels) of thousands of genes in particular cells, tissues or organs at once. To serve this purpose, thousands of spots or beads – each contains picomoles of a specific type of probe (a single strand of DNA with a specific sequence) – are arranged in an array on a solid surface of glass slide, a quartz wafer, or nylon membrane. These probes are the DNA fragments that are used to hybridize with fluorescent- or chemiluminescent-labelled cDNA or cRNA from a sample of interest. The level of expression of a gene is quantified based on the intensity level detected for the corresponding probe.

FIG. 1.4. A DNA microarray workflow. The gray boxes represent the main steps in the workflow, whereas the white boxes represent work associated with each of the steps.
Since its birth about two decades ago and following the completion of several genome projects, various microarray platforms – a single- and two-color platforms, spotting and bead arrays, etc. – have been developed for various species (e.g. mouse, macaque, and human) and purposes (e.g. expression profiling and detecting SNP, copy number, and alternative splicing). Here, we focus on general approaches in single color microarray experiments for expression profiling and their data preprocessing and analyses. A workflow associated with this type of DNA microarray is illustrated in FIG. 1.4. As shown, the main steps in the work consist of executing the microarray experiments, data preprocessing and analyses, and at last, interpreting the results of the analyses in the context of biological experiments.

In the first main step, the most crucial work is to design a good experiment so that meaningful results obtained. A good experimental design is important for proper statistical analysis. This includes determining the control and treated groups, the time-points for sample collection, and the number of technical and biological replicates. Once samples (cells, tissues, or organs) are collected according to the design of the experiments, the next task is to extract total RNAs from the samples. Following the extraction, the quantity and integrity (the presence of ribosomal RNAs) of the total RNAs need to be confirmed. Next, total RNAs are converted into labelled cDNAs or cRNAs and then used to hybridize the probes in the arrays. To remove excess hybridization solution and reduce non-specific hybridization, the microarray slides are washed. Then, arrays in the chip are scanned to produce high resolution images containing pixels that represent intensity of fluorescence induced by focusing a laser on the array. The images are then processed for intensity data extraction: (i) identify the positions of the probes (spots or beads) on the array, (ii) identify the pixels used as part of the probes, (iii) identify the pixels used for background, and (iv) summarize
the intensity data for each probe, as well as background and some other quality control information. The scanning, image processing and intensity extraction are done automatically. For practicality, intensity data extracted from spots in which image processing detects some types of problems (flagged data) can be simply removed from the data sets.

In the second step, i.e. microarray data preprocessing, intensity data are commonly transformed using logarithmic (log2-based) transformations, which make the distribution of the data symmetrical and variance across arrays comparable. However, such transformation has a drawback where variances are inflated when the intensities are close to 0. To solve this problem, a non-linear variance-stabilizing transform was proposed [131-134]. In relation to the Illumina bead array platform that we used, where each probe is represented by about 30 randomly distributed beads on average, a variance-stabilizing transform (VST) that takes advantage of these replicates has been proposed by Lin et al [134]. Following the transformation, quality assessment is performed by producing density plots and boxplots for each array, pairwise correlation between arrays, pairwise MA plots (plotting the average and difference in log intensity for each probe from two arrays), density plots of coefficient of variance, and examining the sample relations with a clustering technique or principal component analysis. The goal of quality check is to identify systematic differences in expression level across a chip and between chips – which can be due to for example, collecting the samples or hybridizing the chips on different day. The systematic difference, if any, can be resolved by performing normalization of the signal (log transform). Several methods for between-array normalization are available, which include scaling, centering, quantile normalization (distribution normalization), and percentile shift normalization. The later method is popular for
Agilent (another platform that we used for our experiment) and it is usually combined with a baseline transformation, where for each probe, the median of the log transforms from all or control samples is subtracted from each of the samples.

Once all sorts of problems with raw intensity data are handled, the next task is to perform microarray data analyses using standard tools. This includes the identification of differentially expressed genes in the preprocessed microarray data. The most common approach in doing this is by double filtering with (log) fold change and statistical tests. For filtering with fold change (relative to control samples), the lowest threshold for the absolute fold change that is commonly used is 1.5 (absolute log fold change of about 0.6), whereas the filtering with statistical test is commonly done by setting a threshold of 0.05 for the corrected p-value of the test.

Some common statistical tests for identifying significantly differentially expressed genes are the t test and one-way or two-way ANOVA, depending on how the samples are grouped. The t test is used when two groups of samples are compared; one-way ANOVA when a category of grouping with more than two groups are compared; while two-way ANOVA when two categories of grouping are compared. Performing t test or ANOVA requires the justification of a number of necessary assumptions, such as the normality and homoscedasticity of the data. If the assumptions are not justified, the non-parametric statistical tests that make no assumptions about the measurements, e.g. bootstrap analysis, Mann-Whitney-Wilcoxon test, Kruskal-Wallis tests, can be performed. In addition to these standard statistical tests, improvements or different kinds of statistical tests have also been proposed to deal with thousands of genes or specific grouping in microarray data. These include the statistical test designed to deal with time series microarray data, such as the one proposed by Park et al [135] that incorporates a two-stage model, one
for removing time effect and another is for identifying significantly differentially expressed genes.

Next, since the statistical test is performed on thousands of genes, the chance of getting false positives, i.e. random genes that falsely appear to be significantly differentially expressed, increases. Hence, the p-value for each gene needs to be corrected to reduce the number of false positives. The most conventional technique to correct the p-values is Bonferroni correction, which directly controls the family-wise type I error by testing each individual hypothesis at a statistical significance level of what it would be used for a single hypothesis divided by the number of hypotheses. Unfortunately, this approach is too strict since it tends to remove all genes that are truly differentially expressed; in other words, it does not provide a good balance between false positive and false negative rates. To address this problem, different approaches that are based on false discovery rate (FDR), i.e. the expected proportion of false positives among all significant genes, are proposed. The FDR-corrected version of the p-value is often referred to as q-value, which can be interpreted as the smallest FDR for which the gene is considered significant. Different ways of defining FDR (and hence the q-value) have been proposed, which include Benjamini-Hochberg’s and Storey’s methods [136, 137].

Finally, once a set of differentially expressed genes is identified, set-level and system-level analyses can be performed on the gene set. One of the most common analyses involves the Gene Ontology (GO), i.e. a hierarchical structure (in the form of a directed acyclic graph (DAG)) that consists of a controlled vocabulary of terms that describes the cellular component and molecular function of genes and biological process they are involved in. The goal of the analysis is to identify over-represented GO, i.e. GO categories that have a significant overlap with the set of differentially
expressed genes. The significance of an overlap between these two sets of genes is commonly tested by using a test based on hypergeometric distribution or its binomial approximation. Recently, a Bayesian approach that takes into account the GO term dependence structure has also been proposed to identify the over-represented GO categories [138]. In addition to the set-level analysis, analysis of differentially expressed genes can also be performed at system-level, with a goal to identify signaling pathways relevant to the gene set. Similar approaches to GO analysis can be employed for identifying the pathways, but approaches that take into account the dependence of the genes in the pathways and adjust for the correlation structure are preferred [139]. One of the most popular tools to perform this task is the Ingenuity Pathway Analysis (IPA) software, which is commercially available. Alternatively, several free tools are also available, including the InnateDB [140] that are specialized to facilitate system-level analysis of genes that govern the innate immune response.

A scheme for automated cell image analysis systems

Another technology that becomes an attractive and widely used research and diagnostic tools is the automated microscopy systems, which allow a larger sampling of cell images for statistical analysis. One application of such technology is to identify small molecules, RNA, or peptides that result in a specific cell phenotype – an approach popularly known as high content screening (HCS). Another application is for differential counting of leukocytes in a smear of stained blood cells. One of the main components of the automated microscopy systems is the software that automates the analysis of the digital image data, i.e. an automated image analysis system. Here, in relation to our work in automating the estimation of neutrophil purity in giemsa-stained NPRO cells (described in Chapter 4), we outline a scheme for an automated
image analysis system and discussed some image processing methods that can be incorporated into the system.

As illustrated in FIG. 1.5, the workflow of an automated image analysis system (gray boxes) usually includes image preprocessing, image segmentation, calculation of selected features, and object classification using a classifier (machine learning algorithm). Image preprocessing aims at enhancing the interpretability aspects of an image, including image contrast. One popular method to improve image contrast is histogram equalization, which attempts to flatten the gray-level histogram through a gray-level transformation. A variant of histogram equalization, known as contrast-limited adaptive histogram equalization [141], improves image contrast locally and simultaneously prevents noise overamplification.

FIG 1.5. A scheme for automated image analysis system. The gray boxes represent the main steps in automating the analysis of image data, whereas the white boxes represent a prior experiment that needs to be carried out to determine the set of features and classifier’s parameters for image object classification.
Next, image segmentation is performed to isolate image regions associated with the objects of interest. One of the most popular techniques for image segmentation is Otsu’s method, which reduces a gray-level image into a binary image by setting a threshold that divides the pixels into two classes with a minimum intra-class variance [142]. This technique has been extended into a multi-level thresholding and an efficient algorithm has been developed to rapidly compute the thresholds [143]. Another approach for image segmentation is a texture-based segmentation, which employs a multi-dimensional clustering technique to group pixels that have similar texture descriptors.

Following the identification of the objects, a set of values for specified object features can be derived for object classification. These include morphological features, Zernike moments and Haralick features. Morphological features, such as the area and perimeter of objects, solidity (the ratio between area and convex area) and various statistics for gray-level-images, are more sensible than other features. Zernike moments are the projections of a binary image onto an orthogonal basis set consisting of complex polynomials [144]. These projections (the moments) are complex numbers and insensitive to rotation of the image. The magnitudes of the moments, which are spatially and rotationally invariant, are used as moment-based region descriptors. At last, Haralick features formulated in [145] describe some intuitive texture aspects of an image, e.g. image complexity, smoothness, and directionality of the pattern in the image. In addition to these features, wavelet transforms that have been used for image compression, e.g. Daubechies and Gabor wavelet transforms, have also been used to derive wavelet features from images.

Not all features need to be used to classify the objects since some features may be redundant or irrelevant and they can increase the classification error. The most
characterizing features can be selected to minimize the classification error. Several schemes for feature selections are available. The most popular ones are the filter and wrapper methods. Filter methods are less complex than the other methods and select features based on a statistics or other ranking criteria (e.g. T-score, F-score, or minimal-redundancy-maximal-relevance criterion (mRMR)). On the other hand, wrapper methods convolve the feature selector with a classifier to directly minimize the classification error of the corresponding classifier. To search the compact feature subsets, wrapper methods can be conducted using backward (where one feature is excluded one at a time without increasing classification error) or forward selection (where one feature is added if it gives the largest error reduction). In general, filters provide more speed and give a generalized set of features, while wrappers can give high classification accuracy but less generalized selected features [146].

Finally, any type of supervised learning algorithms can be used as a classifier that classifies the objects. The term supervised means that we need to separately train the classifier with a set of training data whose classes are known in order to obtain a set of parameters that is used by the classifier to predict the class of new data. The simplest techniques will be the linear regression model (fit by least squares) and $k$-nearest neighbours ($k$-NN). In the first technique, the training data are used to estimate the coefficient of a linear model by minimizing the sum of squared residuals. For a two-class problem, the linear fit gives a hyperplane that splits the feature space into two sub-spaces that correspond to the two classes. On the other hand, the $k$-NN method predicts the class of a new observation by taking a vote from the $k$ observations in the training data set nearest to the new observation. Thus, whereas linear regression model gives a classifier that fits the training data globally, the $k$-NN classifier fits the training data locally to predict the class of new observations. More
sophisticated approaches for classification problems have been developed, which are mostly the variants of these two simple techniques, including projection pursuit and neural network models.

1.4. Outline of the thesis

This thesis is organized as follows: Chapter 2 presents transcriptomic analyses of lung responses to LVI and HVI infections, and the results of the analyses are linked to histopathologic data of infected lungs. In Chapter 3, we present an automated image analysis system for estimating neutrophil purity from high content images of giemsa stained NPROs. Using both automated and manual giemsa stained approaches, as well as flow cytometry based on neutrophil marker antibodies, we examine the purity of neutrophils in NPROs collected within 6 days of differentiation and determine the optimal purity that can be obtained. In chapter 4, the use of NPROs for modeling IAV infection in neutrophils is assessed, and transcriptomic analyses of NPROs infected with LVI and HVI are subsequently presented. Finally, we summarize our results and discuss future prospects of our research in Chapter 5.
Chapter 2

Transcriptional profiling of lungs infected with highly and less virulent IAV (H3N2) viruses

Several studies have been carried out to uncover the global response of lungs infected with IAVs, mainly of H1N1 and H5N1 subtypes (discussed in Section 1.3.3). In general, the outcomes of these studies suggest that highly virulent IAV infections are characterized by early and sustained innate immune response, whereas the less virulent ones only elicit a delayed and transient activation of this response. Here, we uncover global response of murine lungs infected with highly and less virulent IAVs of H3N2 subtype, i.e. LVI and HVI, respectively, and link the results of the analyses to lung histopathology data. The motivation of this work is given in Section 2.1, followed by the description of materials and methods in Section 2.2. In Section 2.3, histopathology data and the results of microarray data analyses are presented. Finally, we discuss and conclude our results in Section 2.4.

2.1 Motivations

Previously, a mouse-adapted IAV (designated as HVI) was generated from human influenza H3N2 (designated as LVI) through serial lung-to-lung passaging
[13]. HVI causes highly virulent influenza in mice, while LVI only causes less virulent influenza. Most interestingly, HVI has been shown to acquire critical mutations including in the HA and NS1 genes, exhibit significantly higher replication kinetics, and cause extrapulmonary spread to other major organs. Moreover, following HVI infection, neutrophils have been shown to be the majority (about 60%) of cell population observed in bronchoalveolar lavage samples collected at 96 h postinfection [154].

Here, we have been motivated to compare transcriptomic profiles between lungs infected with LVI and HVI by exploiting our mouse adaptation model. Since we were interested in early response to infections, the time-points chosen for sample collection were 12, 48 and 96 h postinfection. Genes associated with innate immune response (especially cytokine/chemokine activity and type I IFN response) as well as apoptosis have been the focus of the investigation. In addition, we have also been motivated to identify novel pathways that explain the pathogenesis of lungs infected with the viruses, especially HVI. One noteworthy advantage of this model, unlike previous transcriptomic models of host response to highly virulent IAV infections [1-6, 195], is that it is free from specific strain effects. This is important because infection with different IAV strains causes different regulation of host genes, such as shown when comparing r1918 H1N1 and H5N1 [195]. Therefore, the most critical host factor associated with the virulence may likely be identified when the model is free from strain effect.

2.2 Materials and methods

Adaptation of human influenza H3N2 virus in mice to increase viral virulence. A mouse-adapted influenza A/Aichi/2/68 (H3N2) virus was previously developed [13].
The original stock of virus from the American Type Culture Collection (Manassas, VA, USA) was adapted in female BALB/c mice (7-8 weeks old) through serial lung-to-lung passaging. The non-adapted virus from the original stock and the adapted virus from the tenth passage are designated as LVI as HVI, respectively. All animal infection experiments were approved by the Institutional Animal Care and Use Committee (protocol 035/07) and were conducted in an animal BSL-2 laboratory facility.

**Mouse inoculation experiments and histopathology.** Female BALB/c mice (7-8 weeks old) were anesthetized with Avertin (Sigma-Aldrich) via an intra-peritoneal injection and inoculated intranasally with 35 \( \mu \)L each of egg allantoic fluid containing \( 1.4 \times 10^5 \) TCID\(_{50}\)/ml of LVI, lung homogenate containing \( 2.9 \times 10^6 \) TCID\(_{50}\)/mL of HVI, or uninfected lung homogenate as negative control. Mice were housed in individually ventilated cages maintained at 21°C with 12-hour light-dark cycles. The lungs of control mice and those infected with LVI and HVI were harvested at 12, 48 and 96 h postinfection, fixed in formalin, sectioned, stained with hematoxylin and eosin, and observed under light microscopy.

**Total RNA isolation from infected and mock-infected murine lungs.** For each experimental group (control, LVI, HVI), total RNAs were extracted from the lungs of three individual mice at 12, 48, and 96 h postinfection using RNeasy Plus Mini kit (Qiagen, Hilden, Germany). Mortar and pestle were used to disrupt lung tissue in liquid nitrogen, and the tissue powder was added to RLT lysis buffer (600 \( \mu \)L). The lysate was centrifuged at 12,000×g for 3 min, and the supernatant was subsequently centrifuged using a gDNA Eliminator spin column at 8,000×g for 0.5 min to remove
genomic DNA. One volume of 70% ethanol was added to the flow-through and mixed by pipetting. The mixture (700 μL) was transferred to the RNA-binding RNeasy spin column and centrifuged at 8,000×g for 0.5 min. The flow-through was discarded, and the mixture was washed using 700 μL of RW1 buffer once and 500 μL of RPE buffer twice. The RNeasy spin column was centrifuged at 12,000×g for 1 min to ensure thorough drying and removal of any residual buffer. The RNAs were then eluted in RNase-free water and stored at −80°C for microarray experiments. The isolated total RNAs were quantified by UV spectrophotometry at 260 nm using a Nanodrop ND 1000 spectrophotometer. Purity was ensured by an OD260/OD280 ratio of >1.8. The integrity of the RNAs was verified by electrophoresis in a 1.0% native agarose gel prepared using DEPC-treated 1× tris-acetate-EDTA (TAE) buffer, with the equipment washed with DEPC-treated water.

**Oligonucleotide microarray experiments.** Samples for cRNA hybridization onto single-channel MouseRef-8 v1.1 BeadChip microarrays (consisting of ~25,000 70mer probe sequences, ~30 replicates for each probe) were prepared using a Totalprep RNA Amplification kit (Illumina, San Diego, CA). Initially, equal amounts of high-quality RNAs from three mice in each treatment group were pooled in order to reduce the variability due to genetic difference between mice. Next, 9 μl of the RT Master Mix containing T7 Oligo(dT) primers and RNase inhibitors were added to each pool of total RNAs (500 ng in 11 μl) to synthesize the first strand cDNAs and then incubated at 42°C for 2 h. Subsequently, the second strand cDNAs were synthesized to form dsDNA templates for transcription. To degrade the RNA, 80 μl of the Second Strand Master Mix containing DNA Polymerase and RNase H were added to the sample and incubated at 16°C for 2 h. Purification of cDNAs was then performed.
using a cDNA filter cartridge system which included a series of washes and a final elution of cDNAs using 19 μl of 55°C nuclease-free water. This step removed the RNAs, primers, enzymes, and salts that may inhibit the subsequent in vitro transcription (IVT). IVT for generating multiple copies of biotinylated cRNAs was carried out by adding 7.5 μl of IVT Master Mix to the sample and incubated at 37°C for 14 h. Then, 75 μl of nuclease-free water were added to stop the reaction and the sample was purified in a cRNA filter cartridge using cRNA binding buffer, wash buffer, and 100% ethanol. The cRNAs from each sample were eluted using 100 μl of 55°C nuclease-free water. Subsequently, 5 μl of the nuclease-free water (~750 ng of cRNAs) were mixed with 10 μl of GEX-HYB hybridization buffer. The mixture was preheated to 65°C for 5 minutes and then cooled to room temperature. Finally, cRNAs from each experiment were loaded onto one array in the microarray slide placed in a hybridization chamber filled up with 200 μl of GEX-HCB humidifying buffer. The microarray slides were incubated in a rocker at 58°C for 17 h to facilitate hybridization.

**Microarray data preprocessing and analyses.** Global gene expression in total lungs of mock, LVI and HVI experiments were collected by using Illumina single-channel MouseRef-8 v1.1 BeadChips. Illumina BeadArray Reader was used to scan the chips and produce TIFF images which were subsequently converted into summary intensity data by using Bead Summary package. The summary includes the minimum, maximum, average and standard deviation of signal intensity and the p-value of signal detection. Most of the analyses were done using Bioconductor packages and other packages in the R software library (http://www.r-project.org/). First, probes that have p-value of detection less than 0.025 were selected. Then, lumi package [147] was used
to perform variance-stabilizing transformation and robust spline normalization on the selected probe intensity data. The algorithm implemented in the package takes advantage of technical replicates for each probe in order to handle the dependency of variance on intensity level (a higher intensity gives a higher variance). The new value associated with each probe is originally called VST; however, since the value is close to the log transformed value and to provide no further confusion for most readers, we just refer it as a log intensity value. Next, RIKEN probes were converted to the relevant gene symbol if they were available in the MGI RIKEN provided by the Mouse Genome Informatics (MGI) database and the expression value for each gene was the average of log intensities across relevant probes. This gave 12,181 genes which were then analyzed by using two-way ANOVA to identify genes that had a significant infection effect, time effect and interaction effect. If a gene had no significant interaction effect, then we employed two-way ANOVA without considering the interaction effect to improve the p-value for infection and time effect. The Storey method for multiple tests correction [137] implemented in R software, i.e. qvalue package, was used to adjust p-values of the ANOVA test for each gene. Functional analysis of any particular gene set was performed by using FuncAssociate 2.0 [148].

2.3 Results

2.3.1 Murine lungs infected with LVI and HVI display distinct pathologic features

The lungs of control mice and of animals infected with LVI and HVI were harvested at 12, 48 and 96 h postinfection and subjected to histopathology analyses. Lungs
inoculated with mock control did not exhibit pathologic features at 12 h postinfection, but mild alveolitis was observed at 96 h postinfection (data not shown). As illustrated in FIG. 2.1, lungs infected with either LVI or HVI exhibited minimal pathologic changes at 12 h postinfection. At 48 h postinfection, LVI infection led to alveolitis with mild infiltration of inflammatory cells, whereas HVI infection caused more severe damage as evidenced by bronchopneumonia and emphysematous changes. LVI-infected lungs exhibited recovery at 96 h postinfection, as indicated by the clearance of cellular infiltrate and disappearance of alveolitis. In contrast, HVI-infected lungs at 96 h postinfection were progressively worsening, characterized by massive leukocyte infiltration (predominantly with neutrophils and macrophages), necrotizing bronchitis, hemorrhagic edema, and accumulation of proteinaceous materials in alveolar space.

FIG. 2.1. Histopathology of mouse lungs infected with LVI or HVI at 12, 48, and 96 h postinfection (p.i.). Sections of lung tissues were stained with H&E. At 12 h p.i., both LVI- and HVI-infected lungs exhibited normal morphology. At 48 h p.i., LVI-infected lungs revealed alveolitis with mild infiltration of inflammatory cells, whereas HVI-infected lungs displayed a more severe pathology with bronchopneumonia and emphysema. At 96 h p.i., LVI-infected lungs exhibited recovery, while HVI-infected lungs displayed severe bronchopneumonia, massive leukocyte infiltration, necrotizing bronchitis, hemorrhage and accumulation of proteinaceous materials.
2.3.2 Transcriptomic data analyses and kinetics of LVI versus HVI infections

To identify the alteration of host pulmonary transcriptome associated with the increased virulence of IAV during adaptation, gene expression data were collected from the lungs of three individual mice at 12, 48 and 96 h postinfection for each infection group (mock, LVI and HVI). The expression data were then preprocessed and analyzed by using two-way ANOVA with correction for multiple tests. After multiple two-way ANOVA tests, using a significance level of 0.05 of q-value (adjusted p-value), we identified 599 genes that showed a significant interaction between infection and time, 3,515 genes that exhibited a significant infection effect, and 10,937 genes that had a significant time effect. The overlap between these significant groups is shown in FIG. 2.2.A. Most of the genes with significant infection effect also had a significant time effect and/or interaction effect, which could be expected of the dynamics of gene interactions and the changes of cell populations in the lungs during the infection process. In addition, a large number of genes (7,589) had a significant time effect without a significant infection effect. These genes are expected to be those showing common response to mock, LVI and HVI infection, and differential expression across time-points may be due to non-infectious factors.

Next, we focus on the genes that revealed a significant infection effect, in particular genes that showed an absolute log ratio of intensity (difference between log expression value in LVI-/HVI- and mock-infected group) greater than a threshold of 0.6 in any infected group and at any time-point. This filtering presumes that the effect of the viral and non-viral components to the host gene expression are additive. The threshold of 0.6 was chosen in a way such that the percentage of filtered genes with a significant infection effect was kept high while not too many genes in the array showed a log ratio close to the threshold at any combination of infection and time.
what follows, our analyses and discussions are mainly limited to these filtered genes, which are the differentially expressed genes in LVI and/or HVI.

FIG. 2.2B shows the expression profiles of 1,195 differentially expressed genes identified in the arrays. Overall, the differences between gene expression patterns in LVI and HVI at 48 and 96 h postinfection were more conspicuous compared to that at 12 h postinfection, with a much larger number of over-expressed and under-expressed genes being observed in HVI. The overlap between filtered genes in LVI and HVI (FIG. 2.2.C) provides a quantitative view of common and specific differentially expressed genes in LVI and HVI. At 12 h postinfection, relatively few genes in LVI and HVI (i.e. 36 and 22 genes, respectively) demonstrated log ratio intensities above threshold. Surprisingly, only five genes were common in LVI and HVI at this time-point, and most of the genes in LVI were under-expressed while genes in HVI were over-expressed.

In spite of the differences observed at 12 h postinfection, the set of 88 differentially expressed genes in LVI at 48 h postinfection was only a subset of differentially expressed genes in HVI. These genes were all over-expressed, with the changes in expression levels being consistently higher in HVI. In addition, 573 genes were differentially expressed in HVI, with 168 of them under-expressed. At 96 h postinfection, almost all of the differentially expressed genes in LVI were also the differentially expressed genes in HVI and the number of differentially expressed genes in each of LVI and HVI was more than twice compared to that at 48 h postinfection. For LVI, 182 genes were over-expressed and 24 genes were under-expressed. For HVI, 601 were over-expressed and 462 were under-expressed. Overall, more than 75% (i.e. 77% for LVI and 80% for HVI) of differentially expressed genes
at 48 h postinfection were continually over-expressed or under-expressed at 96 h postinfection.

**FIG. 2.2.** Summary of transcriptional response of murine lungs infected with LVI and HVI. (A) Venn diagram showing the overlap between significant groups of genes as detected by two-way ANOVA analysis with Storey correction for multiple tests (q-value <0.05). (B) Heatmap of 1,195 genes that had a significant infection effect and an absolute log ratio >0.6 at any time-point of any treatment. (C) Venn diagram showing the summary of the overlap of differentially expressed genes in LVI (dark-gray circles) and HVI (light-gray circles) that had an absolute log ratio >0.6 at 12, 48 and 96 h postinfection (p.i.).
2.3.3 Early inflammatory and immune responses of LVI versus HVI infections

Only a few genes were found to be differentially expressed in LVI and HVI gene sets at 12 h postinfection. The overlapping genes included ALDH3A1, IL1B, RETNLG, SAA3, and S100A9. Interestingly, only two of them were differentially expressed in the same manner, i.e. over-expression of SAA3 and under-expression of ALDH3A1. A notable observation was the depression of proinflammatory IL1B, which is normally expected to be over-expressed early after infection such as observed in HVI.

To further investigate the basis of early responses in LVI and HVI at 12 h postinfection, gene ontology (GO) analysis was performed with FuncAssociate 2.0. The GO category of inflammatory response was found to be over-represented in the HVI gene set, which contains over-expressed proinflammatory genes including CCL7, CXCL1, IL1B, ORM2 and SAA3. On the other hand, the LVI gene set was strikingly over-represented by under-expressed chemotactic, including CCL5, CCL6, CORO1A, EAR2, FCGR3, IL1B, RAC2, S100A8, and S100A9. In addition, genes associated with the inflammatory response were also over-represented in LVI. However, some of these genes (i.e. SAA3 and STAB1) were over-expressed, while others (i.e. CCL5, FGCR3 and IL1B) were under-expressed. Overall, our results suggest that immune responses were activated at 12 h postinfection in LVI and HVI, but to varying extents.

2.3.4. HVI infection is characterized by hypercytokinemia and dysregulated TREM1 signaling

GO category for cytokine activity was over-represented in both LVI and HVI at 48 and 96 h postinfection. The heatmap of differentially expressed genes associated with
cytokine activity (FIG. 2.3) indicated that cytokines were highly active in HVI at 48 h postinfection, although the infiltration of immune cells was still relatively mild according to histopathology. At 96 h postinfection, the transcripts of cytokine genes were highly expressed, thus implying a hypercytokinemia or cytokine storm during HVI infection. The most interesting cytokine was the neutrophil chemoattractant CXCL1, which was rapidly over-expressed even at 12 h postinfection in HVI but not LVI, suggesting a major role of neutrophils in early response to HVI infection. On the other hand, the number of cytokine genes and their expression levels were much lower in LVI. Nevertheless, the heatmap alluded that CCL4, CCL7, CCL12, CXCL1, CXCL9, CXCL13 and IL1B (the seven genes at the bottom of the heatmap) play more functional roles during LVI and HVI infections compared to other cytokines.

The development of the hypercytokinemia in HVI was potentially caused by the dysregulation of TREM1 signaling pathway, which is known to amplify the signal from TLRs and NLRs, including the production of various cytokines and chemokines in macrophages and neutrophils (discussed in Section 1.3.4). This pathway was discovered by IPA software to be one of the significant canonical pathways throughout the observed time-points in both LVI and HVI. The major elements of TREM1 signaling pathway are TREM1 protein and its adaptor molecule TYROBP, and other regulatory molecules that include TLR2, TLR7, JAK2, STAT3 and NF-κB2. The changes of the expression levels of these genes are shown in FIG. 2.4A (note that the q-value of NF-κB2 was significant, but its expression changes were not above threshold; it is included here due to its importance). As shown in the figure, more intense over-expression of TREM1 elements was observed in HVI compared to that in LVI.
FIG. 2.3. Overview of cytokine gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.

Of particular interest, the transcript level of TREM1 and TYROBP increased rapidly in HVI, but at a lower rate in LVI. Notably, the transcript level of TYROBP decreased significantly in LVI at 12 h postinfection before it was over-expressed (but not above threshold) at 48 and 96 h postinfection. In addition, TLR2, TLR7 and JAK2
interestingly displayed expression patterns similar to that of TYROBP. Furthermore, the expression level of transcription factors associated with TREM1 pathway (i.e. STAT3 for neutrophils, NF-κB2 for neutrophils and macrophages) were also generally higher in HVI than in LVI. Since STAT3 expression was more evident than NF-κB2, the modulation of TREM1 signaling is likely to occur more intensively in neutrophils rather than in macrophages during HVI virus infection. To summarize the role of TREM1 signaling in controlling the level of cytokines and chemokines, a relevant subnetwork is highlighted in FIG. 2.4B.

2.3.5. Trends of type I IFN-related and apoptotic gene expression

Here, we examine transcriptional profiles for genes involved in type I IFN response and apoptotic activity. The profile for type I IFN-related genes reflects the antiviral activity, while the profile for apoptosis related genes may reflect the pathological state during infection. As shown in FIG. 2.5 and 2.6, both transcriptional profiles of type I IFN-related genes and apoptotic genes were similar between LVI and HVI in terms of the progression of their expression but were different with respect to their fold changes.

Moreover, some of the type I IFN-related genes, including H2-D1 and H2-L, were interestingly under-expressed in LVI at 12 h postinfection. On the other hand, there were no obvious changes in their expression in HVI at that time-point except for the antiviral ISG15. At subsequent time-points, the expression levels of type I IFN-related genes in HVI at 48 h postinfection were higher than those at 96 h postinfection, while for LVI, their corresponding expression levels were generally similar. Overall, type I IFN-induced antiviral response to HVI was rapidly activated and more intense than to LVI.
FIG. 2.4. Differential regulation of TREM1 signaling in LVI and HVI. (A) Dynamic expression profiles of selected genes involved in the TREM1 signaling pathway in LVI and HVI. (B) A proposed gene sub-network implicated in the development of hypercytokinemia. Genes in the sub-network had a significant infection effect (q-value <0.05), and exhibited an absolute log ratio of >0.6 in at least one time-point after LVI or HVI infection. Direct interactions and indirect interactions are represented by solid lines and dashed lines, respectively. Gray nodes represent transiently under-expressed genes in LVI, while white nodes represent over-expressed genes at all time-points in both LVI and HVI.
FIG. 2.5. Overview of type I IFN-related gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.

For apoptotic genes, insignificant changes were observed at 12 h postinfection for both LVI and HVI. However, at 48 and 96 h postinfection, enhanced expression of apoptotic genes were observed at much higher levels in HVI than in LVI. In addition, we also observed that expression of pro-apoptotic genes (e.g. CASP1, CSRNP1, DAXX, GADD45G, IFNG and TNF) was higher at 48 h postinfection than that at 96
h postinfection, while anti-apoptotic genes (e.g. IL6, SOD2 and TNFAIP3) were expressed at higher levels later in time. Overall, over-expression of pro-apoptotic genes at early phase of HVI infection dominates over-expression anti-apoptotic genes, indicating that cellular activity is likely towards cell death rather than cell survival. Thus, enhanced apoptotic activity in HVI mirror the alveolar and bronchiolar damage observed in histopathology data.

**FIG. 2.6.** Overview of apoptotic gene expression patterns at 12, 48 and 96 h postinfection (p.i.) with LVI and HVI.
2.3.6. Under-expressed genes in HVI infected lungs are associated with metabolism and developmental processes

Unlike LVI, hundreds of genes were under-expressed in HVI at 96 h postinfection. This motivated further investigation to uncover pathways associated with this trend. In HVI at 48 and 96 h postinfection, IPA analyses on under-expressed genes revealed an over-representation of pathways associated with metabolism of small molecules such as xenobiotics, carbohydrates, amino acids, lipids and neurotransmitters. Key players in these pathways included the family of cytochrome p450 (CYP), glutathione s-transferase (GST) and aldehyde dehydrogenase (NAD), which are regulated by the nuclear hormone receptor RXR\(\alpha\). Under-expression of CYP, GST and NAD may be related to the cytokine activity via a mechanism referred to as LPS/IL1-mediated inhibition of RXR function. In this case, the deactivation of this pathway may be stimulated by excessive expression of IL1B.

A significant number of under-expressed genes was also identified to be associated with cellular development, growth and proliferation, including fibroblast growth factor 1 (FGF1), vascular endothelial growth factor A (VEGFA), and NOTCH signaling pathway genes, NOTCH3 and NOTCH4. In particular, expression of NOTCH3 and NOTCH4 gradually decreased until 96 h postinfection. Overall, these observations suggest that metabolic impairment in the lungs and the lack of tissue repair may contribute to lung damage during HVI infection. A sub-network that highlights the interactions between genes within metabolic and developmental processes is shown in FIG. 2.7. Interestingly, Two circadian genes, albumin D box-binding protein (DBP) and nuclear receptor subfamily 1 group D member 1 (NR1D1), were also included in the network since these two genes were interestingly under-
expressed even at 12 h postinfection, which may signify the initial loss of regular process in the lungs early after HVI infection.

**FIG. 2.7.** A gene sub-network highlighting the relationships between under-expressed genes in HVI. Genes in the sub-network had a significant infection effect (q-value < 0.05) and showed a log ratio ≈-0.6 at 96 h post-infection. Gray nodes represent genes associated with metabolism of small molecules, while white nodes represent genes associated with developmental processes.

### 2.4 Discussion

Many factors contribute to the virulence of IAV infection. Antigenic changes and key mutations that increase viral replication rate in epithelial cells are well associated with enhanced virulence. The ability of highly virulent IAV to modulate an aberrant innate immune response is also recognized as another factor contributing to heightened pathogenicity, including excessive production of cytokines and chemokines. The
modulation of innate immunity is corroborated with flow cytometry data that showed a significant increase of neutrophils and macrophages in lungs infected with a highly virulent IAV compared to a less virulent one [14]. Furthermore, macrophages infected with IAV may contribute to the hyperinduction of innate immune responses, including type I IFN and cytokine responses. Specifically, in vitro experiments by Lee et al [9] demonstrated that a highly pathogenic IAV could induce much higher expression of type I IFN and cytokine related genes compared to a less virulent one. On the other hand, the role of neutrophils in modulating innate immune response has not been investigated in vitro. However, IAV infection in neutrophils is well known to activate an atypical respiratory burst and depress neutrophil functions (as described previously in Section 1.3.5). Very recently, in vivo transcriptomic responses elicited in neutrophils and other cell types during lethal and non-lethal IAV infection in a mouse model have also been uncovered by Brandes et al [204].

In the case of HVI, we have demonstrated that the increased virulence of the virus is partially attributed to the increased replication rate of the virus in vivo and in vitro. HVI exhibits higher replication potential compared to LVI from 12 to 60 h postinfection in vitro [13]. Furthermore, our results are consistent with excessive recruitment of neutrophils and macrophages in HVI-infected animals as observed by lung histopathology. Specifically, we observed over-expression of neutrophil chemoattractant CXCL1 and macrophage chemoattractant CCL7 at much higher levels in HVI than in LVI. Of particular interest is the far greater over-expression of CXCL1 in HVI even at the earlier time-point of 12 h. Unlike the observations for HVI infection, decreased cellular infiltration and a more subdued inflammatory response were the major characteristics of LVI infection early after infection. Specifically, we observed a temporary repression of some genes associated with leukocyte chemotaxis.
It may be argued that the transient depression of chemotactic genes is beneficial to the host since it delay inflammatory and immune responses. Consequently, such defects avoid excessive infiltration of innate immune cells that may contribute to damage of infected lungs. The most interesting gene is IL1B which was under-expressed in LVI but was over-expressed in HVI early after infection. This cytokine is a key mediator of the inflammatory response and is mainly produced by monocytes or macrophages [149, 150]. In these microarray studies, under-expression of IL1B after LVI infection was observed. Interestingly, IL1B under-expression is also associated with the recovery of macaques infected with highly virulent H5N1 virus [6] and highlights the importance of transcriptional regulation of the IL1B network in determining the outcome of IAV infection.

In addition to key proinflammatory responses, we also investigated critical signaling pathways associated with response to HVI and LVI. Specifically, we observed differential regulation of the TREM1 signaling pathway in HVI versus LVI at early time-points. TREM1 and TREM-like receptors are potent regulators of inflammation via the integration of inflammatory signals with leukocyte adhesion molecules [151]. Triggering via TREM1 activates the transcription by JAK2-STAT3 (in neutrophils) and by NF-κB (in neutrophils and monocytes), and culminates in the production of proinflammatory cytokines, chemokines, reactive oxygen species, degranulation of neutrophil granules, and phagocytosis [111]. Based on the expression patterns of JAK2-STAT3 and NF-κB, we suggest that TREM1 signaling is likely to be more dominant in neutrophils rather than in macrophages during the time frame of our experiments. Over-expression of TREM1 signaling components in neutrophils upon HVI infection may be implicated in hypercytokinemia, while under-expression of some elements of the TREM1 signaling pathway (e.g., TYROBP, TLR2, and
TLR7) in LVI may be associated with defective neutrophil activity. Certain IAVs have been shown to induce the expression of some TLRs, including TLR2 and TLR7 [119, 124]. TREM1 and TLRs can synergistically activate neutrophils and induce the oxidative burst [152]. Hence, further investigations are warranted to elucidate IAV-induced TREM1 signaling in neutrophils and to explore the therapeutic potential of modulation of TREM1 signaling [153].

The roles of neutrophils and macrophages during highly virulent influenza infections are not fully elucidated. Depletion of neutrophils and/or macrophages in mice infected with highly virulent IAVs may culminate in uncontrolled virus replication and spread in the lungs [66, 69]. Thus, the recruitment of these inflammatory cells is important to clear and contain the virus. Neutrophils are also important in ameliorating lung injury by reducing pathogenic effects on vascular permeability and pulmonary inflammation [68], while macrophages are important for inducing epithelial lung repair by producing hepatocyte growth factor [154]. However, both neutrophils and macrophages are susceptible to IAV infection. Infection of neutrophils may lead to their activation or ultimately cell death [155]. As neutrophils become activated, the release of cytokines and free radicals contribute to lung damage. Moreover, comparison between pathological manifestations in models of neutrophil- and macrophage-depleted mice infected with sub-lethal doses of PR8 H1N1 indicated the more important role of neutrophils rather than macrophages in acute lung injury during highly virulent influenza infection [71]. A better understanding of how neutrophil activation and death mechanisms influence the infection outcome may lead to novel strategies for intervention against severe influenza.
Chapter 2

Our transcriptomic studies suggest that metabolic impairment and lack of lung repair constitute major factors that promote lung damage. Pathway analyses of under-expressed genes in HVI revealed possible mechanisms by which cytokine dysregulation may cause impaired metabolism through inhibition of RXRα function to express metabolic genes such as CYP, GST, and NAD. The model has been well-established in hepatic cells in relation to liver cholestasis. Kim et al. [156] showed that IL1B treatment could decrease the transcript and protein of the nuclear hormone receptor RXRα which regulates CYP, GST, and NAD in vitro. Moreover, IL1B also activates JNK-mediated, CRM1-dependent nuclear export of RXRα [157]. A similar process may underlie under-expression of metabolic genes. In addition to metabolic derangement, we observed the host’s impaired ability to repair lung damage during HVI infection. Many genes encoding growth factors associated with lung remodeling were under-expressed, including FGF1 which is important in lung epithelial cell proliferation [158] and VEGFA which is crucial for embryonic lung development [159, 160] and promoting lymphangiogenesis during lung development [161]. Moreover, some elements of NOTCH signaling which is important in deciding cell fate in the formation and development of various organs [162] were also under-expressed.

Next, to uncover whether differential regulation of pathways (subnetworks) and biological processes inferred from our microarray data underlie general differences in murine global pulmonary responses to highly versus less virulent (lethal versus non-lethal) IAV infections, we performed a meta-analysis involving our microarray data and other relevant microarray datasets downloaded from the Gene Expression Omnibus (GEO) database. The procedures and the outcomes of the meta-analysis are given in Appendix A. Firstly, as shown in FIG. A1 in the appendix, we
can see that the cytokine activity, IFN response and apoptotic activity in the lungs infected with highly virulent IAVs (HVI (H3N2), PR8 (H1N1), r1918 (H1N1) and VN1203 (H5N1)) are generally more intense than those infected with less virulent IAVs (LVI (H3N2), CA04 (2009 pandemic H1N1, which is non-pathogenic in mouse models) and TX91 (H1N1)). Secondly, as shown in FIG. A2, the expression profiles of the TREM1 subnetwork and the subnetworks associated with metabolic and developmental processes in highly versus less virulent infections are also distinctive. Most interestingly, we could observe the over-expression of signaling molecules in TREM1 subnetworks, especially TYROBP and TLR2, which were much more pronounced during highly virulent IAV infections. Nevertheless, few genes did clearly show opposite expression pattern within the virulence group. These include ALDH3A1, which was under-expressed early after HVI and PR8 infections, but over-expressed in r1918 and VN1203. In addition, we also observed that CA04 seemed to be able to induce considerably high levels of cytokine and IFN related genes on day 3 postinfection although apoptotic activity, up-regulation of TREM1 signaling, and down-regulation of metabolic and developmental processes were less evident. Overall, despite few differences at individual gene levels across the infections that might be due to the difference between the microarray platforms used for the experiments, the results of the meta-analysis suggest that the regulation of the genes in the subnetworks and biological processes identified from our LVI and HVI models could determine the outcome of IAV infections.

Following the meta-analysis, we propose a set of candidate signature genes whose expression may serve to differentiate highly virulent influenza from less virulent one, particularly when viral load is reaching its peak on 3 to 4 days postinfection. A heatmap that summarizes the expression of the candidate signature
genes on those days is given in FIG. A3 in the appendix. Fifteen genes were selected to indicate the activity in TREM1 and metabolic/developmental subnetworks, as well as the levels of cytokine activity, IFN response and apoptotic activity.

Firstly, viral load may be related to the level of the IFN-inducible ISG15, which was highly expressed even at the early time-point in HVI. ISG15, a ubiquitin-like modifier, is a crucial antiviral protein against various pathogens including IAV. Interestingly, ISG15 was identified as a gene with the greatest discriminating power to distinguish symptomatic from asymptomatic subjects with influenza [163].

Secondly, IL1B expression may be indicative of the progression of the disease state, since under-expression of this gene is associated with LVI infection, as well as with host survival after highly pathogenic H5N1 infection in a macaque model [6].

To address the recruitment of neutrophils and macrophages during highly virulent influenza infection, neutrophil chemoattractant CXCL1 and macrophage chemoattractant CCL7 are noteworthy. CXCL10, CCL4 and CCL5, which are a chemoattractant for macrophages, natural killer cells and a variety of other immune cells, may also potentially serve as candidate signature genes. Of main interest is CXCL10, an interferon inducible protein that has been shown to be significantly elevated during H5N1 infection in a ferret model [4]. CXCL10 possesses T-helper 1 (Th1) chemotactic properties, and patients suffering from severe infection with the new pandemic variant of IAV secrete significantly high levels of Th1 cytokines [83]. Furthermore, with respect to the TREM1 signaling that could amplify the production of the chemoattractants, TYROBP and TLR2 expression should be taken into consideration at the final outcome to reflect total neutrophil activation and inactivation.
Next, highly virulent influenza infection is commonly accompanied by increased apoptotic activity, and the expression level of growth arrest and DNA-damage-inducible 45 gamma (GADD45G) may serve as a quantitative measurement as previously proposed [1]. In addition, the expression of two apoptosis-related cysteine peptidases, i.e. CASP1 and CASP4, is worth being taken into account. In addition to their roles in apoptosis, CASP1 and CASP4 also play important roles in inflammation. CASP1 is known for its ability to proteolytically cleave and activate the inactive precursor of IL1B, whereas CASP4 (also known as CASP11) is required for CASP1 activation [164].

In relation to defect in metabolic and developmental processes, we consider the expression of ALDH2, SCGB3A1 and DBP. ALDH2 is well studied in association with the metabolism of alcohol. It is an enzyme that detoxifies aldehydes to carboxylic acids, and its absence could increase oxidative stress [165]. The secretoglobin SCGB3A1, whose function is not well characterized, is predominantly expressed in the lung and trachea [166]. At last, DBP, which was first identified as a circadian gene in rat hepatocytes [167], was under-expressed early after HVI infection, and hence, it may reflect the commencement of homeostatic impairment in the lung following the infections.
Chapter 3

High content image approach for estimating the purity of neutrophils in giemsa-stained NPROs

Following the outcome of the genomic analysis in Chapter 2, we put interest in uncovering the global response of neutrophils to IAV infection. For this, we focused on in vitro model based on neutrophils differentiated from MPRO cells, which are designated as NPROs. Here, we propose and evaluate an automated image analysis system for estimating the purity of neutrophils in high content images of giemsa-stained NPROs. The performance of the automated system was evaluated by comparing its estimates with estimates from manual differential cell counting of giemsa-stained NPROs. In addition, flow cytometry based on specific and non-specific neutrophil marker antibodies were also employed to validate the estimations by giemsa-staining approach. In particular, we utilized these approaches to uncover the optimal neutrophil purity in NPROs within 6 days of culturing with excess versus limited supply of medium (or, in other words, at low versus high concentration of cells, respectively). The outcome on the purity analyses in this chapter has been the basis for carrying out the experiments related to modeling IAV infection in neutrophils, which are discussed in the next chapter.
This chapter is organized as follows: Section 3.1 presents the motivation of this work. Section 3.2 describes materials and methods that include the image acquisition and segmentation, feature extraction and selection methods, and machine learning approach for classifying cell images. Section 3.3 provides the evaluation of image analysis and machine learning approaches that were incorporated in the proposed automated image analysis system, followed by their application to uncover the trends of neutrophil yields in NPROs within 6 days of differentiation with excess versus limited supply of fresh medium. A comparison between the trends output by the automated system and the trends output by manual counting and flow cytometry approaches is presented. Finally, Section 3.4 discusses and interprets the trends and the performance of the automated image analysis system.

3.1. Motivations

NPROs, which refer to neutrophils from induced differentiated MPRO cells, have been shown to be a potential source of neutrophils for in vitro experiments. In particular, NPROs have similar properties to normal neutrophils in terms of chemotaxis, respiratory burst, phagocytosis, and ability to release neutrophil extracellular traps (NETs) [15, 17]. Hence, NPROs are likely adequate to model interaction between neutrophils and pathogens, including bacteria and viruses. Nevertheless, one problem in using NPROs is that the cells contain not only mature neutrophils, but also immature neutrophils or neutrophil precursors. Regrettably, no information about neutrophil content in NPROs is available until now, despite the importance of this information in confirming the reliability of NPROs as a neutrophil source. So far, only cells that undergo a granulocytic differentiation have been quantified from giemsa-stained NPROs [16].
Here, we set a goal to quantify neutrophil yields in NPROs and at the same time, determined the maximum neutrophil yield that can be obtained during NPRO maintenance. Specifically, we were interested in investigating whether the volume of fresh medium added to NPROs during the maintenance affects the outcome of the differentiation. We hypothesized that with limited addition of fresh medium, more neutrophils could be obtained since MPRO cells may likely to differentiate into neutrophils or die rather than undergo cell division.

For neutrophil quantification, two approaches are available: differential counting of giemsa-stained cells and flow cytometry. The cost for preparing giemsa-stained NPROs is much cheaper than preparing samples for flow cytometry since it merely requires a drop of cheap chemical dye solution, rather than expensive neutrophil marker antibodies. Furthermore, giemsa-stained NPROs can be prepared in 10 minutes, while the antibody staining procedure for flow cytometry requires at least 30 minutes. However, thousands of cells can be inspected within seconds by a flow machine, while only a few hundreds of giemsa-stained cells are reasonably inspected manually within 10 minutes by an experienced lab scientist.

To overcome the drawbacks of the manual approach, the use of an automated image analysis system that analyzes high content images of giemsa-stained NPROs provides an attractive alternative. The automated system permits a higher number of cells to be counted and eliminates the laborious work of manually counting when handling a large number of samples. Some sophisticated automated image analysis systems have been available for differential blood cell counting in a smear of giemsa-stained blood cells. However, their performance has been shown to be dependent on the source of the blood samples [185]. In addition, the systems have been specifically developed to analyze a smear of a human blood sample. Hence, different algorithms
and training data sets may be required to deal with NPROs, which are cells of mouse origin and contain different mixtures of cells. Overall, these motivated us to develop and assess an automated image analysis system to quantify neutrophil yields in high content images of giemsa-stained NPROs, taken at varying magnification.

Additionally, we were also motivated to quantify neutrophils by using flow cytometry based on specific and non-specific neutrophil marker antibodies, i.e. anti-Ly-6G and anti-Gr-1, respectively. In addition to confirming the presence of neutrophil markers in NPROs, these quantifications are useful to validate the estimations based on giemsa-stained NPROs.

3.2. Materials and methods

**MPRO culture and differentiation into NPROs.** MPRO cell lines (ATCC Number CRL-11422) were cultured in IMDM containing 4 mM L-glutamine, 3 g/L sodium bicarbonate, 10 ng/mL murine GM-CSF, and 20% heat-inactivated FBS at 37°C with 5% CO₂. The culture medium was replenished in 1-2 days to bring cell concentration into 5×10⁵ cells per ml. To induce the differentiation of MPRO cells into NPROs, the growth medium was supplemented with 10 μM ATRA. Certain volumes of fresh medium for differentiation were added at specific days during differentiation to maintain the NPROs. Specifically, we maintained NPROs at low or high concentration of cells, i.e. with an excess amount of nutrients (we added a volume equal to initial volume (1V) on day 2 after differentiation, 2V on day 3, and another 1V on each day 4 and 5 to keep the cell concentration between 1-2×10⁶ cells per ml) or, alternatively, a limited amount of nutrients (1.5V on day 3 and 0.5V on each day 4...
and 5, leading to cell concentration between $3-5\times10^6$ cells per ml). The viability of NPROs was determined based on trypan blue exclusion.

**Flow cytometry.** To estimate neutrophil purity with flow cytometry, $1\times10^6$ MPRO or NPRO cells were washed with $1\times$ PBS before and after 20 minutes incubation at 4°C with 100 µl of plain stain buffer or stain buffer containing 1:50 dilution of anti-mouse PerCP-CyTM5.5 rat anti-mouse Ly-6G (BD Pharmingen) or FITC anti-mouse Gr-1 (MACS, Miltenyi Biotec), fixed and resuspended in sheath buffer. Flow analysis was performed by using Cytomics FC 500 Series Flow Cytometry System (Beckman Coulter) and data were analyzed by using CXP software version 1.0 or WinMDI version 2.9 (Bio-Soft NET, [http://en.bio-soft.net](http://en.bio-soft.net)). The threshold for singly stained samples was determined by allowing 5% error for its unstained counterparts and proportions of cells positive for the staining were obtained by subtracting 5% from the percentage of stained samples passing the threshold.

**Giemsa staining and imaging.** For giemsa staining, 50 µl of each sample containing $2.5-5.0\times10^4$ MPRO or NPRO cells were cytopspun onto a glass slide and cytopspinned cells were then stained based on May-Grünwald-Giemsa staining. Brightfield giemsa-stained, cytopspinned cells were scanned using a MIRAX MIDI slide scanner (Carl Zeiss, Germany) that creates stitched composites of entire slides as a single image. Ten fields of scanned cells were sampled at 40× magnification with MIRAX Viewer and the percentage of neutrophil-like cells was estimated (to reduce the laborious work of manual differential cell counting, but not for the automated one, a smaller image containing fewer cells was randomly cropped from each field). The
scanned images were rescaled to 20× and 10× in order to assess whether the automated approach can be applied to images at lower magnification.

**Cell, nucleus and cytoplasm segmentation.** We employed an iterative method to segment the cells. For each iteration, color images were converted into gray scale images and contrast-limited adaptive histogram equalization [141] was used to improve contrast in the images. A global image threshold using Otsu’s method [142] was then applied to identify objects (cell regions); small objects were not considered and holes in each object were filled. Next, minimum bounding boxes associated with each object were determined and used to obtain color images of each object. A threshold for the size of the bounding box was determined to identify images containing a single cell. Images were pre-processed in the next iteration if their size was bigger than the threshold; and if no significant reduction was found for the identified objects, the images were not considered. Four level Otsu’s segmentation was applied to images of a single cell to determine cell, nucleus and cytoplasm regions.

**Feature extraction and selection.** Sixty one morphological features of detected nucleus regions, which include 30 Zernike moments described in [144] and shape measurements such as area, perimeter, and solidity, were derived. In addition, seven statistics associated with pixel value measurement of gray scale segmented nucleus and cytoplasm regions were included. Table 3.1 list all of the 68 features extracted from the nucleus and cytoplasm regions associated with an image of a single cell. Feature selection was performed by using a minimal-redundancy-maximal-relevance (mRMR) criterion based on mutual information that is described in [146].
<table>
<thead>
<tr>
<th>No.</th>
<th>Object</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nucleus</td>
<td>NumberOfSegments</td>
<td>Integer specifying the number of regions or parts (segments) representing the nucleus</td>
</tr>
<tr>
<td>2</td>
<td>Nucleus</td>
<td>EulerNumber</td>
<td>Integer specifying the number of regions or parts (segments) representing the nucleus minus the number of holes in those segments</td>
</tr>
<tr>
<td>3</td>
<td>Nucleus</td>
<td>CentroidLocation</td>
<td>Binary number specifying the location of the centroid of a nucleus relative to its own region (0 if the centroid is outside nucleus, 1 if the centroid is inside the nucleus)</td>
</tr>
<tr>
<td>4</td>
<td>Nucleus</td>
<td>Area</td>
<td>Integer specifying the actual number of pixels of a nucleus</td>
</tr>
<tr>
<td>5</td>
<td>Nucleus</td>
<td>AreaPerConvex-SegmentsArea</td>
<td>Scalar specifying the proportion of the pixels in the convex hulls of nucleus segments that are also in the nucleus region, i.e. the ratio of nucleus area to the sum or area of all convex nucleus segments</td>
</tr>
<tr>
<td>6</td>
<td>Nucleus</td>
<td>AreaPerFilledArea</td>
<td>Scalar specifying the proportion of the pixels in the filled nucleus (holes in nucleus segments filled in) that are also in the nucleus region, i.e. the ratio of nucleus area to the filled nucleus area</td>
</tr>
<tr>
<td>7</td>
<td>Nucleus</td>
<td>Solidity</td>
<td>Scalar specifying the proportion of the pixels in the convex hull of whole nucleus that are also in the nucleus region, i.e. the ratio of nucleus area to nucleus convex hull area</td>
</tr>
<tr>
<td>8</td>
<td>Nucleus</td>
<td>Extent</td>
<td>Scalar specifying the ratio of nucleus area to the total area of nucleus bounding box</td>
</tr>
<tr>
<td>9</td>
<td>Nucleus</td>
<td>EquivDiameter</td>
<td>Scalar specifying the diameter of a circle with the same area as the nucleus region, i.e. sqrt(4*Area/pi)</td>
</tr>
<tr>
<td>10</td>
<td>Nucleus</td>
<td>Perimeter</td>
<td>Scalar specifying the distance around the inner and outer boundary of the nucleus region, i.e. the distance between each adjoining pair of pixels around the inner or outer border of the nucleus regions</td>
</tr>
<tr>
<td>11</td>
<td>Nucleus</td>
<td>PerimeterPerArea</td>
<td>Scalar specifying the ratio of nucleus perimeter to nucleus area</td>
</tr>
<tr>
<td>12</td>
<td>Nucleus</td>
<td>Eccentricity</td>
<td>Scalar specifying the eccentricity of the ellipse that has the same second-moments as the nucleus region, i.e. the ratio of the distance between the foci of the ellipse and its major axis length (the value is between 0 and 1; 0 and 1 are degenerate cases - an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment.)</td>
</tr>
<tr>
<td>13</td>
<td>Nucleus</td>
<td>MajorAxisLength</td>
<td>Scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the nucleus region.</td>
</tr>
<tr>
<td>14</td>
<td>Nucleus</td>
<td>MinorAxisLength</td>
<td>Scalar specifying the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the nucleus region.</td>
</tr>
<tr>
<td>15</td>
<td>Nucleus</td>
<td>Orientation</td>
<td>Scalar specifying the angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse that has the same second-moments as the nucleus region.</td>
</tr>
<tr>
<td>16 to 45</td>
<td>Nucleus</td>
<td>Zernike moments (i.e. Zer0-0, Zer1-1, Zer2-0, ..., Zer9-9)</td>
<td>30 scalars specifying the magnitudes of Zernike moments (as described in Boland et al [144]) through degree 9</td>
</tr>
<tr>
<td>46 to 61</td>
<td>Nucleus</td>
<td>Extrema coordinates (i.e. x.top-left, y.top-left, x.top-right, y.top-right, ..., y.left-top)</td>
<td>16 scalars specifying the x- or y-coordinates of 8 extrema points (top-left, top-right, right-top, bottom-right, bottom-right, bottom-left, left-bottom, and left-top) in the nucleus region relative to its centroid</td>
</tr>
<tr>
<td>62</td>
<td>Nucleus</td>
<td>MaxIntNucleus</td>
<td>Integer specifying the value of the pixel with the greatest intensity in the gray-scale nucleus region</td>
</tr>
<tr>
<td>63</td>
<td>Nucleus</td>
<td>MinIntNucleus</td>
<td>Integer specifying the value of the pixel with the lowest intensity in the gray-scale nucleus region</td>
</tr>
<tr>
<td>64</td>
<td>Nucleus</td>
<td>AvgIntNucleus</td>
<td>Scalar specifying the mean of all the intensity values in the gray-scale nucleus region</td>
</tr>
<tr>
<td>65</td>
<td>Cytoplasm</td>
<td>MaxIntCytoplasm</td>
<td>Integer specifying the value of the pixel with the greatest intensity in the gray-scale cytoplasm region</td>
</tr>
<tr>
<td>66</td>
<td>Cytoplasm</td>
<td>MinIntCytoplasm</td>
<td>Integer specifying the value of the pixel with the lowest intensity in the gray-scale cytoplasm region</td>
</tr>
<tr>
<td>67</td>
<td>Cytoplasm</td>
<td>AvgIntCytoplasm</td>
<td>Scalar specifying the mean of all the intensity values in the gray-scale cytoplasm region</td>
</tr>
<tr>
<td>68</td>
<td>Nucleus</td>
<td>DiffAvgInt</td>
<td>Scalar specifying the difference of AvgIntNucleus to AvgIntCytoplasm</td>
</tr>
</tbody>
</table>
Classifying images of a single cell. Two classes of images of a single cell were defined, one represents the (mature) neutrophil and the other represents the non-neutrophil (other type of cells). A $k$-nearest neighbor ($k$-NN) classifier, combined with mRMR feature selection, was trained to classify cell images. For this, we manually selected 1000 images of a single cell that could reasonably be classified as neutrophils and non-neutrophils (500 images for each class) and derived the morphological features of their nucleus and pixel measurement as described previously.

The automated image analysis system. The proposed automated image analysis system follows the scheme illustrated in FIG. 1.5. First, the system uses the iterative segmentation to obtain the bounding box of the objects considered as a single cell and segment the nucleus and cytoplasm regions as previously described. Then, features selected for cell classification were extracted from the segmented regions. Finally, after classifying the cells with a trained $k$-NN, the estimate for neutrophil purity is calculated.

3.3. Results

3.3.1 Diversity of cell stages in NPROs

Unlike peripheral blood which consists of more or less discrete stages of cell population (i.e. red blood cells, platelets, and leukocytes that are composed of neutrophils, eosinophils, basophils, lymphocytes, and monocytes), NPROs consist of maturing neutrophil cells at varying stages. The stages observed in NPROs often do not clearly belong to the defined stages of neutrophil differentiation (promyelocyte, myelocyte, metamyelocyte, band cell, and segmented neutrophil), but also in between
those defined stages. In addition to this, mitotic stages of promyelocyte and myelocyte could also possibly be observed in NPROs. Furthermore, although the events have not been validated, NPROs may also contain eosinophils. To illustrate the diversity of the cells, snapshots of giemsa-stained NPROs collected immediately (day 0) and on day 3 and 5 after ATRA stimulation with limited medium supplementation are shown in FIG. 3.1A. For a comparison with the stages of neutrophil maturation, two best representatives of NPROs at each stage were selected as shown in FIG. 3.1B.

(A)

(B)

FIG. 3.1. Giemsa-stained NPROs. (A) The diversity of induced differentiated MPRO population observed on day 0, 3 and 5 after ATRA-induced differentiation. (B) Representatives of NPROs at each stage of neutrophil maturation: promyelocyte, myelocyte, metamyelocyte, band cell, and mature neutrophil.
In addition to stages observed for live cells, dead cells are expectedly present in NPRO culture since the life time of neutrophils is short (1-2 days). The percentage of dead cells in NPRO culture can be estimated with trypan blue since viable or live cells exclude it whereas dead cells absorb it and are colored blue. In giemsa-stained NPROs, only dead cells at the later stage (cells have been shrinking and chromatins are condensed) can be recognized easily, but not cells at early stage of apoptosis, necrosis or other types of cell deaths. FIG. 3.2 shows the NPRO viabilities based on trypan-blue exclusion test over 6 days of differentiation for both excess and limited medium supplies. On day 3-6 after differentiation, maintenance with excess medium supply resulted in a high cell viability, i.e. >90% (average of all samples), whereas maintenance with limited medium supply resulted in a slightly lower viability that was about 85% on average.

FIG. 3.2. The viabilities of NPROs over a six day period of maintenance with excess and limited supply of new medium. The viabilities were determined based on the trypan blue exclusion test. The viabilities are shown as means ± SE of four independent determinants. The volumes added during the maintenance are shown in the table above the horizontal axis ($qV = q \times \text{initial volume for differentiation at day 0}$). The differences between the viabilities of NPROs maintained with excess and limited supply of new medium were tested using the two-tail t-test (* p-value <0.05).
3.3.2 Iterative global image thresholding to isolate a single cell from cell clusters in high content images of giemsa-stained NPROs

Cell clusters are often observed in cytospinned giemsa-stained MPRO or NPRO cells. Thus, a segmentation technique is required by an automated image analysis system to single out the cells for further inspection. Here we used a segmentation algorithm that incorporates an iterative global image thresholding and contrast-limited adaptive histogram equalization (described in Materials and Methods) to isolate images of a single cell from the clusters. As illustrated in FIG.3.3, the algorithm is able to split a bigger cluster into smaller clusters and eventually single cells. The objects that are considered as an image of a single cell are determined by filtering the size of the bounding box of the objects. This filter could remove most unwanted images that contain more than one cell, dead cells or cell debris; nevertheless, the cells whose bounding box bigger than the threshold are also removed (exemplified by the cell in the red bounding box no.3 in FIG. 3.3A).

The ability of iterative global image thresholding to segment cells from sampled regions at different magnification was tested using 50 fields of cell clusters scanned at 40× magnification, as well as those fields from these 50 fields rescaled to 20× and 10× magnification (FIG. 3.3B for snapshots of the cell clusters). The success rate of the algorithm, i.e. the ratio between the numbers of objects (cells) singled out by the algorithm versus manual counting, was 0.71, 0.61 and 0.35 for images at 40×, 20× and 10× magnification, respectively. These ratios were about 0.14 higher than the ratios obtained when the algorithm was run without iteration, which were 0.57, 0.46 and 0.22 for 40×, 20× and 10×, respectively. Note that all segmented objects were assumed to be a cell; based on manual screening, the case where an object was not a cell was rarely found.
FIG. 3.3. (A) An example of cell segmentation results output by the iterative global image thresholding. For each iteration, the objects in a blue bounding box are the images of a single cell successfully detected by the iterative thresholding algorithm, whereas the objects in a red bounding box are processed in the next iteration. (B) Fifty fields of giemsa-stained NPRO cells for testing the performance of iterative global image thresholding to isolate a single cell from cell clusters.
3.3.3 Training the \( k \)-nearest neighbor (\( k \)-NN) classifier and selecting features to classify cell images into neutrophils and non-neutrophils

Images of a single cell were initially obtained by employing our segmentation algorithm on a set of images of cytopspun giemsa-stained NPROs collected between day 0 and 6 after induced differentiation. Then, we manually selected and classified 1000 images of a single cell at 40\( \times \) magnification, i.e. 500 images for the mature neutrophil class and another 500 for the immature neutrophil class (samples of training images are shown in FIG. 3.4). The nucleus and cytoplasm of each image were segmented by employing four level Otsu’s segmentation and 68 features associated with these two objects, as listed in Table 3.1, were. For each class, 40% of the images were randomly selected as training data and the rest were used as test data for training the \( k \)-nearest neighbors (\( k \)-NN). To select the best \( k \) for \( k \)-NN, all features were used to classify the cell image. To improve the performance of the \( k \)-NN classifier, feature selection by the mRMR criterion was performed. As in the previous section, these experiments were also performed on images rescaled to 20\( \times \) and 10\( \times \) magnification.

(A) Neutrophils  
(B) Non-neutrophils

FIG. 3.4. Samples of training image data for neutrophils (A) and non-neutrophils (B).
As shown in FIG. 3.5A, training k-NN with varying number of neighbors (k values) indicated that the k=10 gave a considerably better performance in all cases of cell magnification. The best numbers of features for classification with 10-NN were 4, 9 and 9 for 10×, 20× and 40× magnification, respectively (FIG. 3.5B). The corresponding selected features and mean test errors over 100 experiments that decreased slightly towards the higher magnification were shown in Table 3.2. In
addition, recall and precision for neutrophil and non-neutrophil class were kept high (>0.8) for all magnifications despite their slight decrease towards the lower magnification (FIG. 3.5C). Overall, the performance of the classifier is relatively high for all the scales (10×, 20× and 40× magnification) and this is very likely due to the characteristic preservation and strong influence of the first four features listed in the second column of Table 3.2 in differentiating neutrophils from non-neutrophils.

Table 3.2. Selected features for classification and their ranking for each scale. Mean test error for classification at corresponding scale using the selected features are shown at the last row.

<table>
<thead>
<tr>
<th>Object</th>
<th>Feature</th>
<th>40×</th>
<th>20×</th>
<th>10×</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleus</td>
<td>PerimeterPerArea</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Cytoplasm</td>
<td>AvgIntCytoplasm</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Nucleus</td>
<td>CentroidLocation</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Nucleus</td>
<td>Area</td>
<td>8</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Nucleus</td>
<td>Solidity</td>
<td>9</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Nucleus</td>
<td>NumberOfSegments</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Nucleus</td>
<td>Zer2-0</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Nucleus</td>
<td>Zer3-1</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Nucleus</td>
<td>Zer4-0</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Mean test error for classification</strong></td>
<td></td>
<td>0.087</td>
<td>0.091</td>
<td>0.110</td>
</tr>
</tbody>
</table>

3.3.4 Giemsa-staining approach for estimating neutrophil purity in NPROs: automated versus manual

The proposed automated image analysis system incorporated the iterative segmentation algorithm, the 10-NN classifier, and the relevant feature set that depends on the scale of the cell image (as listed in Table 3.2). The relevant features associated with nucleus and cytoplasm of all 1000 manually classified cell images mentioned previously were used to predict the class of new images of a single cell. The automated system was used to estimate neutrophil purity in four biological
replicates of NPROs collected on day 0 to 6 after maintenance with excess and limited medium supplies. However, note that in the first two days of differentiation, the way of maintaining NPROs was the same for both maintenance approaches. Hence, neutrophil yields obtained during differentiation with excess and limited medium supplies were only compared after 3 days of differentiation.

For neutrophil estimation, two glass-slides of cytospinned giemsa-stained cells were prepared for each NPRO sample. Subsequently, ten fields of circular cytospin area (two each from the center, top-right, bottom-right, bottom-left, and top-left) were sampled at 40× magnification from each glass-slide. The automated estimation was done on those sampled images, as well as the ones rescaled to 20× and 10×. In addition, a smaller image containing fewer cells was randomly cropped from each sampled field taken at 40× magnification for manual neutrophil estimation. For each sample, the average numbers of cells being analyzed were 1623, 1569 and 883 for automated estimation at scale 40×, 20× and 10×, respectively, and 346 for manual classification.

FIG. 3.6 shows the trends of neutrophil purities in NPROs collected from day 0 to 6 after differentiation that were estimated manually or using the automated image analysis system at varying magnifications. As shown in the figure, there was a very good agreement between the trends of the automated and manual estimates at any magnification (the correlation between estimates was 0.89, 0.88 and 0.79 for 40×, 20× and 10× magnification, respectively) despite the lower number of cells being counted and increased standard error towards the lower magnification. However, the daily average of the automated estimates tended to be higher than that of the manual estimates in both ways of NPRO maintenance; a significant difference was observed on day 1 after differentiation for automated estimation at 20× and 10× magnification.
Moreover, as confirmed with the Wilcoxon matched test on all pairs of the manual and automated estimates, there was evidence that the automated approach at any magnification gave higher estimates than the manual approach (p-values < 0.01). Finally, when comparing neutrophil yields obtained from NPROs maintained with excess and limited medium supply using unpaired t-test, there was no evidence that those two maintenance approaches gave different neutrophil purity in any day between day 3 to 6 after differentiation, except when comparing manual estimates on day 6 after differentiation (the differences were tested using unpaired t-test).

3.3.5 Comparison between giemsa-staining and flow cytometry approaches for estimating neutrophil purity in NPROs

Neutrophil purity in NPROs collected on day 0 to 6 after differentiation in excess and limited medium supplies was also estimated by flow cytometry based on anti-Ly-6G, a specific neutrophil marker. The estimates, which were an average of four flow measurements on a different set of NPRO cultures, are shown in FIG. 3.7 in comparison with the estimates obtained by manual giemsa approach. As shown, estimates given by giemsa approaches were often higher than those obtained by anti-Ly-6G-based flow cytometry; significant differences were observed on day 1 for maintenance with excess medium and on day 6 for maintenance with limited medium. The difference between giemsa and flow estimates was confirmed with two-way ANOVA test (p-value associated with the estimation approaches < 0.01). Interestingly, unlike giemsa estimates, the flow estimates on day 5 and 6 were similar. Nevertheless, the flow cytometry approach also did not suggest any significant difference between neutrophil purities in NPROs maintained in excess and limited medium supplies.
FIG. 3.6. Comparison between automated (at scale 40×, 20× and 10×) and manual estimates for neutrophil purity in cytopspun giemsa-stained NPROs collected on day 0 to 6 after differentiation. NPROs were maintained with excess or limited medium (left and right panels, respectively). The estimate at each day of differentiation are shown as means ± SE of four independent biological replicates. For each magnification, the differences between the automated and manual estimates at each day of differentiation were tested using the paired two-tail t-test with Bonferroni correction (* adjusted p-value <0.05).
FIG. 3.7. Anti-Ly-6G-based flow cytometry estimates (average of four replicated samples) for neutrophil purity in NPROs collected on day 0 to 6 after differentiation, in comparison with manual estimates for neutrophil purity in cytopspined giemsa-stained NPROs. NPROs were maintained with excess or limited medium (left and right panels, respectively). The estimate at each day of differentiation are shown as means ± SE of four independent biological replicates. For each magnification, the differences between the flow and manual estimates at each day of differentiation were tested using the unpaired two-tail t-test with Bonferroni correction (* adjusted p-value <0.05).

FIG. 3.8. A representative of flow cytometry analysis of NPROs which were unstained, stained with PerCP-Cy5.5 anti-Ly-6G or FITC anti-Gr-1 only, and stained with both PerCP-Cy5.5 anti-Ly-6G and FITC anti-Gr-1.

Furthermore, we compared the estimates for neutrophil purity in NPROs given by flow cytometry based on anti-Ly-6G and non-specific neutrophil marker antibody, i.e. Gr-1 (FIG. 3.8). NPROs were maintained in limited medium supply and collected on day 5 after differentiation. Our experiments revealed that the Gr-1-based estimate
for neutrophil purity in NPROs collected on that day was higher than anti-Ly-6G-based estimate, i.e. 32% versus 18% (average of four replicated samples). By double staining NPROs with anti-Gr-1 and anti-Ly-6G, we showed that Ly-6G$^+$ cells were mainly Gr-1$^{hi}$, and there was a population of Gr-1$^{int}$ cells that were Ly-6G$^-$ and might represent immature neutrophils.

### 3.4 Discussion

First, we showed that a simple iterative global image thresholding could be used to increase the number of single cells segmented in images of giemsa-stained. Then, by using the 10-nearest neighbour classifier on selected features, the images of the single cells could be classified into neutrophils and non-neutrophils with considerably high accuracy. Finally, we showed that an automated image analysis system that combines the segmentation approach and the classifier could estimate the trend of neutrophil purity in NPRO samples over 6 days of differentiation. As expected, the best performance of the algorithm, classifier and automated system was achieved when they were applied on images at scale 40×. Interestingly, although the performance at scale 10× was not satisfying, the performance at scale 20× was similar to that at scale 40×.

For the automated image analysis system, its estimate for neutrophil purity is often higher than manual estimate. At scale 40×, the daily differences between the averages of automated and manual estimates were less than 2% for day 0 to 2, and often more than 4% and up to 10% for day 3 to 6. The tendency of the automated system to give higher estimates could be partly explained by the fact that non-neutrophils were more often misclassified than neutrophils, as it can be inferred from
FIG. 3.5C. As a consequence of the tendency, the automated system also often gave estimates that have a higher variance than manual estimates.

Despite these problems, the automated image analysis system generally performed satisfactorily at scale 40× and 20× for the samples we worked on. It gave an almost similar conclusion when comparing the two approaches in maintaining NPROs. Additionally, it has an advantage that it can inspect a lot more giemsa-stained cells than the manual approach reasonably does for the same amount of time. Hence, at this point, we viewed the automated system as a good complement, or even substitution, for the manual approach to inspect the outcomes of differentiating MPRO cells.

Next, in order to examine the potential of NPROs as a neutrophil source, we set a goal to determine the best time to harvest NPROs as well as the highest neutrophil yield that NPRO culture can provide. For this, NPROs were maintained over 6 days of differentiation and their daily neutrophil contents were inspected with manual and automated differential counting of giemsa-stained cells, as well as flow cytometry based on anti-Ly-6G and anti-Gr-1 staining. Neutrophil yields in NPROs maintained with excess and limited supplies of new differentiation medium were compared. Initially, maintenance with excess medium supply was intended to keep cell viability high, whereas maintenance with limited medium supply was thought to induce promyelocyte differentiation rather than division, despite its potential to cause more cell death. As shown in our FIG. 3.2 and FIG. 3.7, this trend was indeed observed although the differences were rarely significant.

Of particular interest, anti-Ly-6G-based estimates were usually lower than those estimated by giemsa-staining approaches, although correlations between time series estimates from both approaches were above 0.85. The inconsistency was worse
when comparing the estimates on day 6 after differentiation – which to our best guess, the loss of Ly-6G surface markers is due to neutrophil senescence. Additionally, a possible explanation is that Ly-6G is expressed at the latest stage of neutrophil maturation, right after the cells have a segmented nucleus. Perhaps, the gap between giemsa- and anti-Ly-6G-based estimates might correspond with the gap between anti-Ly-6G- and anti-Gr-1-based estimates that we also demonstrated for NPROs collected on day 5. Anti-Gr-1 is less specific than anti-Ly-6G since it recognizes Ly-6G as well as another Ly-6 family, i.e. Ly-6C, which is present in not only neutrophils but also other immune cells such as macrophages and CD8\(^+\) T cells [187, 188]. By double staining NPRO cells with Gr-1 and Ly-6G, we showed that Ly-6G\(^{hi}\) cells were mainly Gr-1\(^{hi}\), and there was a population of Gr-1\(^{int}\) cells that were Ly-6G\(^{low}\) and might represent immature segmented neutrophils that have a particular level of Ly-6C expression before they expresses Ly-6G.

Despite the gaps between approaches, both giemsa-based- and anti-Ly-6G-based estimates showed almost similar trends. Based on these trends and the trend for NPRO viabilities, we suggest that the best time to harvest NPROs is on day 5 after differentiation. Although both manual and automated giemsa-based estimates on day 6 were higher than those on day 5, the anti-Ly-6G-based estimates showed an opposite pattern and moreover, NPRO viability was at its lowest on day 6 and the culture might have more senescent neutrophils; hence, harvesting NPROs on day 6 was not advised. On day 5 after differentiation, the overall average of neutrophil purity was about 20-32%. In particular, it was 32% according to automated giemsa-based estimate at scale 40×, 27% according to manual giemsa-based estimate, 20% according to anti-Ly-6G-based flow cytometry estimate, and 32% according to anti-Gr-1-based flow cytometry estimate.
In addition to maximizing neutrophil yields, keeping the viability of harvested NPRO high is another issue to be addressed. This is best done during the maintenance of NPRO cells since removing dead cells cannot be done simply by centrifugation, which is due to the nature of NPRO cells as a suspension culture. Furthermore, the success of obtaining NPROs with a highly viable and purer neutrophil population is influenced by various factors that include the composition of complete medium and the way of maintaining NPRO cells. Without medium update, the viability of NPROs dropped significantly on day 4 after differentiation and almost all cells died on day 5. Adding new medium can keep the viability high since it contains GM-CSF, which is known to promote neutrophil survival by inhibition of programmed cell death [189]. Regarding to our approach in maintaining NPROs, we are able to keep the viability above 90% while providing excess supply of new medium, and successfully maintained it at about 85% during limited medium supply, on average.

Taken together, we have assessed various alternatives in estimating neutrophil purity in NPRO culture by both giemsa-staining or flow cytometry approaches. Previously, the estimation was mainly done by manual differential counting of giemsa-stained NPROs, or alternatively, using a flow cytometry approach based on non-specific neutrophil marker antibody Gr-1 or MAC-1 [130]. Here, we rigorously examined neutrophil yields in NPROs not only by using those approaches, but also using an automated image analysis system and a flow cytometry technique based on a specific neutrophil marker antibody, anti-Ly-6G. As a main result, our experimentation indicated that NPROs is best harvested on day 5 after induced differentiation, where on average the neutrophil yield and cell viability are about 20-30% and 85-90%, respectively. This percentage is still unsatisfactorily, and hence,
further experiments to increase neutrophil yields in NPROs need to be carried out. For this, investigation with different retinoids listed in [16] to stimulate MPRO differentiation, a different source of serum and the use of a mitotic inhibitor might be considered.
Chapter 4

Transcriptional profiling of neutrophils infected with highly and less virulent IAV (H3N2) viruses

In Chapter 2, we have performed genomic analysis on lungs infected with highly and less virulent IAV (H3N2) viruses, i.e., the HVI and LVI. The outcomes of the analysis have provided insights about biological processes and pathways involved in the progression of highly virulent influenza. Nevertheless, the outcomes only reflect the dynamics of gene regulation at organ level, and conclusion about which cell type contributes to the dynamics cannot be drawn. Hence, to uncover the contribution of specific cell types to the global pulmonary response, genomic analysis needs to be performed at cellular level.

Here we put interest in uncovering transcriptomic changes in neutrophils following infection with HVI and LVI. The roles of neutrophils during \textit{in vivo} infection with IAV of varying virulence have been investigated, and several effects of IAV infection in neutrophils have been uncovered \textit{in vitro} despite its virulence (discussed in Section 1.3.2 and 1.3.5, respectively). However, hitherto, no \textit{in vitro} study has investigated global neutrophil response to IAV infection. Hence, we have become the first to uncover global neutrophil response to IAV infection, and specifically, we compare differential global effects of neutrophils infected with highly virulent and less virulent IAV.
versus less virulent IAVs. For the infection model, we employed neutrophils from induced differentiated MPRO cells (NPROs) as a neutrophil source instead of using neutrophils isolated from bone marrow or blood.

We highlight the motivation of the genomic study in Section 4.1, followed by descriptions of materials and methods in Section 4.2. Results related to transcriptomic responses of NPROs infected with LVI and HVI, as well as validating the use of the cells for modeling IAV infection in neutrophils, are presented in Section 4.3. Finally, Section 4.4 concludes novel insights on the contribution of infected neutrophils to global response observed in lungs during highly virulent IAV infections.

4.1. Motivations

Highly virulent IAV infections are characterized by the dysregulation of the innate immune response in various influenza animal models, including our murine infection model using HVI of H3N2 subtype (discussed in Chapter 2). In particular, several type I IFN-stimulated and cytokine/chemokine genes are strongly over-expressed immediately after infection with highly virulent IAVs. However, analyses of bronchoalveolar lavage from animals infected with highly virulent IAVs show a significant increase in cellularity and numbers of neutrophils and macrophages compared to those infected with less virulent IAVs [14]. This has also been shown in lungs infected with HVI [154]. It is therefore logical to investigate the global transcriptomic responses of these two innate immune cell types following infection with IAVs of varying virulence, and to elucidate their contributions to differential lung pathogenesis.

Currently, activated macrophages following highly virulent IAV infection have been considered to play a greater role than neutrophils in causing the
dysregulation of innate immune responses. This is corroborated by a study that has demonstrated high expression levels of type I IFN-inducible and cytokine/chemokine genes following a highly pathogenic H5N1 virus infection in human monocyte-derived macrophages [9]. However, murine alveolar macrophages infected with the highly virulent mouse adapted PR8 of H1N1 subtype surprisingly failed to induce strong expression of these genes [10]. These opposing results might be attributed to the ability of H5N1 virus to replicate in human alveolar macrophages, but not the pandemic swine origin H1N1-2009 virus [116]. Thus, the extent of the contribution of macrophages to dysregulation of innate immune response following highly virulent IAV infection appears to be varying. Other types of immune cells may also contribute to the dysregulation of innate immune response.

We hypothesized that neutrophils might contribute to the dysregulation of innate immune response. This proposal is also driven by our previous findings, where genes and pathways associated with neutrophils activities, e.g. neutrophil chemoattractant CXCL1 and TREM1 signaling, were activated early after HVI infection (discussed in Section 2.3.4). To provide evidence for our hypothesis, we were motivated to reveal global response of neutrophils infected with LVI and HVI. For this, we have decided to develop in vitro infection model by using cultured NPROs, containing both mature and immature neutrophils, which is technically easier to be obtained, cheaper and less laborious than using neutrophils isolated from mouse blood or bone marrow. The use of NPROs is also without ethical issues, and moreover, NPROs have been shown to provide valid neutrophil model [15] and their use may reflect a more realistic scenario in which immature neutrophils shown to be circulating during acute influenza [205]. Since NPROs were used for the first time to model IAV infection in neutrophils, we were also motivated to initially validate the
model by checking whether IAV infection in NPROs elicits similar characteristics to IAV infection in normal neutrophils. These include showing that IAV infection in NPROs is abortive and augments apoptosis.

4.2 Materials and methods

**MPRO cell culture and differentiation into NPROs.** MPRO cell lines (ATCC Number CRL-11422) were cultured in IMDM containing 4 mM L-glutamine, 3 g/L sodium bicarbonate, 10 ng/mL murine GM-CSF, and 20% heat-inactivated FBS at 37°C with 5% CO₂. The culture medium was replenished in 1-2 days to bring cell concentration into 5×10⁵ cells per ml. To induce the differentiation of MPRO cells into NPROs, the growth medium was supplemented with 10 μM ATRA. NPROs were maintained with excess or limited amount of nutrients (i.e. at low or high concentration of cells), as described in Materials and Methods of Chapter 4. For most of the experiments with IAVs, especially microarray experiments, NPROs were harvested on day 5 after maintained with limited amount of nutrients.

**MDCK cell culture.** MDCK cells (ATCC Number CCL-34) were cultured in a TC-75 flask at 37°C with 5% CO₂ in EMEM containing 10% FBS. Cells were subcultured once confluent by first removing medium and washing the cell monolayer with 1× PBS. Cell monolayer was then trypsinized with 1× trypsin-EDTA solution (Invitrogen, Carlsbad, CA). Fresh EMEM was added into the flask and solution was flushed to detach cells completely. Next, cell suspension was centrifuged and cell pellet was resuspended with fresh growth medium, and finally cells were cultured in a new flask.
Virus strains, determination of virus titers, and neutrophil infection. Influenza virus strain A/Aichi/2/68 H3N2 was purchased from the ATCC, propagated in eggs, and shown to cause less virulent influenza in mice. The virus strain was also adapted to cause highly virulent influenza in a mouse model through serial lung-to-lung passaging until the tenth passage [13]. Non-adapted virus from the original stock and the adapted virus from the tenth passage are designated as LVI as HVI, respectively. For further use, HVI was passaged using 4-6 week-old female BALB/c mice, where lung homogenates were collected two days after infection. All animal experiments were conducted in a BSL-2 laboratory, and approved by the Institutional Animal Care and Use Committee. Following further passaging of HVI in mice, HVI of the fourteenth passage was available for use. Since HVI replicates more efficiently than LVI, we also propagated the HVI in MDCK cells (termed as “MPI”), and used MPI to perform most experiments that analyzed infection of NPRoS. Virus titers were determined either by plaque assay or TCID\textsubscript{50} assay [13]. For infection, NPRoS were incubated with IAV for 1 h at 37°C with 5% CO\textsubscript{2}, and then resuspended in IMDM containing 25 ng/ml GM-CSF to 5 × 10\textsuperscript{6} cells per ml.

Labelling IAV with lipophilic dye for identifying infected cells. Plain EMEM or EMEM containing 10\textsuperscript{7} pfu of MPI per ml was incubated with 1:20 dilution of the fluorescent lipophilic dye (1,1’-dioctadecyl-3,3,3’,3’-tetramethyl-indodicarbocyanine or DiD) in a 6-well plate for 1 h at 37°C with 5% CO\textsubscript{2}. After incubation, the inocula were filtered using 0.45 μm Minisart\textregistered NML syringe filters (Sartorius Stedium), and frozen at -80°C until further use. The inoculum without virus but incubated with DiD, which was also filtered, served as the DiD mock control. Following incubation of 10\textsuperscript{6} cells with DiD-labelled MPI for 1 h at MOI of 0.1, 1 or 10, infected cells were
identified by flow cytometry or confocal microscopy. For flow cytometry, cells were washed with PBS, stained with FITC Ly-6G (BD Pharmingen), fixed, and resuspended in sheath buffer. For confocal microscopy, cells were washed with PBS, cytopspun onto glass slides, permeabilized with Triton X-100, stained with DAPI, and examined with the Olympus FluoView FV1000 laser scanning confocal microscope using a 100×/1.45 oil objective, with 405nm solid state laser diode, and 633nm HeNe laser as the excitation source (for visualizing DiD dye).

Comparing kinetics of IAV replication in MDCK and NPRO cells. To compare the viral replication kinetics, confluent MDCK cell monolayer in 12-well plates was inoculated with LVI, HVI or MPI at MOI of 0.1 (without adding TPCK trypsin) and incubated for 1 h at 37°C with 5% CO₂. Inocula were removed, replaced with 1 ml of plain EMEM, and 100 μl of each supernatant were collected after 3, 6, 12 and 24 h for virus plaque assay. The replication kinetics of MPI in MDCK and NPRO cells following infection at MOI of 0.1 were compared in 24-well plates. Infected NPROs were incubated in fresh IMDM containing 25 ng/ml GM-CSF (800 μl, 5 × 10⁶ cells per ml). Supernatants (50 μl) of both infected cell types were harvested after 6, 12 and 24 h for virus plaque assay.

Immunofluorescence detection of IAV-infected cells. At 6, 9, 12 and 24 h after infection of 10⁶ NPRO cells with MPI at MOI of 1, 3-5 × 10⁴ cells were cytopspinned onto a glass slide. Cells were fixed with formaldehyde, permeabilized with 1% Triton X-100 and blocked with 2% FBS in PBS for 10 min and 1 h, respectively. After washing with PBS, cells were incubated overnight with 1:250 dilution of primary rabbit antibody raised against influenza A/Aichi/2/68 H3N2 virus. Cells were then
washed twice with PBS, incubated with 1:250 dilution of anti-rabbit Alexa 555 (Molecular Probes) for 1 h at room temperature, washed, and DAPI was applied. Twelve fields of infected NPROs were captured for analyzing the kinetics of viral protein synthesis using the Olympus BX60 fluorescence microscope. Staining was also performed on NPROs infected with LVI after 9 h and HVI after 3 h (MOI of 1 was used). Infected MDCK cells stained in similar way served as positive control.

**Early and late apoptosis assays.** Flow cytometry based on annexin V and propidium iodide (PI) staining was performed on 5 replicates of paired uninfected and MPI-infected (at MOI of 1) samples collected at 0, 3, 6, 9, 12 and 24 h after incubation. Infected or uninfected NPROs (~1.5×10⁶ cells) in 24-well plates were washed with PBS, and incubated with 1:50 dilution of both FITC annexin V and PI (BD Pharmingen) in 100 µl of binding buffer for 20 min at room temperature in the dark. Then, 400 µl of binding buffer was added to each sample, cells were gently resuspended, and analyzed by flow cytometry as previously described. TUNEL assay was also performed on cytospinned samples collected at 12 h after incubation using an In Situ Cell Death detection kit, POD (Roche Applied Science) according to the manufacturer’s protocol. The percentage of apoptotic cells was ascertained in a similar way as for the manual neutrophil estimation by giemsa-staining described in Section 3.2.

**Oligonucleotide microarray experiments.** NPROs were infected with LVI, HVI (at MOI of 0.1) and MPI (at MOI of 1) for 1 h in IMDM containing 10 ng/ml GM-CSF. The control was represented by NPROs incubated with uninfected murine lung homogenate (mock infection). NPROs were resuspended into 24-well plates (3 × 10⁶
cells per well). After 3 and 9 h of incubation, each supernatant was removed and total RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer’s recommendations. Two biological replicates of total RNAs were prepared for each experimental group. The quantity and purity of total RNA were measured by the Nanodrop ND-100 spectrophotometer, while RNA integrity was verified by the Agilent 2100 Bioanalyzer. Total RNA was labelled with the One Color Low Input Quick Amp Labeling kit, version 6.5 (Agilent) following the manufacturer’s instructions. Briefly, 100 ng of each RNA sample was converted into double-stranded cDNA with an oligo-dT primer containing the recognition site for T7 RNA polymerase. In vitro transcription with T7 RNA polymerase generated cyanine 3-CTP labeled cRNA. The labelled cRNA (600 ng) was hybridized onto Agilent SurePrint G3 Mouse GE 8x60K Microarray (design ID 028005) for 17 h at 65°C, 10 rpm in Agilent hybridization oven. The microarray slide was then washed in wash buffer 1 for 1 min at room temperature, and another minute in wash buffer 2 at 37°C before scanning with the Agilent High Resolution Microarray Scanner (C-model). Raw signal data were extracted from the TIFF image with Agilent Feature Extraction Software (version 10.7.1.1).

Microarray data preprocessing and analyses. GeneSpring GX 11.5 (Agilent) was utilized for normalizing and filtering the Agilent microarray data. Percentile shift normalization (set at 75th percentile) was applied on log-transformed data. Baseline transformation to median for each batch of data from each replicate was performed. Batch effect correction was also carried out to correct markers that showed consistent signals within batches but large variations between batches. Probes flagged as detected in any experiment were included in a working list (22,911 transcripts). Two-
Chapter 4

way ANOVA was applied on this list using R software (http://www.r-project.org/) to identify genes that had a significant infection, time and interaction effects. If a gene had no significant interaction effect, we employed two-way ANOVA without considering the interaction effect to improve the p-value for infection and time effects. The Benjamini-Hochberg method for multiple test correction implemented in the R software was utilized to adjust p-values of the ANOVA test for each gene. Gene Ontology (GO) analysis was performed using FuncAssociate 2.0 [148], while pathway analysis was carried out using IPA software.

Quantitative real-time RT-PCR analyses. Total RNA was converted into cDNA using MMLV reverse transcriptase (Promega), and SYBR Green PCR analyses were performed on two technical samples using the LightCycler system (Roche) according to the manufacturer’s protocol. Primer pairs for selected genes were synthesized based on PrimerBank [168, 169], i.e. GAPDH (6679937a1), DDX58 (2737002a1), IRF1 (6680467a1), IRF7 (8567364a1), IRF9 (6680474a1), IFIH1 (23956208a1), STAT1 (27502700a1), STAT2 (9910572a1), IFIT3 (6754288a1), ISG15 (7657240a1), ISG20 (15805028a1), OASL2 (16924024a1), and CXCL10 (10946576a1); except IFN-β primers (forward primer: 5'-CCACAGCCCTCTCCATCAACTATAAGC-3', reverse primer: 5'AGCTC TTCACTGGAGAGCAG TTGAGG-3'). The uniformity of each amplified product was confirmed by the melting curve displaying a single consistent peak. The GAPDH housekeeping gene was used for normalizing the expression data of other genes. The specificity of each primer pair was confirmed by a single diagnostic band in gel electrophoresis of real-time RT-PCR products.
4.3. Results

4.3.1 Neutrophils from induced differentiated MPRO cells (NPROs) are infectable by IAV

The propagation of HVI in MDCK cells allowed us to produce numerous stocks of high titer IAV. The cost of the propagation is also much cheaper than the cost of the propagation of HVI in mice or propagation of LVI in eggs. Considering this, we used the virus from these stocks for investigating the model of IAV infection in NPROs, and here, we addressed whether IAV is able to infect mature and immature neutrophils in NPROs. As shown in FIG. 4.1, the virus (designated as MPI) interestingly showed a multiplication kinetic that is comparable to that of HVI and significantly higher than that of LVI.

![Graph showing viral replication kinetics of LVI, HVI, and MPI in MDCK cells](image)

**FIG. 4.1.** Comparison between viral replication kinetics of LVI, HVI, and MPI in MDCK cells. The data represent mean ± SE of three independent measurements of log virus titers. The difference between log virus titers for each pair of viruses at each time point was tested using two-tail unpaired t-test (* p-value <0.05).

IAV was labelled with a fluorescent lipophilic DiD dye to evaluate the percentage of NPROs being infected. Using $10^6$ cells in 200 μl, flow cytometry
detected about 4% and 11% infected cells (average of four biological replicates) after 1 h incubation with DiD-labelled MPI at MOI of 0.1 and 1, respectively. NPROs positive for DiD staining could still be detected at 24 h postinfection. Furthermore, DiD-labelling of MPI indicated that both Ly-6G-positive and Ly-6G-negative (i.e. mature and immature) neutrophils could be infected (FIG. 4.2). Finally, immediately after incubation, confocal microscopy could visualize infected NPROs with DiD dye clearly dispersed in the cytoplasm (FIG. 4.3).

FIG. 4.2. A representative of flow cytometry analysis of NPROs infected either with unlabelled or DiD-labelled MPI, and unstained or stained with FITC Ly-6G.

FIG. 4.3. Confocal microscopy on NPROs examined 1 h after incubation with DiD-labelled MPI. Two representative examples are depicted: the left panels show nuclei stained with DAPI (blue), middle panels exhibit the cytoplasm of infected cells containing DiD label (red), while right panels are the merged images.
4.3.2 NPROs do not support active IAV replication

Virus titers in supernatants of infected NPROs at various time-points remained relatively unchanged, but were significantly increased in supernatants collected from infected MDCK cells as positive control (FIG. 4.4A).

Representative immunofluorescence staining of infected NPROs at 9 h postinfection using antibody raised against H3N2 virus indicated that viral proteins were synthesized in the cytoplasm as well as nucleus (FIG. 4.4B). Interestingly, the percentage of NPROs positive for viral staining was very low at 6 h postinfection (less than 1%), but modestly increased from ~2.5% at 9 h postinfection to 5.5% at 24 h postinfection (FIG. 4.4C). Taken together, these findings suggest that IAV infection of NPROs is abortive, i.e. the virus does not actively replicate within NPROs, although viral proteins can be synthesized.

4.3.3 IAV infection of NPROs augments early and late apoptosis

The effects of MPI infection on NPRO cell death and apoptosis were assessed by flow cytometry based on annexin V (A5) and propidium iodide (PI) staining. This assay determines the proportions of viable cells (A5⁻PI⁻ cells), and of cells undergoing early apoptosis (A5⁺PI⁻ cells). Cell viability of infected NPROs was generally lower than that of uninfected samples, which was significant at 12 and 24 h postinfection (data not shown). Furthermore, the proportion of NPROs undergoing early apoptosis in the infected group was significantly greater than in the uninfected group at 6 h postinfection onward (FIG. 4.5A). In addition, the TUNEL assay indicated a significantly higher percentage of apoptotic cells in the infected group (FIG. 4.5B). Thus, IAV infection induces NPRO death and apoptosis.
FIG 4.4. (A) Comparison of virus titers of supernatants from infected MDCK cells and NPROs at 6, 12 and 24 h postinfection with MPI at MOI of 0.1. The data represent means + SE of three independent measurements of log virus titers. The means of log virus titers vs detection limit was tested using one-tail unpaired t-test (* indicates p-value < 0.05). (B) Detection of viral protein synthesis in infected NPROs via immunofluorescence staining (IF) using polyclonal antibody against H3N2 virus. Viral proteins could be observed in the cytoplasm and/or nucleus of neutrophils at 9 h after incubation with MPI. (C) Kinetics of viral protein synthesis based on the percentage of cells positive for viral staining. The percentages are shown as means ± SE of three independent IF stained samples of NPROs that were initially infected with MPI at MOI of 1.
FIG. 4.5. Induction of enhanced apoptosis in NPROs infected by IAV. About $1.5 \times 10^6$ cells were initially infected with MPI at MOI of 1, and compared with uninfected cells. (A) Flow cytometry based on A5 and PI staining indicated that the proportions of infected cells undergoing early apoptosis (A5-positive and PI-negative) were significantly higher than uninfected cells at multiple time-points. (B) TUNEL assay also revealed augmented late apoptosis of infected NPROs at 12 h postinfection, as illustrated by representative images, and quantification of the proportion of TUNEL-positive cells. The proportions are shown as means ± SE of five independent determinants for (A) and four determinants for (B). The differences between the proportions of apoptotic cells in infected and uninfected samples were tested using the one-tail paired t-test (* p-value <0.05).

4.3.4 Transcriptomic data analyses of NPROs infected with MPI, LVI and HVI

To investigate global transcriptional response of neutrophils to H3N2 virus, gene expression data were obtained at 3 and 9 h postinfection from NPROs mock-infected with mouse lung homogenate (control), and infected with LVI, HVI and MPI.
Normalized expression data were analyzed by two-way ANOVA with Benjamini-Hochberg correction for multiple tests to identify genes with significant infection, time and interaction effects. Using a significance level of 0.05 for adjusted p-value, we identified 133 genes with significant infection effect, 4,463 genes with significant time effect, and 138 genes with significant interaction between infection and time effects. The overlap between these significant groups of genes is depicted in FIG. 4.6A.

We focused mainly on the expression profiles of 129 genes that revealed a significant infection effect with an absolute log ratio intensity relative to mock infection of >0.6 at any time-point of any treatment (referred to as differentially expressed genes) as shown in FIG. 4.6B. The most striking result was that NPROs responded to HVI more rapidly than to LVI and MPI. In particular, many genes were over-expressed as early as 3 h postinfection with HVI, whereas their enhanced expression was only observed at 9 h postinfection for LVI and MPI. Interestingly, the maximal expression changes for most genes in all infections were quite comparable even though the MOI for MPI was ten-fold higher than that for LVI and HVI. Notably, more intense expression changes in response to MPI coincided with relatively higher viral protein synthesis detected at 9 h postinfection as described earlier, whereas only very few NPROs infected with LVI were positive for viral protein staining at this time-point (data not shown).

4.3.5 Early and stronger type I IFN response characterizes HVI infection of NPROs

Analysis with IPA software revealed activation of IFN regulatory factor (IRF) and IFN signaling pathways as the two most significantly over-represented pathways. The
most interesting sub-network in these pathways was that associated with IFN-β (IFNB1) since this gene, but not type I IFN-α and type II IFN-γ, was differentially expressed in our microarray data. In addition, FuncAssociate 2.0 listed “cellular response to IFN-β” as the most significant over-represented Gene Ontology category. FIG. 4.7 depicts the heatmap summarizing the expression changes of mainly type I IFN-inducible genes. It is notable that the expression of most IFN-related genes in HVI was more intense at 3 h postinfection compared to that at 9 h postinfection, while the reverse pattern was true for LVI and MPI. FIG. 4.8 illustrates the sub-network associated with IFN-β and highlights: (1) the role of two RNA sensing molecules DDX58 and IFIH1 in activating IRF3 and IRF7 to induce IFN-β expression and initiate antiviral innate immunity; and (2) the role of IFN-β signaling in elevating the expression of IRF7 and antiviral genes via activation of STAT1, STAT2 and IRF9 complex.

To validate the microarray data, we performed quantitative RT-PCR for several IFN-related genes that were up-expressed and of our interest, including DDX58, IFIH1, IFN-β, IRF1, IRF7, IRF9, STAT1, STAT2, ISG15, ISG20, IFIT3, OASL2 and CXCL10 (FIG. 4.9). For several of these genes, i.e. DDX58, STAT2, ISG15, ISG20, OASL2, and CXCL10, there was high correlation between microarray and RT-PCR (0.96, 0.90, 0.86, 0.93, 0.95, and 0.95, respectively). For all infections, RT-PCR detected over-expression of IRF1 especially at 3 h postinfection, and of IRF7 and IRF9 later at 9 h postinfection. Furthermore, RT-PCR also indicated early induction of DDX58, STAT1, STAT2, ISG15, ISG20 and OASL2 genes particularly for HVI. Similar to microarray data, although correlation score was not high, IFN-β gene was mainly over-expressed in MPI and LVI but not in HVI. Overall, there was generally good consistency between RT-PCR and microarray data, with both
approaches providing evidence for earlier expression of type I IFN-inducible genes in NPROs infected with HVI compared to LVI and MPI.

![Venn diagram](image)

**FIG. 4.6.** Analyses of gene microarray data of IAV infection of NPROs. (A) Venn diagram showing the overlap between significant groups of genes as analyzed by two-way ANOVA with Benjamini-Hochberg correction for multiple tests (adjusted p-value <0.05). (B) Heatmap of 129 genes with a significant infection effect, and an absolute log ratio of >0.6 at any time-point of infection with LVI, HVI and MPI.
FIG. 4.7. Overview of IFN-related genes in response to infection with LVI, HVI and MPI at 3 and 9 h postinfection (p.i.).

FIG. 4.8. A subnetwork of canonical pathways associated with IFN-β regulation and signaling. Genes that were over-expressed in LVI, HVI and MPI infections by microarray data analysis are shaded grey; genes over-expressed in LVI and MPI but not in HVI shaded black.
FIG. 4.9. Validation of microarray data for several IFN-related genes by real-time RT-PCR. Log fold changes (relative to mock control) shown in the graphs reflect the average of two biological samples (each biological sample was an average of two technical measurements).

4.3.6 Expression of apoptosis-related genes during MPI, LVI and HVI infections of NPROs

Since NPRO apoptosis was evident, here we highlight apoptotic gene expression eventhough GO term for ‘apoptosis’ was not found to be significantly over-represented. Microarray data also elucidated higher expression levels of genes associated with apoptotic activity in infected versus uninfected NPROs, including CASP4, EIF2AK2, NOD1, XAF1 and genes related to retinoic acid-mediated apoptosis signaling, i.e. PARP9, PARP10 and PARP14. It should be emphasized that increased expression of these genes occurred earlier at 3 h postinfection with HVI, in comparison with 9 h postinfection with LVI and MPI (Table 4.1). Specifically for
MPI infection, elevated expression of these genes was concomitant with the enhanced apoptosis observed in infected NPROs.

Table 4.1. Changes in expression of genes associated with apoptotic activity in NPROs infected with IAVs. The gene expression changes for MPI, LVI and HVI infections at the time-points are expressed as log ratios.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Biological activity</th>
<th>MPI (3h.p.i. 9h.p.i.)</th>
<th>LVI (3h.p.i. 9h.p.i.)</th>
<th>HVI (3h.p.i. 9h.p.i.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASP4</td>
<td>execution-phase of cell apoptosis</td>
<td>-0.9 1.9</td>
<td>-0.7 2.8</td>
<td>1.2 6.5</td>
</tr>
<tr>
<td>EF2AK2</td>
<td>inhibition of protein synthesis</td>
<td>0.7 2.7</td>
<td>1.2 3.6</td>
<td>4.2 1.2</td>
</tr>
<tr>
<td>NOD1</td>
<td>enhance caspase 9-mediated apoptosis</td>
<td>0.4 0.9</td>
<td>0.6 2.4</td>
<td>1.4 0.4</td>
</tr>
<tr>
<td>PARP9</td>
<td>retinoic acid-mediated apoptosis signaling</td>
<td>0.5 2.8</td>
<td>0.7 3.2</td>
<td>1.2 1.3</td>
</tr>
<tr>
<td>PARP10</td>
<td>retinoic acid-mediated apoptosis signaling</td>
<td>-0.1 1.7</td>
<td>0.4 2.8</td>
<td>1.3 0.9</td>
</tr>
<tr>
<td>PARP14</td>
<td>retinoic acid-mediated apoptosis signaling</td>
<td>0.5 2.6</td>
<td>0.8 4.0</td>
<td>4.1 1.5</td>
</tr>
<tr>
<td>RNF34</td>
<td>anti-apoptotic function</td>
<td>0.0 1.7</td>
<td>0.8 2.3</td>
<td>2.4 0.8</td>
</tr>
<tr>
<td>XAF1</td>
<td>inhibition of anti-caspase activity</td>
<td>1.2 3.9</td>
<td>1.6 5.1</td>
<td>5.0 3.2</td>
</tr>
</tbody>
</table>

4.4 Discussion

We succeeded in employing the NPROs in suspension culture as a source of highly viable neutrophils with a significant proportion of mature neutrophils (20%-30%, as concluded in the previous chapter). We demonstrated the potential of NPROs for modeling IAV infection in neutrophils, to gain novel insights into the relationships and interactions between IAV virulence and neutrophils. In particular, we showed that IAV was able to infect NPRO cells by successfully labelling with DiD for tracking individual IAV in living cells [170]. Notwithstanding that mature and immature neutrophils could be infected, IAV infection in NPROs was abortive and induced apoptosis. Of interest, we revealed that NPROs infected with LVI and HVI elicited mainly type I IFN responses, but with differences in kinetics.

Typically, type I and type II IFNs confer protection against viruses and bacteria, respectively. The protective function of type II IFN-γ produced by mature and immature neutrophils early after infection has been demonstrated in infection
models of *Listeria monocytogenes* [171] and *Streptococcus pyogenes* [172]. However, the protective role of type I IFN-α/β produced by neutrophils remains unclear. Despite their potent antiviral and antibacterial activities, over-activation of type I and II IFNs may also be detrimental. Over-activation of neutrophil IFN-γ stimulated by repetitive exposure to inhaled particulate antigents contributes to the pathogenesis of hypersensitivity pneumonitis [173]. Neutrophils have also been attributed as the major source of excess IFN-α in patients with lupus. Recent reports demonstrated that chromatin could induce neutrophils to produce IFN-α [174] and IFN-α, on the other hand, could prime neutrophils to release NETs [86]. This suggests the presence of a feedback loop that may cause the over-activation of IFN-α and at the same time, increase the abundance of NETs observed in patients with lupus. The protective role of IFN-β produced by neutrophils, or its detrimental effects related to its over-or under-activation are still not completely understood. Nevertheless, over-activation of genes downstream of IFN-α/β in neutrophils, which was more pronounced than that in other cell types (including macrophages, CD4 and CD8 T cells), has been suggested to contribute to the pathogenesis of tuberculosis [175].

Our infection model of NPROs with HVI revealed rapid activation of mostly type I IFN-inducible genes, including the ubiquitin-like modifier ISG15, an antiviral gene that was noticeably highly expressed at 12 h after HVI infection of mice (shown in FIG. 2.5 in Chapter 2). However, not all type I IFN-inducible genes were differentially expressed in IAV infection of NPROs. For example, we did not observe any change in expression of the MX1 gene, which is implicated in many *in vitro* and *in vivo* influenza models, and determines the effectiveness of type I IFN against IAV [176]. Nevertheless, the rapid IFN response in neutrophils may represent a potent contributor to the dysregulation of innate immunity and pathogenesis of HVI-infected
animals. In influenza models, activation of type I IFN signaling sensitizes hosts to secondary bacterial pneumonia infection due to impairment of neutrophil responses [177], and potentiates virus-induced apoptosis via FADD/CASP8 death signaling pathway [178].

The rapid response of the type I IFN pathway in HVI was concurrent with the early expression of the viral dsRNA sensing molecule, DDX58 or RIG-I. This may highlight the importance of viral factors in accelerating the type I IFN response. Interestingly, the rapid response was accompanied by the repression of IFN-β gene expression. This repression is likely due to the synthesis of viral NS1 protein that inhibits IFN-β transcriptional induction via various mechanisms, such as disrupting the activities of DDX58 [179], NF-κβ [180], IRF3 [181], IRF7 [182] and TRIM25 [183]. Furthermore, one noteworthy note is that the rapid type I IFN response of NPROs to HVI could not be explained by its higher replication rate in MDCK cells alone: the replication kinetics of HVI and MPI in MDCK cells were similar, but the type I IFN response in MPI was relatively slower. Hence, the differences in gene expression kinetics of NPROs infected by HVI and MPI signify the importance of comparing the original highly virulent IAV isolate in murine lung homogenate against that propagated in MDCK cell lines. In view of the real-life scenario of IAV infection, this reflects the role of antigenic drift that might have consequences in changing host response to the virus.

The expression levels of several genes encoding cytokines and chemokines were upregulated in neutrophils in response to IAV infection. Of particular interest is CXCL10, an IFN-γ-inducible protein that acts as a potent neutrophil chemoattractant [184], and is highly expressed in various animal models of infection with highly virulent IAV. In highly pathogenic H5N1 infection of ferrets, inhibition of CXCR3 (a
cognate receptor of CXCL10) reduces lung infiltration and delays mortality [4]. Interestingly, in mice infected with HVI, CXCL10 is not elevated, whereas neutrophil chemoattractant CXCL1 is highly expressed even at 12 h postinfection (FIG. 2.3 in Chapter 2). In spite of this difference, CXCL1 and CXCL10 potently form a chemokine-driven feedforward circuit that recruits more and more neutrophils to infection site and triggers the neutrophils to release similar chemoattractants. Evidence for such feedforward circuit has been provided, but driven by CXCL2 in the case of PR8 H1N1 infection [204].

In addition, we also demonstrated that IAV infection induces NPRO apoptosis. Specifically, the annexin V and TUNEL assays confirmed that MPI-infected NPROs undergo cell death via apoptosis that was significantly higher than that in uninfected samples, particularly at 6-24 h postinfection. These apoptosis results were complemented by transcriptomic data of infected NPROs that revealed differential regulation of apoptosis-related gene, including IFN-inducible EIF2AK2 (eukaryotic translation initiation factor 2-alpha kinase 2), and genes involved in retinoic acid-mediated apoptosis signaling. Significantly, HVI infection triggered more rapid over-expression of these apoptosis-related genes (at 3 h postinfection) in contrast to MPI and LVI infections (at 9 h postinfection). In other reports, influenza-induced neutrophil apoptosis has been associated with the increased neutrophil expression of Fas antigen and Fas ligand [20]. The respiratory burst may also play a role in the induction, but significant contribution was only observed after simultaneous incubation with influenza virus and S. pneumoniae [19]. The induction of neutrophil apoptosis by influenza virus may also be induced by viral proteins, especially multifunctional NS1 [194]; hence, abortive infection in neutrophils may still induce apoptosis. Interestingly, the course of influenza-induced neutrophil apoptosis seems to
be compatible with other cellular or viral gene/protein expression, as suggested by results of MPI experiments and report in [20]. We also note that host-associated factors may contribute to neutrophil apoptosis induction only if the cells are infected; host factors expressed from other infected or uninfected cells, e.g. MDCK host-associated proteins in MPI inoculum, may not be enough to induce neutrophil apoptosis. This has been demonstrated for IFN-β, which is able to induce apoptosis of IAV-infected fibroblast cells, but not apoptosis of uninfected cells [203].

Finally, as presented in Appendix B, we have compared the list of differentially expressed genes from our study with the list identified by reanalyzing microarray data of murine macrophages infected with highly virulent PR8 (H1N1) or less virulent FJ02 (H3N2) viruses collected by Zinman et al [10]. The reanalysis revealed 287 differentially expressed genes in macrophage experiment, but significant changes were mainly observed in FJ02, but not PR8, infected macrophages – consistent with the analysis result reported in [10]. Interestingly, only 18 genes were found in the overlap between differentially expressed genes in NPROs and macrophages. These genes are mainly IFN-related genes, which include DDX58, GBP2, GCH1, IFI204, IFIH1, IFIT3, IRF1, ISG15, OASL1 and RSAD2. Furthermore, GO analysis revealed that unlike infected neutrophils, infected macrophages also regulate genes associated with cell proliferation in addition to genes related to immune and defense responses. Thus, the comparative study mainly indicates that IAV infection could elicit IFN response in both neutrophils and macrophages, and that macrophages could elicit a broader genomic response.

Overall, our model of IAV infection using NPROs has offered new insights into the roles of neutrophils during the pathogenesis of severe influenza. In particular, we propose that the rapid type I IFN responses of neutrophils infected with HVI may
contribute to dysregulation of innate immunity in the lungs. For future studies, it may be worthwhile to investigate gene expression responses of neutrophils, in comparison with macrophages, to different strains of highly virulent IAV in order to evaluate the contribution of both cells to innate immune dysfunction. Furthermore, we may also use NPROs to investigate factors that cause abortive infection of IAV, which to some extent have been investigated in dendritic cells – showing that it is likely due to defect in virus assembly rather than failure in the release of the virus from cell surface [202]. Another area of considerable interest would be to explore neutrophil defects attributed to IAV infection, including demonstrating the impact on phagocytosis and developing a secondary bacterial infection model. Lastly, this model may also be exploited in future studies on the interactions of neutrophils with other related pathogens or diseases.
Chapter 5

Conclusion and future prospects

5.1 Conclusion

Our studies have succeeded in advancing the understanding of global host response to highly and less virulent IAV infections. We highlight the roles of TREM1 and IFN signalings in neutrophils in promoting dysregulation of innate immune response in lungs infected with highly virulent IAV. The dysregulation is linked to defects in lung homeostasis and repair via LPS/IL1-mediated inhibition of RXR function pathway. Finally, we show that comparatively enhanced expression of apoptosis-related genes was elicited in lungs and neutrophils early after highly virulent IAV infection.

In our studies, we have also demonstrated the usability of NPROs as a neutrophil source for modeling IAV infection in neutrophils. Hence, we promote the use of NPROs for future neutrophil studies, not only in association with IAV, but also with other pathogens. Moreover, we contribute to the development of an automated image analysis system to quantify neutrophil yields in giemsa-stained NPROs. The automated system will be helpful in executing experiments that attempt to optimize neutrophil yields in NPROs.

Altogether, insights from our infection model in animals and NPROs may be useful for directing future influenza and neutrophil research, as well as the
development of host-oriented drug discovery for influenza. Nevertheless, the outcomes of our studies still need to be verified with further *in vitro* and *in vivo* experiments. These include the verification of the level of the proteins in the signaling subnetworks and the response of normal neutrophils (isolated from blood) to influenza infection. One major concern of using NPROs is the impurity of the neutrophils, and hence, experiments with purer neutrophils are desirable. Another important note is that the outcomes were based on mouse models and thus, their generalization to human models must be made with precautions. Regardless of this concern, mouse model will continue to provide biological insights about influenza since it allows us to perform research that is often not possible to be done with human.

5.2 Future prospects

5.2.1 Large scale investigations on systems-level host response to IAV infection

Recently, large scale genomic studies on host response to IAV infection have become more popular as microarray experiments get cheaper and computational tools for data analysis are more established. These studies collect microarray data from control (mock-infected) and IAV-infected samples (and other samples infected with other pathogens) at several time-points postinfection, or alternatively, from different level of biological systems [10, 11]. For future studies, large scale genomic studies that dissect global response of different types of cells to IAV infection, and global host response to infection of IAVs of different subtypes and virulence, are warranted. In relation to dysregulation of innate immune response observed during highly virulent IAV infection *in vivo*, large scale genomic studies on innate immune cells
(mainly neutrophils and macrophages) infected with IAVs of varying subtypes are desired. These studies may reveal the contribution of each cell to the dysregulation, and specifically, elucidate whether each cell respond differently to different subtype of IAV or not.

Next, elaborating various ‘omic’ (including genomic, proteomic, glycomic and lipidomic) approaches to understand how cellular responses are integrated has been a major goal in systems biology studies. In this case, neutrophil provides an attractive model for signaling pathways since the cells are terminally differentiated and not encumbered by machinery necessary for cellular replication. Our genomic studies with NPROs have also demonstrated this, where mainly type I IFN-related genes are over-expressed during IAV infection. Hence, in addition to genomic (transcriptomic) studies, it will also be interesting to employ other ‘omic’ approaches to obtain a complete view on the molecular basis of neutrophil response to IAV infection.

**5.2.2 Using NPROs for modeling secondary bacterial infection in neutrophils**

Secondary bacterial infection contributes to increased morbidity and mortality following IAV infection. The increased susceptibility to a secondary bacterial infection has been partly attributed to influenza-induced neutrophil dysfunction, as demonstrated *in vivo* by McNamee and Harmsen [190]. This is parallel to *in vitro* studies showing that phagocytosis and killing of the bacteria by neutrophils infected with IAV are reduced [20, 191-193]. Hence, in addition to developing a model of primary IAV infection in neutrophils, it is also interesting to use NPROs for modeling secondary bacterial infection following IAV infection. One study that might be of interest is to investigate the relationships between activation of type I IFN-inducible genes and reduction in bacterial killing by neutrophils. This proposal is based on
previous *in vivo* experiments showing that influenza-induced type I IFN response sensitize hosts to secondary bacterial infection [177].

### 5.2.3 Obtaining NPROs with high neutrophil yields

The optimal quantity of neutrophils obtained in NPROs is currently about 20-30% on average. For some experiments, e.g. investigation on the release of NETs, NPROs containing purer neutrophils are warranted. Hence, efforts that focus on developing technique to increase neutrophil yields in NPROs are required. One method that may be promising is to add a mitotic inhibitor into the complete media for inducing MPRO differentiation (e.g., Ara-C, which has been shown to partially induce neutrophil differentiation from FDCP-mix cell line expressing mutant RAS [206]). In addition, a technique for isolating neutrophils from NPROs may also be considered, e.g. by using Percoll density gradient described in [201].

For the above attempt, the use of an automated image analysis system to estimate neutrophil purity in giemsa-stained NPROs provides a cheap and convenient alternative, especially for handling a large number of samples that are likely to be the case during the optimization of neutrophil yields in NPROs. In this thesis, we propose an automated system that performs the task, and demonstrate that it gives comparable estimates to manual quantification. Nevertheless, further improvement of the automated system may be considered, e.g. extracting more cell features and incorporating more sophisticated algorithms for segmentation and classification.

### 5.2.4 Host-oriented drug discovery for influenza

The results from our genomic studies may become the basis of host-oriented drug discovery for severe influenza. In particular, we suggest the use of drugs that
inhibit signals from TREM1 and IFN receptors in infected lungs such that innate immune response is suppressed. The use of growth factors (e.g. keratinocyte growth factor, fibroblast growth factor, and hepatocyte growth factor), which has been suggested for treatment of lung damage caused by toxic compounds [207] and could induce the proliferation of alveolar epithelial cells [154], may be considered to promote lung repair. Moreover, combining multiple drugs to target the virus and/or host pathways involved in influenza progression is also highly encouraged.

In relation to NPROs, a high content screening (HCS) approach can be utilized to identify substances or drugs with desired effects on neutrophils during IAV infection. Some possible objectives include the identification of molecules that prevent neutrophil defects following IAV infection, suppress type I IFN response elicited by neutrophils, and stimulate/suppress the release of NETs.

### 5.3 Final remarks

This thesis delivers novel insights on systems-level host response to IAV infections. In particular, the use of NPROs provides unprecedented knowledge about the interaction between neutrophils and IAV: neutrophils rapidly elicit type I IFN response following a highly virulent IAV infection. This event that may contribute to the dysregulation of innate immune response observed in vivo, especially due to the modulation of cytokine/chemokine expression. Furthermore, an automated image analysis system that can help to quickly quantify neutrophil purity in giemsa-stained NPROs is introduced to anticipate a larger number of samples being quantified. Finally, we hope that the thesis, as a whole, will have an impact on future research related to influenza and neutrophils, as well as improve the strategy in treating the disease.
References


34. Nelson MI, Viboud C, Simonsen L, Bennett RT, Griesemer SB, St George K, Taylor J, Spiro DJ, Sengamalay NA, Ghedin E, Taubenberger JK, Holmes EC,


References


References


distinct features of Streptococcus pneumoniae and influenza A virus infection,”


References


encodes a member of the novel uteroglobin-related protein gene family,” *Cytogenetic and Genome Research*, vol. 97, no. 1-2, pp. 120-127, 2002.


References


Appendix A

Comparative analysis on global responses of murine lungs infected with different strains of IAVs

To uncover whether biological processes and pathways (subnetworks) observed in HVI (or LVI) were also elicited in mouse lungs infected with different highly (or less) virulent IAV strains, we performed a meta-analysis involving our microarray data and other relevant microarray datasets downloaded from the Gene Expression Omnibus (GEO) database. These include microarray dataset from Cilloniz et al [195] for VN1203 and r1918 (GEO series GSE48890), Go et al [208] for 2009 pandemic H1N1 (GSE40092), and Brandes et al [204] for PR8 and TX91 (GSE42638). The gene expression in the GEO datasets has been represented as log-normalized intensity values. We focused on expression data from mock and IAV infected wild type mice that were collected on or before day 4 postinfection. For PR8 dataset, we focused on expression data associated with lethal infection dose of 100LD50.

For each microarray dataset, probe IDs were converted into MGI gene symbols. If probe IDs map to multiple genes, then their log-ratio measurements were duplicated for each gene. Probe IDs not mapping to any gene symbol were excluded. Next, the log-ratio measurements for the same gene symbols were averaged. To identify genes with a significant infection effect, two-way ANOVA (for single-
channel microarray data, i.e., GSE40092 and GSE42638) or linear model and empirical Bayes methods [210] (for two-channel microarray data, i.e., GSE48890) have been performed. The p-values of the tests were adjusted using Storey method [137].

Next, we applied the Fisher’s method for combining p-values implemented in MADAM package [186] to identify genes with significant infection effect across the experimental studies (including our microarray experiments for HVI and LVI). Benjamini-Hochberg method was used to correct for multiple testing. Biological and technical replicates for log intensity values were averaged. In what follows, we focused on the genes that had adjusted combined p-values < 0.05 and showed an absolute log ratio of intensity (difference between log intensity values in IAV-infected versus mock-infected group) greater than a threshold of 0.6 in any infected group and at any time-point.

The heatmaps that compare the expression of genes associated with biological processes and subnetworks activated in the murine lungs infected with various strains of IAV are shown in FIG. A1 and A2. Specifically, FIG. A1 compares the expression of genes associated with cytokine activity, IFN response and apoptotic activity (the genes shown in FIG. 2.3, FIG. 2.5 and 2.6, respectively), whereas FIG. A2. compares the expression of genes in the TREM1 subnetwork (shown in FIG. 2.4B) and genes in the metabolic and developmental subnetwork (shown in FIG. 2.7). In addition, as shown in FIG. A3, we also present a heatmap for candidate signature genes that may differentiate highly versus less virulent IAV infections on day 3-4 postinfection. Discussion on these heatmaps could be found in section 2.4 in Chapter 2.
FIG. A1. Heatmap of expression values of genes associated with cytokine activity, IFN response and apoptotic activity before or on day 4 postinfection with various IAV strains (highly virulent: HVI, PR8, r1918 and H5N1; less virulent: CA04 (2009 pandemic H1N1), TX91 and LVI). Rows represent genes in the same order as in FIG. 2.3 (for cytokine activity), 2.5 (IFN response) and 2.6 (apoptotic activity). White expression level indicates that the genes were not represented in the associated studies or not identified as differentially expressed genes across the experimental studies. (d: day)
FIG. A2. Heatmap of expression values of genes in the TREM1 subnetwork (shown in FIG. 2.4B) and metabolic and developmental subnetworks (shown in FIG. 2.7) before or on day 4 postinfection with various IAV strains (highly virulent: HVI, PR8, r1918 and H5N1; less virulent: CA04 (2009 pandemic H1N1), TX91 and LVI). White expression level indicates that the genes were not represented in the associated studies or not identified as differentially expressed genes across the experimental studies. (d: day)
FIG. A3. Heatmap of expression level of candidate signature genes on day 3–4 postinfection with various IAV strains (highly virulent: HVI, PR8, r1918 and H5N1; less virulent: CA04 (2009 pandemic H1N1), TX91 and LVI). White expression level indicates that the genes were not represented in the associated studies or not identified as differentially expressed genes across the experimental studies. (d: day)
Appendix B

Comparative analysis on global responses of murine neutrophils versus macrophages infected with IAVs

To investigate whether NPROs (neutrophils) and macrophages respond differently to IAV infection, we reanalyzed the raw microarray dataset collected by Zinman et al [10], which can be downloaded from ArrayExpress (accession number: E-MTAB-427). We focused on microarray data related to uninfected and IAV (PR8 (H1N1) and FJ02 (H3N2)) infected murine alveolar macrophages at 1, 2 and 6 hours postinfection. Arrays whose expression data were not comparable with the majority were not included in the analysis (4 arrays were removed, leaving 29 arrays for our analysis). Following background correction and between-array normalization, the linear model and empirical Bayes methods [210] were used to identify genes with a significant infection effect. In what follows, we only consider the 287 genes that had p-values < 0.05 and showed an absolute log ratio of intensity (difference between log intensity values in IAV-infected versus mock-infected group) greater than a threshold of 0.6 in any infected group and at any time-point. These genes, consistent with the report in [10], were mainly differentially expressed following FJ02 infection, but not following PR8 infection. The heatmap of expression values of differentially expressed (DE) genes identified from the reanalysis and our neutrophil study is shown in FIG.
B1. The overlap between DE genes in these two studies consists of BC006779, CASP4, CCL4, CCL5, DDX58, GBP2, GCH1, IFI204, IFIH1, IFIT3, IRF1, ISG15, OASL1, PARP14, RSAD2, SAMD9L, USP18 and ZC3HAV1. Outcomes from this appendix are discussed further in section 4.4 in Chapter 4.

**FIG. B1.** Heatmap of expression values of differentially expressed (DE) genes in NPROs (neutrophils) and macrophages following infection with IAVs (HVI, LVI and MPI for NPROs, and PR8 and FJ02 for macrophages). White expression level indicate that the genes were not represented or not identified as differentially expressed genes in the associated study. (h: hour).