# Exploring biomarkers to predict pig disease resilience traits under a natural disease challenge

by

Ziqi Yang

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Animal Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

# Abstract

The intensification and consolidation of modern pig production is exposed to higher risks of endemic or pandemic infections. The complexity of the polymicrobial challenge and increasing concerns on antibiotics resistance make it pivotal to find an efficient way of controlling infections besides using vaccines. Breeding for disease resilience (maintaining productive performance during pathogen infections) could be a solution to circumvent this problem. This study was focused on three types of biological information in blood: serum acute phase proteins (APPs), whole blood transcriptome, and serum metabolome, which were reported with the potential for disease diagnosis and livestock production assessment. It is unknown whether they could be used to predict pig disease resilience before exposure to pathogens. The feasibility was tested using those molecules separately to identify biomarkers associated with disease resilience in a natural disease challenge model which simulates the polymicrobial environments in commercial farms. Identification of such biomarkers could help characterize disease resilience and provide a theoretical guide for commercial pig breeding.

Plasma concentrations of alpha-1 acid glycoprotein (AGP), haptoglobin (HP), and C-reactive protein (CRP) were determined in 60 pigs before and after challenge using ELISA. The resilient pigs had a relatively low level of AGP in plasma before challenge. The concentrations of HP and CRP, but not AGP, were induced dramatically upon challenge in all the groups of pigs. Resilient pigs showed a slow response of both HP and CRP at the early stage of the challenge but had a sharp increase of CRP at the later stage. Correlation analysis between APPs and various performance traits suggested that they are

more proper for assessing rather than predicting pig health and productivity under challenge even though AGP concentration before challenge showed some trends to correlate with productivity-related traits.

Fifty eight pigs were chosen from their phenotypes after challenge and the whole blood transcriptome before challenge was determined by RNA-Seq. They were grouped into 4 groups (Resistant, Resilient, Susceptible, and Early_dead) in response to the natural disease challenge. Only two significant transcripts from the differential expression (DE) analysis were found higher in the Susceptible group compared to the others (q-value < 0.1). They were mapped to the *IgC* gene and the *SLAMF9* gene, respectively. A larger cohort of 209 pigs was utilized for validating the findings. Results from both of the cohorts supported a hypothetical hierarchical model for the baseline immunity: Resistant≧Resilient>Early_dead>Susceptible. The larger cohort with samples post disease challenge revealed that all the pigs activated innate immune response early after infection. Of note, the Resilient group exhibited a unique strategy of restricting the potency and energy expenditure of the immune response, implying that the resilient pigs maintain the productive performance under disease challenge via consuming less energy from their immune system. However, the results suggest that it may not be feasible using pre-challenge whole blood transcriptome to identify biomarkers for disease resilience in the context of this study.

Plasma metabolomic profiles of 460 healthy pigs were determined by NMR spectroscopy before the natural disease challenge, and the pigs were then divided into four groups as in the RNA-Seq experiment, or groups defined by single or double trait records. Succinate and dimethylglycine in unchallenged pigs were found with

significantly higher concentrations in the Early_dead group compared to the others. However, batch effect was found as the major causative factor for the metabolome variations, and pyruvic acid was found as the only hit with significantly lower concentration in the Early_dead group than the others. Machine learning was performed to test whether an integrated metabolite profile could be used to predict the pig phenotypes, but prediction accuracy was far from ideal. Together, a snapshot of the plasma metabolome only provided a limited prediction value for pig resilience, probably due to the prominent impact of the batch factor.

In summary, this thesis created a framework for investigating proxy traits to assess or predict pig resilience, examined the viability of using various molecular information, including APPs, transcriptome, and metabolome derived from peripheral blood to predict pig disease resilience phenotype in a natural disease challenge model. Additionally, it improved knowledge of potential molecular processes influencing how pigs differentially respond to polymicrobial challenges, shedding light on how to use disease resilience as a breeding objective to meet the rapidly expanding demand for healthy pork products.

# Preface

This thesis is an original work by Ziqi Yang. No part of this thesis has been previously published. The research conducted for this thesis forms part of a collaborative project (Project name "Alberta Livestock Genome Program–ALGP2") which follows the Canadian Council on Animal Care guidelines. The project was fully under the oversight of the Centre de Développement du Porc du Québec and the herd veterinarian together with project veterinarians. This project was funded by Genome Canada, Genome Alberta, PigGen Canada, Swine Innovation Porc, and Alberta Agriculture and Forestry. PigGen Canada was also intimately involved in the design of the project and the development of protocols for the natural disease challenge model used in this study. This work is also part of the AMR – One Health Consortium, funded by the Major Innovation Fund program of the Alberta Ministry of Economic Development, Trade and Tourism. The animal protocol used in this study was reviewed and granted approval by the Animal Protection Committee of the Centre de Recherche en Sciences Animales de Deschambault (15PO283) and the Animal Care and Use Committee at the University of Alberta (AUP00002227).

The project and protocols for the natural disease challenge model were designed by Fortin, Harding, PigGen Canada, Dyck, Dekkers, and Plastow. Fortin oversaw the sample collection and scheduling. Harding, was in charge of veterinary oversight on the project. For the data chapters Chapter 2, 3, 4 of this thesis, I designed, performed the technical experiments and data analysis under the guidance of the project principal investigator Plastow. For CBC data acquisition and plasma isolation, Xuechun Bai, Jiehan Lim, and

myself participated in the blood sample processing and measurement, and Professor Catherine Field, provided equipment and technical support. Plastow was in charge of the database and genotyping for the project.

# Acknowledgments

First and foremost, I would like to extend my heartfelt gratitude to my supervisor, Dr. Graham Plastow, for his unwavering support throughout my PhD journey and life in Canada. His patient guidance not only facilitated the construction of my knowledge but, even more importantly, enlightened me on how to think independently and approach research questions with confidence. Whenever I encountered challenges during my experiments, Dr. Plastow always emphasized my progress first and offered warm care and understanding during moments of fatigue and stress. His encouragement served as an empowering force, enabling me to face challenges and explore new scientific possibilities with zeal. Engaging in scientific discussions with him was a delight, as his erudition and sense of humor were not only insightful but also made the experience enjoyable. I am deeply grateful for all of this and more, as it significantly contributed to a wonderful training experience. Dr. Plastow's professional guidance will undoubtedly have a profound impact on my future career and life.

I also want to express my deep gratitude to my supervisory committee members, Dr. Leluo Guan and Dr. Catherine Field. Their invaluable advice played a critical role in the design, interpretation, and troubleshooting of my experiments. I truly treasure their constructive criticisms, as they helped refine my scientific logic and corrected my experiment directions.

Special thanks to Dr. Paul Stothard who opened the window for me to process my RNA-Seq data from his course. I am grateful for the insightful discussions with him on the technical challenges I encountered during my analysis. I would also like to express my appreciation to Susan Goruk and Dr. Marnie Newell for their training and assistance in processing the blood samples to obtain the peripheral cell data for my study.

My gratitude extends to my lab colleagues and friends: Yan Meng, Dr. Xuechun Bai, Dr. Tianfu Yang, Dr. Elda Dervishi, Jiehan Lim, and Dr. Jiyuan Li. Their help, advice, and support have been invaluable throughout my research journey.

Lastly, but by no means least, I would like to acknowledge my husband, Kang Yu. His patience in helping me edit my thesis and discussing my experiment results has been immeasurable. His love and tolerance have provided consistent strength during the challenges I faced as an international student on this arduous road.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| **$^{1}$H NMR** | Proton NMR |
| **AGP** | Alpha-1 Acid Glycoprotein |
| **ANOVA** | Analysis Of Variance |
| **APC** | Antigen-Presenting Cells |
| **App** | *Actinobacillus pleuropneumoniae* |
| **APPs** | Acute Phase Proteins |
| **APR** | Acute Phase Response |
| **ARDS** | Acute Respiratory Distress Syndrome |
| **ASCA** | Anova-Simultaneous Component Analysis |
| **ASE** | Allele-Specific Expression Analysis |
| **ASFV** | African Swine Fever Virus |
| **Bas** | Basophils |
| **BLUPF90** | Best Linear Unbiased Prediction For Family 90 |
| **CAGE** | Cap Analysis Of Gene Expression |
| **cAMP** | Cyclic Adenosine Monophosphate |
| **CBC** | Complete Blood Count |
| **CCAC** | Canadian Council On Animal Care |
| **cDNA** | Complementary DNA |
| **CDPQ** | Centre De Développement Du Porc Du Québec Inc. |
| **CE** | Capillary Electrophoresis |
| **COVID-19** | Coronavirus Disease 2019 |
| **CRP** | C-Reactive Protein |
| **CRSAD** | Centre For Research In Animal Science Deschambault |
| **DE** | Differential Expression |
| **DETs** | Differential Expression Transcripts |
| **DEGs** | Differentially Expressed Genes |
| **DSS** | Sodium 2,2-Dimethyl-2-Silapentane-5-Sulfonate |

| | |
|---|---|
| **DT** | Decision Tree |
| **EDTA** | Ethylene Diamine Triacetic Acid |
| **ELISA** | Enzyme-Linked Immunosorbent Assay |
| **Eos** | Eosinophils |
| **eQTL** | Gene Expression Quantitative Trait Locus |
| **ER** | Endoplasmic Reticulum |
| **EST** | Expressed Sequence Tags |
| **FI** | Feed Intake |
| **FPKM** | Fragments Per Kilobase Of Transcript Per Million Mapped Reads |
| **FT-IR** | Fourier-Transform Infrared Spectroscopy |
| **GBLUP** | Genomic Best Linear Unbiased Prediction |
| **GC** | Gas Chromatography |
| **GC/MS** | Gas Chromatography/Mass Spectrometry |
| **GO** | Gene Ontology |
| **GR** | Globin Reduction |
| **GWAS** | Genome-Wide Association Study |
| **HBA** | Alpha Hemoglobin |
| **HBB** | Beta Hemoglobin |
| **HMDB** | Human Metabolome Database |
| **HN** | Healthy And Normal Weight |
| **HO** | Healthy And Overweight |
| **HP** | Haptoglobin |
| **HU** | Healthy And Underweight |
| **IF** | Immunofluorescence |
| **IFN** | Interferon |
| **IHC** | Immunohistochemistry |
| **IL** | Interleukin |
| **IL-1** | Interleukin-1 |
| **IL-6** | Interleukin-6 |

| | |
|---|---|
| **ISH** | In Situ Hybridization |
| **KNN** | K-Nearest Neighbors |
| **LC** | Liquid Chromatography |
| **LC/MS** | Liquid Chromatography/Mass Spectrometry |
| **LD** | Linkage Disequilibrium |
| **LR** | Logistic Regression |
| **Lym** | Lymphocytes |
| **MAF** | Minor Allele Frequency |
| **MDA** | Mean Death Age |
| **MDW** | Mean Death Weight |
| **mGWAS** | Genome-Wide Association Studies With Metabotypes |
| **MHCI** | Major Histocompatibility Complex I |
| **Mhyo** | *Mycoplasma hyopneumoniae* |
| **Mono** | Monocytes |
| **MPSS** | Massively Parallel Signature Sequencing |
| **MPSW** | Mean Pre-Slaughter Weight |
| **MS** | Mass Spectrometry |
| **MSA** | Mean Slaughter Age |
| **MTN** | Mean Treatment Number |
| **Neu** | Neutrophils |
| **NGS** | Next-Generation Sequencing |
| **NK** | Natural Killer |
| **NMR** | Nuclear Magnetic Resonance |
| **OPLS-DA** | Orthogonal Projection To Latent Structure-Discriminant Analysis |
| **PBMCs** | Peripheral Blood Mononuclear Cells |
| **PCA** | Principal Component Analysis |
| **PCR** | Polymerase Chain Reaction |
| **PCV1** | Porcine Circovirus Type 1 Virus |
| **PCV2** | Porcine Circovirus Type 2 Virus |

| | |
|---|---|
| **PCV2-SD** | Porcine Circovirus Type 2 Virus -Systemic Disease |
| **PEDV** | Porcine Epidemic Diarrhea Virus |
| **PHGC** | PRRS Host Genetics Consortium |
| **Pig-MAP** | Pig-Major Acute Phase Protein |
| **PLS-DA** | Projection To Latent Structures Discriminant Analysis |
| **PMWS** | Postweaning Multisystemic Wasting Syndrome |
| **PRDC** | Porcine Respiratory Disease Complex |
| **PRRS** | Porcine Reproductive And Respiratory Syndrome |
| **PRRs** | Pattern Recognition Receptors |
| **PRRSV** | Porcine Reproductive And Respiratory Syndrome Virus |
| **qPCR** | Real-Time Quantitative PCR |
| **QTL** | Quantitative Trait Locus |
| **RBC** | Red Blood Cells |
| **RDA** | Range Of Death Age |
| **RDW** | Range Of Death Weight |
| **RELA** | Rel-Like Domain-Containing Protein A |
| **RF** | Random Forest |
| **RNS** | Reactive Nitrogen Species |
| **ROC** | Receiver Operating Characteristic |
| **ROS** | Reactive Oxygen Species |
| **RPSW** | Range Of Pre-Slaughter Weight |
| **RSA** | Range Of Slaughter Age |
| **RTN** | Range Of Treatment Number |
| **RT-PCR** | Reverse Transcriptase PCR |
| **SAA** | Serum Amyloid A |
| **SAGE** | Serial Analysis Of Gene Expression |
| **SAM** | Significance Analysis Of Microarray |
| **SARS-CoV-2** | Severe Acute Respiratory Syndrome Coronavirus 2 |
| **SIV** | Swine Influenza Virus |

| | |
|---|---|
| **SLAMF9** | Signaling Lymphocytic Activation Molecule Family Member 9 |
| **SN** | Sick And Normal Weight |
| **SO** | Sick And Overweight |
| **SPF** | Specific-Pathogen-Free |
| **SU** | Sick And Underweight |
| **SVM** | Support Vector Machine |
| **TGEV** | Transmissible Gastroenteritis Virus |
| **TNF-α** | Tumor Necrosis Factor-Alpha |
| **TUNEL** | Terminal Deoxynucleotidyl Transferase Dutp Nick End Labeling |
| **WBC** | White Blood Cells |

# List of Appendices

# Chapter 1: Introduction

## 1.1 Overview of the pig industry and common disease

### 1.1.1 The global pig sector

Pork is one of the most popular and consumed types of livestock meat in the world. Pig production is global except for some regions with cultural and religious reservations regarding the consumption of pork. Global patterns of meat production have changed over time since the 1960s (Figure 1.1). Pork production is constantly increasing as China, the country with the largest population in the world, traditionally consumes pork as the main type of meat. With the rapidly increasing demand for pork, the form of pork production has gradually shifted from family farms to large corporations, China became the largest pork-producing country after the 1970s and led to worldwide pork consumption gradually surpassing beef (Figure 1.1) (Brown, 2012).

### 1.1.2 How pig production is impacted by disease

The economy of modern pig production farms is affected by numerous factors. Some factors cannot be controlled by the farmers, such as market price fluctuations, and strict policies regarding food safety and quality assurance (Agriculture and Agri-Food Canada, AAFC). Nevertheless, many aspects can be controlled by the farmers'

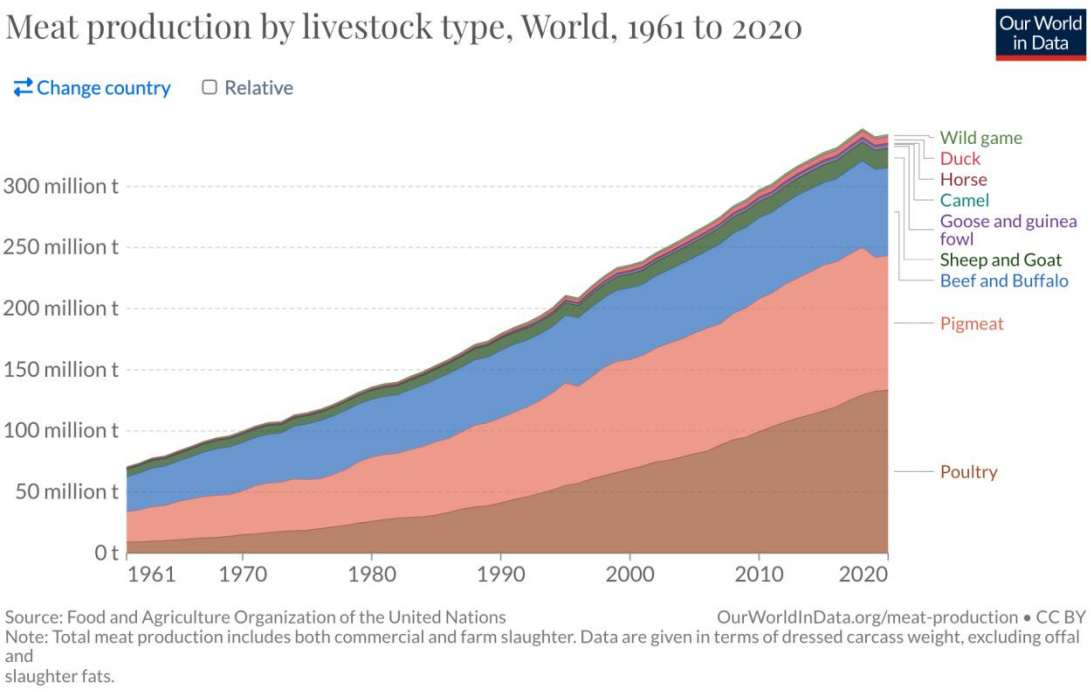



**Figure 1.1** World meat production of livestock from 1961-2020 (Ritchie et al., 2017).

management strategy to improve productivity, such as disease control and treatment. Currently, pig farming operations have been transformed into more intensive and specialized industries to meet the requirement of efficient management and production (Brisson, 2014). However, this economic scale will also lead to large-scale disease outbreaks and potential infections. For example, by the Porcine Respiratory Disease Complex (PRDC), which is caused by multi-factorial bacterial and viral pathogens including Porcine Reproductive and Respiratory Syndrome virus (PRRSV), Swine Influenza virus (SIV), and *Mycoplasma hyopneumoniae* among others (Martelli & Segalés, 2021).

The occurrence of disease has a direct negative impact on the economy of pig farms. For example, in 2005, the annual loss caused by PRRSV in the US was estimated at $561 million in the whole swine industry (Neumann et al., 2005). In 2013, these calculations were updated when the current clinical effects of PRRSV, pig inventory, production systems, and performance were taken into account, and the total annual loss due to PRRSV in the US was increased to $664 million (X. Li et al., 2013). The direct negative impacts caused by PRRSV outbreaks on pig farms consist of many components including mortality, slow growth, abortion, etc. However, the most intuitive impact on consumers is a sharp increase in pork market prices. For example, according to the information from the US Department of Agriculture, the average price of Spiral ham was $3.08 in 2014 which increased by 17% from $2.63 in 2013 possibly induced by Listeria contamination. Moreover, the outbreak of the Porcine Epidemic Diarrhea virus (PEDV) from 2013 to 2014 wiped out as many as seven million pigs in 32 states of the US and pushed the pork price to a higher record although PEDV does not sicken humans or impact the food safety

of pork products (Schulz & Tonsor, 2015). While it may take some time to enhance the treatment efficacy and cost of the commercial medications, strategies for prevention and treatment of such viral infections are being explored, including vaccinations and antiviral drugs (Kimman et al., 2009; Song & Park, 2012; Zuckermann et al., 2007). Given the fact of exposure to multiple pathogens can be challenging for the swine industry, recent studies were focused on characterizing a phenotype called "disease resilience". Instead of trying to reduce the infection rate and treat the infected pigs, it aims to minimize economic loss by sustaining pigs' regular production performance per se when they are confronted with those pathogens (Guy et al., 2012). This thesis tested the feasibility to predict disease resilience phenotype using different types of biological information and the detailed approaches are described in Chapter 2, 3, and 4.

## 1.1.3 Common pig pathogens

### 1.1.3.1 PRRSV

PRRSV is considered as a major pig pathogen with a worldwide distribution (Lunney et al., 2016). It is a single-stranded RNA virus belonging to the *Arteriviridae* (Benfield et al., 1992; Wensvoort et al., 1991). As indicated by its denomination, the pathogenesis of PRRSV mainly involves respiratory disease in newborn and growing pigs while inducing reproductive problems in pregnant sows, but this virus can cause a more complicated disease when acting as a co-factor in PRDC (Lunney et al., 2016). Porcine reproductive and respiratory syndrome (PRRS) caused by PRRSV has a very high economic impact on the swine industry in North America by causing an increase in mortality, a decrease in growth in growing pigs, and stillbirth and abortion in gestation sows (Zimmerman et al.,

1997). PRRS could reduce the pig's productive performance (e.g. growth rate) by not only reducing the average daily intake but also the digestibility, feed efficiency, and protein accretion rates in grower-finisher pigs (Schweer et al., 2017). There was evidence that the reduced productive performance caused by PRRSV is various concerning the interplay of the virus with the immune system of the hosts with different genetic backgrounds (Boddicker et al., 2012). The vaccination of sows for PRRSV was effective in improving reproductive performance in endemic PRRS farms (Jeong et al., 2017). However, the vaccination of growing pigs did not confer complete protection against the disease induced by the virus and failed to improve the poor productive performance of the infected pigs (Savard et al., 2016). The rapid mutation and evolution of the PRRSV genome, and the immune evasion mechanisms developed by the virus, make it difficult to confer complete protection of the pigs by vaccination (Kimman et al., 2009).

### 1.1.3.2 Porcine Circovirus Type 2 viruses (PCV2)

In the late 1990s, PCV2 was first isolated from the pigs affected by the Postweaning Multisystemic Wasting Syndrome (PMWS) and it was antigenically and genetically distinct from the non-pathogenic Porcine Circovirus Type 1 virus (PCV1) (Allan et al., 1998). The main infection site of PCV2 is lymphoid tissues and this causes lymphoid depletion and immunosuppression of the infected pigs (Meng, 2013). PCV2 has also been associated with PRDC and its impact is probably leading to the occurrence of PCV2-systemic disease (PCV2-SD) which is a multifactorial disease caused by the PCV2 infection (Ticó et al., 2013). Transmission of PCV2 may happen by direct contact with the infected pigs and the infection by PCV2 may persist in swine for several months

under either experimental or field conditions (Opriessnig et al., 2004). PCV2 could cause overt weight loss, late-term abortions, and stillbirths in pigs, and the infected pig is easy to diagnose from a clinical point of view by checking for lesions of lymph nodes, thymus, lung, and liver (Segalés, 2012). The most widely used control measures for PCV2 include vaccination, using antibiotics to prevent concurrent bacterial infections, improvement of biosecurity and sanitation, isolating affected pigs and disinfecting the pens after every use, decreasing stressors (e.g. high stocking density, inadequate ventilation, inadequate temperature control), and control of concomitant viral infections, especially PRRS. Other prevention and control measures used on young pigs before the anticipated onset of the disease include injection of vitamins and vaccination against common pathogens (Grau-Roma et al., 2011). Some evidence supported that different breeds of pigs showed differential susceptibility and disease severity in responding to PCV2, which implies the genetic background may influence the pathogenesis of this virus (Allan et al., 1998).

#### 1.1.3.3 *Mycoplasma hyopneumoniae* (Mhyo)

*Mycoplasma hyopneumoniae* typically shortened to "Mhyo" is a species of bacteria known to cause porcine enzootic pneumonia which is a highly contagious and chronic disease (Robinson, 1966) and could cause a significant reduction in the growing weight of pigs (Maes et al., 2021). Mhyo is also known as a major pathogen involved in PRDC (Holko et al., 2004). Mhyo attaches to the cilia of epithelial cells in the lungs of swine, which lead to the stop of beating, clumping, and even the loss of cilia, and eventually the death of epithelial cells. Lesions can be found in the lungs of pigs with porcine enzootic pneumonia (Desrosiers, 2001). This damage may increase the susceptibility of the

secondary infection in the respiratory system, by pathogens such as the Swine Influenza Virus (SIV) (Thacker et al., 2001). The treatment of porcine pneumonia is limited to antibiotics, but the efficacy is not sufficient to completely remove the infection. Vaccines have been found to reduce the severity of the disease but could not prevent the disease development in infected pigs (Maes et al., 2008).

### 1.1.3.4 Swine Influenza Virus (SIV)

Swine influenza (also known as "swine flu") is an acute, highly contagious, respiratory disease that results from the infection of SIV (Kothalawala et al., 2006). The classic type A infection of SIV isolates with mild virulence may favor the replication of *Haemophilus parasuis* (Mussá et al., 2012), *Actinobacillus pleuropneumoniae* (Tobias et al., 2009), and Mhyo (Desrosiers, 2001), any of which may complicate the SIV outbreak. In winter, it has been reported that low relative humidity produced by indoor heating and cold temperatures would favor influenza virus spread (Lowen et al., 2007). Thus, in North America, SIV outbreaks are most common in fall or winter and usually start from the onset of particularly cold weather. An outbreak typically starts with one or two individual cases and then spreads rapidly within a herd, mainly through aerosolization and pig-to-pig contact (Tellier, 2009). The clinical symptoms of SIV infection contain depression, fever (to 108°F [42°C]), anorexia, coughing, dyspnea, weakness, prostration, and a mucous discharge from the eyes and nose (Vincent et al., 2008). Vaccination and strict control of the pathogen could effectively prevent influenza outbreaks. Besides, antimicrobials may reduce secondary bacterial infections and expectorants may help relieve symptoms in severely affected herds (Parmar et al., 2011). Cross infection of SIV

is known to happen between pigs and humans and the response in most healthy people are mild or sub-clinical, but genetic reassortment of the virus in the hosts may generate novel strains that bring potential risk to human health and the pig industry (Krueger & Gray, 2012). Nevertheless, the name of "swine flu" and the concerns on zoonosis have changed people's perception on pork and may reduce pork consumption (Goodwin et al., 2009; Johnson, 2009).

#### 1.1.3.5 *Actinobacillus pleuropneumoniae* (App)

*Actinobacillus pleuropneumoniae* or App is a Gram-negative, facultatively anaerobic, respiratory pathogen found in pigs and causes pleuropneumoniae, a severe and contagious respiratory disease (Cruijsen, 1995), which could bring big economic losses to the swine industries in most countries. To date, 16 serotypes of App have been identified based on the capsular polysaccharide composition (Bossé et al., 2017). Transmission of App is mainly by nose-to-nose contact while the aerosol transmission is limited, and many recovered pigs are carriers (Tobias, 2014). The clinical symptoms found in infected herds include severe respiratory distress, thumps, fever, anorexia, reluctance to move, and sometimes open-mouth breathing with a blood-stained, frothy nasal and oral discharge. It is common for App that concurrent infection happens with mycoplasma (Caruso & Ross, 1990), pasteurellae (Pohl et al., 1983), PRRSV (Pol et al., 1997), or SIV (Mussá et al., 2012). Although some vaccines showed promising effects on reducing clinical symptoms, the control of the App is still difficult. Pigs are not fully protected from infection and transmission after vaccination and survivors are frequently remaining to be carriers (Ramjeet et al., 2008). Management strategies including segregated early weaning, "all-

in/all-out" management, reduced stocking rates when possible, and improved ventilation are recommended. The isolation of App from the upper respiratory tract can be useful for the detection of carrier pigs and complements serological screening (Sidibé et al., 1993). Future prevention strategies may be more focused on screening genetic markers in pigs that are associated with resistance to App (Sassu et al., 2018).

### 1.1.4 Rationale for a multiple-pathogen challenge study

The goal of agricultural production, which includes the pig sector, is to feed everyone on the planet, and demand for pork, both in terms of quantity and quality, is steadily rising. Antibiotic resistance and swine disease outbreaks, however, restrict the growth of pig production and raise consumer concerns about the safety of their food. The overuse of antibiotics in disease control or prophylaxis increases the global emergence of antibiotic-resistant bacteria and the problems of residues in animal products (Koch et al., 2017; Nisha, 2008). The extensive use of antibiotics in pig farms dramatically increased the selection pressure for resistant genes in bacterial pathogens, leading to more frequent spread of antibiotic-resistant bacteria while limiting the efficacy of antibiotics in combating disease outbreaks. This has become an One Health issue and should be tackled carefully to prevent uncontrollable infections in swine industry and lower the risk for evolution of strains pathogenic to humans (Monger et al., 2021). Meanwhile, the increasing scale of swine farming makes management more challenging. In intensive pig production, a series of measures are used to protect the health of pigs exemplified by modifications of herd environment and population management. It worth noting that despite of reducing the impact of some swine infections (e.g. Trichinella, Toxoplasma,

Sarcoptic mange) by these strategies, there might be negative effect on the prevention of some other diseases (e.g. PRRS, Influenza, PCV2 disease) (Davies, 2012).

Considering the diverse pathogen interactions and secondary infection described above, infections in pigs are commonly complicated and the interplay of pathogens may drive novel disease spread (new hosts and geographic distribution) and genetic mutations of pathogens (Engering et al., 2013). Thus, the strategies of disease control in modern pig farms should not only focus on the main infectious pathogen but also build interventions against multiple pathogens that tend to be cross-infected exemplified by the well-known multi-microorganisms (polygenic) disease PRDC as introduced above. However, it is practically difficult to monitor multiple types of pathogens in farms and it seems impossible to select pigs with resistance to multiple pathogens because of their differential pathogenesis mechanisms. How to reduce the impact of diverse pathogens on swine farming has become a new direction to increase production efficiency. Indeed, in addition to the selecting traits associated with productivity, the worldwide pig breeding objective has extended to also focus on pig robustness in the context of the burden from multiple pathogens and complicated environmental factors (Hermesch et al., 2015). The definition of resilience was introduced to describe the feature of robust pigs which exhibit the potential to maintain reasonable productivity under the challenge of infections (Albers et al., 1987; Knap, 2005). Therefore, to better assess the traits that represent resilience, multiple infection patterns need to be sorted and designed in the common disease study.

## 1.2 Pig disease study and experimental models

### 1.2.1 Commonly-used control and treatment of pig disease

Livestock diseases have big impacts on various aspects of livestock production including reduced productivity, economic loss, the uncertainty of food security, and concerns about animal welfare and human health (Reddy, 2015). Myers has systematically summarized and discussed the important aspects of swine diseases concerning disease recognition, prevention, and control, and the two key components for the management of swine disease are: 1) prevention of disease prior to pig arrival and 2) treatment for sick pigs (White, 2005). Pig care prior to pig arrival which includes barn set up (clean, disinfected, dry, and warm) and health plan (vaccination and feed/water/injectable medication) are extremely important. After the pigs' arrival, it is recommended to sort and settle down the pigs by size, particularly the 2~5% of the smallest pigs. It is important to leave space for special treatment of pigs and it is necessary to move and isolate the sick pigs to specific pens or barns from the healthy pigs. To handle sick pigs more effectively, a pig-marking system needs to be put in place. Different symptoms and sick time should be documented and identified. The sick pigs should then receive targeted medical care, which should also be marked and recorded. Additionally, sick pigs require extra nourishment and a temperature-controlled environment. (e.g extra warmth for pigs with fever). Besides, an all-in-all-out production system could be applied during the treatment and vaccination to prevent contaminating the next batch (Davies, 2012).

Scientific farm management can largely control infections and reduce the risk of

disease outbreaks. However, the clinical symptoms of a specific disease could be affected by not only the genetic viability of the pathogens themselves but also significantly influenced by the farm management characteristics (Goldberg et al., 2000). On one hand, combining the progress of studies on pig pathogens' virulence and evolution and the development of drugs and vaccines, it would be helpful to standardize the management protocol of pig farms to efficiently control infections. On the other hand, studies regarding the genetic level of pigs such as the genetic (or gene) editing of pigs (Whitworth et al., 2016) and genetic selection (Henryon et al., 2001) with specific disease resistance are undoubtedly beneficial for the control of pig diseases.

## 1.2.2 Swine immune system and disease response

### 1.2.2.1 Immune responses

The immune system in mammals has developed to protect them from a variety of pathogenic microbes in the surrounding environment that can disrupt host homeostasis (Chaplin, 2010). As with other mammals, the immune system of pigs can recognize the invasion of external pathogens and induce immune responses by the combination of innate and acquired immune responses. Once the pathogen invades pigs through the respiratory tract, esophagus, or wound, it would initially induce the innate immune response, which mainly includes but is not limited to the tissue physical barriers, the activation of the innate immune cells, and extracellular functional molecules (e.g. cytokines, chemokines, the complement system, and antimicrobial peptides) (Knap, 2005). The acquired immune response is induced later than the innate immune response through the process of recognizing and presenting the pathogen-derived antigen to

lymphocytes (Luster, 2002). The innate and acquired immune responses are strongly connected by the broad interplays between the innate and adaptive immune cells.

Innate immunity is the first line to recognize and defend against invading pathogens. The innate immune response is not specific to pathogen types and it is essential to effectively activate the acquired immunity. The innate immune response is based on rapid and short-term activation of tissue cells (epithelial cells and resident tissue cells), early production of pro-inflammatory cytokines, and recruitment and activation of innate immune cells (macrophages, natural killer (NK) cells, dendritic cells, etc.). Moreover, besides cellular innate immunity, some soluble plasma components, such as complement, antibodies, APPs, and antimicrobial peptides, play crucial roles in mediating humoral innate immune responses (Mair et al., 2014).

Acquired or specific immunity is divided into humoral immunity and cell-mediated immunity and is characterized by specific responses to particular antigens. As with innate immunity, acquired immunity can also distinguish what the "self" is from the "non-self" (foreign antigen) and which is tolerated, and which must be eliminated (Knap, 2005). However, the latter takes a longer time for selective activation and expansion of the adaptive immune cells that can recognize the specific pathogens, and thus are responding more slowly but more potently and effectively than the innate immunity. Furthermore, the effect of adaptive immunity lasts longer than innate immunity since partial antigen-specific adaptive immune cells have "memorized" the antigen after the first exposure (primary immune activation) and differentiated into long-lived memory cells. Those cells persist in the bloodstream and can respond with the same antigen subsequently with greater efficacy and rapidity (secondary immune activation) (Taylor et al., 2012).

#### 1.2.2.2 Immune cells

The response to pathogens is mediated by the complex interactions and activities of a large number of different cell types involved in the immune response. Leukocytes, frequently referred to as white blood cells, are one of the most crucial cell types involved. They come in two main types that work together to identify and remove pathogens. Those two basic types of leukocytes are phagocytes, which can engulf and destroy invading cells, and lymphocytes, which can mediate cellular and humoral immune responses against the invaders to recognize and clear them (Panawala, 2017).

Many different cells are thought to be phagocytes including macrophages, monocytes, neutrophils, dendritic cells, and mast cells. One common type is neutrophils, which phagocytose and destroys pathogens that are opsonized by the complement system. The number of blood neutrophils increases usually as a result of bacterial infection (Koenderman et al., 2014). Other types of phagocytes may use different mechanisms to ensure that the body responds appropriately to certain types of invaders. Macrophages are tissue-resident phagocytes and differentiate from blood monocytes. They are recruited to the site of infection and are able to phagocytose the opsonized pathogens and apoptotic infected cells, and produce cytokines that modulate the whole innate immune response. Macrophages are also important effector cells of cellular immunity that function as antigen-presenting cells (APC) (also dendritic cells) and initiate acquired immune responses (Parkin & Cohen, 2001).

The two main types of lymphocytes are B cells and T cells. The development of lymphocytes takes place in the bone marrow, where they remain to develop into B cells,

or they move to the thymus to develop into T cells. B cells and T cells have different functions, and they are the main effector cells in the humoral and cellular immune response, respectively. B cells are stimulated through the surface antigen-specific receptor (B cell receptor) and differentiated into plasma cells which are responsible for producing specific antibodies. T cells are also activated through the surface antigen-specific receptor (T cell receptor) while this process requires antigen presentation from APCs. T cells are mainly classified into T helper cells and T cytotoxic cells dependent on their surface molecules and primary functions. T helper cells play the central role in mediating the activation and resolution of the humoral and cellular immune response while T cytotoxic cells are killer cells that directly mediate the cellular immune response to the infected cells. Both the T and B cells can form memory cells to maintain a long-term specific immunity. NK cells are important innate lymphocytes and primarily recognize and kill the cells infected by an intracellular pathogen (e.g. virus) that reduces the major histocompatibility complex I (MHCI) expression or up-regulates stress molecules on the surface of the target cells. They also participate in specific cellular immunity through a process mediated by the antibody Fc receptors called antibody-dependent cellular cytotoxicity (Gerner et al., 2009; Henryon et al., 2001; Whitworth et al., 2016).

### 1.2.2.3 Acute phase response and proteins

Acute phase response (APR) is an important part of the host's early defense in the innate immune system. APR induces a series of complex systemic reactions that help eliminate infections, restore homeostasis and promote healing in the host under different

situations including trauma, infection, stress, tumor formation, and inflammation. APR is widely distributed across all mammals and its counterpart can also be found in invertebrates but is more robust potentially to better adapt with a less-potent adaptive immune system. APR can be triggered by numerous pro-inflammatory cytokines generated in an innate immune response (Cray et al., 2009).

APPs are blood proteins synthesized mainly by liver cells and are characterized by the immediate and rapid change of their serum concentrations during the APR. In pigs, the inflammatory response from the pathogen infection leads to the release of acute phase proteins from liver hepatocytes into the circulation (González-Ramón et al., 2000). It has been shown that APP secretion is mediated by the pro-inflammatory cytokines including interleukin-1 (IL-1), interleukin-6 (IL-6), and tumor necrosis factor-alpha (TNF-α) (Baumann & Gauldie, 1990). APPs are classified as positive (increase) and negative (decrease) based on their concentration change in response to the pro-inflammatory cytokines in APR (Jain et al., 2011). Most APPs are known as positive APPs and depending on the fold change when facing the challenge, they are categorized as major (10-100 fold), moderate (2-10 fold), or minor (slight change) APPs (Ceron et al., 2005). However, the pattern and classification of a specific APP are various in different species (Murata et al., 2004).

APPs are the contributors to the APR to limit infections and maintain host homeostasis in an antigen-non-specific manner (Gerner et al., 2009). APPs play various roles in the innate immune response. For example, haptoglobin (HP) binds free hemoglobin in the circulation and facilitates the hemoglobin endocytosis by macrophages to protect from the toxicity of hemoglobin during intravascular hemolysis (Nielsen &

Moestrup, 2009); while the AGP works as a carrier of basic and neutrally charged lipophilic compounds, and is implicated to play immune-modulatory or anti-inflammatory roles (Taguchi, 2013). Most APPs are regarded as non-specific responders in APR under different inflammatory conditions. However, some APPs appear to be associated with specific types of pathogens. For example, the increased level of serum haptoglobin but not AGP was observed after an experimental PRRS virus challenge (Asai et al., 1999).

Although research on animal serum proteins was carried out in studies in the mid-20th century, the broad use of APPs in veterinary medicine was not reported until the early 1990s (Gerner et al., 2009). From then to nowadays, APPs are regarded as useful biomarkers in disease diagnosis with their features of dramatic and rapid change after infection or inflammation since APR happens at the relatively early stage of the immune response. For instance, in pigs, the increased serum level of AGP was associated with some infectious diseases including pneumonia or meningitis (Itoh et al., 1993). APPs including HP, CRP, Pig-Major Acute Phase Protein (Pig-MAP), and Serum Amyloid A (SAA) were strongly elevated in pigs challenged with acute infection of *Actinobacillus pleuropneumoniae* under experimental condition (Heegaard et al., 1998). Transthyretin and transferrin in serum were found as negative APPs in responding to *Streptococcus suis* infection (Campbell et al., 2005) and acute experimental salmonellosis (Kramer et al., 1985), respectively, and may be developed as markers to monitor those infections. Summarized positive and negative APPs classified by induction level are listed in Table 1. In dogs, CRP, HP, and SAA are regarded as important diagnostic markers of steroid-responsive meningitis-arteritis (Lowrie et al., 2009). AGP was identified as a biomarker

for feline infectious peritonitis (Duthie et al., 1997) and potentially for lymphoma (Winkel et al., 2015) in cats. For ruminants such as cattle, HP is known as a major responding APP in the APR and is broadly applied to diagnose and predict the development of many inflammatory diseases including mastitis (Gultiken et al., 2012), enteritis, peritonitis, pneumonia, and endometritis (Eckersall & Bell, 2010). In dairy cows and sheep, a mammary-associated SAA3 isoform that can be produced by the infected mammary gland and secreted to milk during mastitis, making it a potential biomarker for this economically important disease (Campbell et al., 2005).

APPs are not only an early systemic sign of animal disease but also can indicate the health status of animals. For example, HP is considered a promising marker of the herd health status of pigs because researchers found that the pigs with a high concentration of HP at the sub-clinical stage later showed some clinical signs of disease (e.g. depression or gauntness, respiratory difficulties, and inflammatory lesions) (Chen et al., 2003; Harding et al., 1997). Besides, the increased serum level of CRP is widely used to identify inflammatory lesions in pigs (Eckersall et al., 1996) and is also used to assess the general health status (Bürger et al., 1998). The levels of HP and SAA were found to be much higher in dairy cows with acute disease compared to the healthy ones (Tourlomoussis et al., 2004). Furthermore, APPs can also be used as an index to assess productive performance. It has been reported that HP was negatively correlated with the growth rate of pigs (Hiss & Sauerwein, 2003) and different plasma AGP levels have been discovered among different pig lines (Clapperton et al., 2005) and breeds (Clapperton et al., 2007).

In recent years, experimental veterinary medicine has become more interested in using APPs to assess animal health. Regarding animal models of diseases, animal health

monitoring, and objective evaluation of animal health, the benefits of employing APP detection are extensively supported in the human and veterinary medical literature. Unfortunately, due to the actual limitations of the current detection technology, the clinical application of APP analysis is not universal. Ongoing challenges include the need for automated analysis and standardization of testing across laboratories.

Moreover, using APP as diagnostic and prognostic tools should be cautious and needs consideration of the potential effects of the environment, handling process, and other types of stress besides the diseases of interest. Continuous research aiming to understand the exact role played by APPs in disease pathophysiology and during the innate immune response to infections will accelerate and extend the applications of APPs to benefit the health, welfare, and production of animals (Cray et al., 2009; Jain et al., 2011; Murata et al., 2004). Therefore, APPs are promising potential biomarkers that could be exploited in pig production to predict or assess pig disease response.

**Table 1.1** Summary of classified APPs in pigs (Carpintero et al., 2005; Eckersall et al., 1996; Heegaard et al., 2013; Lampreave et al., 1994; Pomorska-Mól et al., 2013; Skovgaard et al., 2009; Sorensen et al., 2006).

| | **Positive APP** | **Negative APP** |
|---|---|---|
| **Major (> 10-fold change)** | CRP, SAA | ? |
| **Moderate (2-10 fold change)** | HP, Pig-MAP | Albumin, Apolipoprotein, Fetuin |
| **Minor (<2 fold change)** | Ceruloplasmin, Transthyretin | AGP, Transferrin |

## 1.2.3 Methods of pig disease study

Numerous studies on pig diseases are being conducted with the goal of improving animal welfare and addressing the direct and indirect financial losses brought on by diseases. These studies examine the pathophysiology and etiology of the diseases, as well as their diagnosis and treatment options. Below is a basic introduction to current research methodologies (together with their benefits and drawbacks).

### 1.2.3.1 Animal model

The first step in designing an experiment for disease research is to select an appropriate animal model in line with the objectives of the study. The animal model of pig disease research is mainly divided into three groups relating to different stages of pig production: pregnant sows, weaned piglets (21-23 days old), and growing/finishing pigs. Pregnant sows are mainly used to detect diseases that impact reproduction traits (Papatsiros et al., 2006), including but not limited to days of pregnancy, birth weight, and the number of stillbirths (Alexopoulos et al., 2005). The weaning period for modern pig production is crucial. Pigs in this period are undergoing stress and adaptation to the stressors including environment and diet change which may cause low or variable growth rates and increased risk of infection by enteric pathogens (King et al., n.d.). The status of the weaned pigs is subsequently influenced by daily feed intake and weight gain, and especially by health status (Fano et al., 2007; Magar et al., 2000). The growing/finishing pigs are the top priority for the intuitive observation of how the disease affected the pig production (de Lange et al., 2001; Donovan, 2005). The performance of the pig during the growing-finishing period has a significant impact on farm efficiency and profitability

but they are at the risk to be infected or co-infected by some respiratory or digestive pathogens including but not limited to PRRSV, PCV2, SIV, and Mhyo which were introduced above (Fablet et al., 2018). Moreover, considering the complicated environmental factors in swine herds, most experiments are often conducted under clean and sterile or specific-pathogen-free (SPF) conditions to avoid the influence of environmental factors on experimental results (Alexopoulos et al., 2005). Thus, animal models should be carefully designed with well-considering the research questions and experimental subjects.

#### 1.2.3.2 Infection method

To study the host response to a specific pathogen associated with infectious disease, researchers usually intentionally apply the infection (or pathogen challenge) to the animal model. There are different choices to introduce pathogen effects to experimental subjects depending on the purpose of the experiment, animal welfare, and pathogen characteristics. Firstly, based on the impact on animal health, infection methods can be initially divided into two categories: denatured and live pathogen infections. Introducing denatured pathogens can elicit a degree of host immune response and form antigen-specific memory cells. However, this method only has a slight impact on pig health because of the use of attenuated or inactivated pathogens so it is hard to study the clinical symptoms of a particular infectious disease (Fachinger et al., 2008). Genetic material cloned from infectious pathogens and live pathogens can lead to both the apparent clinical symptom for susceptible pigs and immune response (Fenaux et al., 2002; Shimizu et al., 1996). There are different approaches to introducing the pathogen of interest to the host that

requires careful consideration to avoid inducing potential adverse effects on the animal model. The oral route is an economical and convenient way to directly introduce infectious agents from the mouth or premixed into the diet. Lalla et.al inoculated a periodontal pathogen to a mouse model by gavage to investigate its effect on atherogenesis (Lalla et al., 2003). The limitation of the oral route may include the slower onset of action, degradation of the introducing agents by the digestive tract, passive reflux in the stomach, and particular requirements for technical skills (Turner et al., 2011). The intranasal route is commonly-used to introduce pathogens that can cause respiratory disease. Nelson and colleagues inoculated a virulent PRRSV to piglets and examined the antibody response to the virus (Nelson et al., 1994). However, the dosage is hard to be accurately controlled through the intranasal route because of the sneezing response of animals (Cray et al., 2009). Another way to deliver the infectious agents is through injection and this can be further detailed based on the location of the delivery such as intravenous, subcutaneous, intraperitoneal, intramuscular, or intradermal route. Choosing from those different injection routes is dependent on the experimental design, for example, the induction of a local or systemic effect. Boddicker explored the genetic basis of the different responses (resistant or susceptible) to PRRSV by conducting the experimental infection of a virulent PRRSV intramuscularly and intranasally on pigs aged between 18 and 28 days (Schweer et al., 2017). Associated common complications of these delivery routes should be adequately considered, and a well-designed administration plan, strong technical skills, and asepsis can minimize the adverse effects derived from the infection approaches (Cray et al., 2009). The current study was conducted utilizing a natural disease challenge model to simulate disease challenges for

pigs in a natural environment from commercial farms, as opposed to the experimental models mentioned above. Chapter 2 provides a detailed description of this paradigm.

#### 1.2.3.3 Sampling method

Pig sampling is one of the most important stages in the design of the experiment and may affect the results of subsequent tests. A number of aspects, including the goal of experiment, the choice of animals, the type of sample, the sample size, the sampling methods, whether the sample is fresh or fixed, and the transportation of sample must all be taken into account. Other crucial factors include the cost (pigs and facility) and simplicity of sample collection. Studies of immune responses during infectious diseases can provide suitable conditions to understand pathogenic mechanisms and also how the host immune system prevents disease. For the experiments assessing immune response, blood and several common immune organs (liver, spleen, lymph node) are the primary sample collection objects. The Sang-Myeong Lee group assessed the cellular and humoral responses in pigs experimentally challenged with PRRSV by isolating blood, lung, bronchoalveolar lavage, and bronchial lymph node samples (Nazki et al., 2020).

In addition, the clinical signs caused by pathogen infection and the replication of the pathogen in the body is also essential for research purposes, the sampling methods in some pig disease studies need to be properly designed to diagnose the disease accurately. Selection of pigs based on observation, body temperature, and treatment history, is an important step prior to sample collection (Bergland et al., 2005), which is also a major consideration in this study to assess the pig response in our disease challenge model. Pig selection can be straightforward in some cases to differentiate the pigs with infectious

disease. For example, gray or purple lungs can help confirm the degree of infection of Mhyo in pigs (Desrosiers, 2001). Pigs infected by *Actinobacillus pleuropneumoniae* have fibrous ribs and lungs (Sanford & Josephson, 1981), and SIV will cause necrotizing bronchiolitis (Desrosiers, 2001). Sampling from those specific tissues in necropsy will help the diagnosis of a specific disease. However, different infectious agents may have similar clinical signs. A good example is diarrhea which is a common sign of enteric disease but may have various causative agents. PEDV is one of the agents that causes diarrhea, and sampling for PEDV is usually gut tissue and fecal materials, then make a definite diagnosis by detecting specific pathogen nucleotide or protein expression (Song & Park, 2012). Serum sampling was also reported for PEDV diagnosis by detecting serum antibodies (Carvajal et al., 1995). Interestingly, a study suggested that PEDV may have airborne transmission and they detected viral-specific RNA from air samples from the room with infected pigs (Alonso et al., 2014). Moreover, some situations such as porcine respiratory disease may involve more than one causing agent. Thus, sampling pigs at recurring intervals can help identify all the agents involved (Bergland et al., 2005).

### 1.2.3.4 Data acquisition from samples

As a general area of disease research, pathology involves four major components for a particular disease: cause, the developing mechanism (pathogenesis), cellular structural change (morphological change), and consequence of change (clinical presentation). In the experiments aiming to investigate disease pathology and its effects, following pig sampling, data acquisition we will adopt different methodologies based on these four major disease components as experimental purposes and different measurement

objectives. Two important aspects that are necessary to be clearly defined are the studying system (pigs at individual, group, or breed level) and the studying output (pig behavior, production, reproduction, or health).

Blood or a localized lesion is typically obtained for testing for the presence of potential pathogens in order to identify the causal agents. For the identification of particular antigen or serum antibodies of PRRSV, PEDV, SIV, PCV2, and Myho, as examples, the standard testing techniques include immunofluorescence (IF), immunohistochemistry (IHC), and enzyme-linked immunosorbent assay (ELISA). Polymerase chain reaction (PCR) methods such as nested PCR, reverse transcriptase PCR (RT-PCR), and real-time quantitative PCR (qPCR) are widely used to sensitively detect gene fragments from those pathogens. Non-specific methods such as macroscopic and microscopic examination can be used to support the diagnosis (Donovan, 2005; Fano et al., 2007; Maes et al., 2008; Nodelijk, 2002; Opriessnig et al., 2004).

Pathogen entrance, replication, and latency, as well as the accompanying host immune responses, all playing a role in disease etiology. For the development of antiviral medications and vaccines, understanding the viral entrance and replication mechanisms is essential. Researchers must first identify the viral and host proteins responsible for virus life cycle, this usually requires data from experimental methods ranging from target screening to functional validation in multiple disciplines including molecular biology, biochemistry, genetics, virology, and immunology. For example, CD163 was known as the main host receptor for PRRSV identified by complementary DNA (cDNA) library screening of swine macrophages (Calvert et al., 2007), and the viral glycoprotein GP2a and GP4 were determined to interact with CD163 for viral entry by a series of molecular

biological experiments including molecular cloning, expression, and co-immunoprecipitation (Das et al., 2010). On the other hand, it is also meaningful to characterize the host immune response and how the virus co-evolved to evade the host immune surveillance. For instance, infected cells and immune cells secrete a variety of cytokines that regulate inflammatory responses to the invading pathogens (Murtaugh et al., 1996), and laboratory methods (ELISA, qPCR) to detect cytokines in pig tissues or blood are essential tools for studying the pathogenesis of swine diseases (Suradhat & Thanawongnuwech, 2003; Van Reeth et al., 1999). Many viruses develop immune evasion mechanisms through co-evolution with their hosts. Researchers noticed that pigs infected by PRRSV exhibited weak specific humoral immune responses, and by combining molecular biology, biochemistry, and virology tests, they figured out that the viral glycan shielding of the viral proteins GP3 and GP5 dramatically minimized the immunogenicity (Vu et al., 2011).

Acute infection usually results in tissue damage and also the functional and morphological change of some immune cells. For example, macrophages are the main targets for PRRSV replication in the lung, and the viral infection induces alterations of cellular functions including the phagocytic and microbicidal capacity which can be assessed by phagocytosis assay (Jakab et al., 1980) and microbial killing assay (Lincoln, 1995), respectively. Besides, PRRSV infection causes morphological changes in alveolar macrophages that are characterized by apoptosis-like phenotype (cell rounding, surface blebbing, and membrane rupture), and this can be observed by scanning electron microscopy (Chiou et al., 2000). Evidence of apoptosis is also featured by multiple nucleosomal-sized DNA bands which can be displayed by agarose gel electrophoresis

(Elmore, 2007). A double-labeling experiment using terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) and *in situ* hybridization (ISH) can be used to visually determine whether the apoptotic cells were primarily virus-infected or were uninfected bystander cells (Sirinarumitr et al., 1998).

The consequence of a particular pig disease is mainly embodied in the resulting clinical symptoms and altered productive performance. Taking PRRSV as an example, as briefly introduced above, the main clinical presentations of PRRSV are reproductive impairment or failure in gestating sows (abortions, weak and stillborn piglets, infertility) and respiratory disease in weaned pigs (fever, pneumonia, lethargy, and failure to thrive). Studies such as comparing viral strains and testing therapeutic strategies usually assess the clinical score derived from the clinical presentations (Halbur et al., 1996). Meanwhile, the increased mortality and decreased growth rates resulting from PRRSV infection are the direct reasons for the economic loss in the pig industry. Thus, many studies focus on screening biomarkers to identify the virus-resistant or resilient pigs, and the production and health-related traits are the most important parameters to measure and define the phenotypes of interest (Lunney & Chen, 2010). The project in this thesis is also focusing on exploring the bio-markers derived from different biological aspects aiming to differentiate pigs that are resilient to the experimental challenge from a mixture of swine pathogens. The next section will introduce the definitions and studies of disease resilience, tolerance, and resistance, and their emerging significance in the pig industry.

## 1.3 Disease resistance, tolerance, and resilience

### 1.3.1 Definitions of resistance, tolerance, and resilience

Many factors, such as the negative effects (e.g. immunological imbalance) of excessive selection on production traits, increased consideration of animal welfare, environmental challenges (e.g. global climate change), and reducing the economic spending on medications, are pushing animal breeding strategies to include animal's robustness which including disease resistance, tolerance, and resilience (König & May, 2019).

Initially, the concept of disease resistance which is the ability to limit the pathogen burden was first proposed and generally used in plant disease control studies (Hammond-Kosack & Jones, 1997). Resistance can be considered as an active host action to limit the pathogen burden through mechanisms involving inhibition of pathogen infection and proliferation. Host resistance is generally quantified by a measure of infection severity such as within-host pathogen burden (e.g., viral or bacterial counts or parasite density) (Doeschl-Wilson et al., 2012).

Tolerance was introduced as a new method of controlling the detrimental effect of a disease (Fornoni et al., 2004; Restif & Koella, 2004), which refers to the ability to reduce the impact of pathogens on host performance without necessarily affecting pathogen burden. Tolerance can be quantitatively assessed by the change in production or reproduction performance (e.g. growth rate, feed intake, or litter size) of livestock in responding to the change in pathogen burden (Doeschl-Wilson et al., 2012).

An early study of sheep parasites by Riffkin and Dobson pointed out that immunocompetent sheep with better resistance to *Haemonchus contortus* did not need to overcome the pathogenic effects of the parasites which survived from the host immune response. Therefore, they suggested that the ability to withstand the pathogenic effect without clearing the infection was better termed "resilience" (Riffkin & Dobson, 1979). The definition of resilience was summarized by Knap and Bishop later (Knap & Bishop, 2000), which refers to the host's ability to maintain a reasonable level of productivity during a “disease challenge" (Bisset & Morris, 1996), i.e. to maintain the situation where those metabolic costs for disease control do not lead to a significant reduction of production-related metabolic processes. In practice, it may be easier to assess resilience using productive traits compared with assessing resistance or tolerance which require determining pathogen burden (Gray, 1997; Lincoln, 1995). For example, pigs in the PRRS Host Genetics Consortium (PHGC) model showed a wide variation in weight gain, with some pigs gaining weight at a relatively normal rate while others failed to thrive during the 42-day infection period. The pigs that gain weight normally can be considered as “resilient” under viral infection (Rowland et al., 2012; Schweer et al., 2017). Morris and his colleagues selected resilient sheep based on the post-weaning age to receive necessary anthelmintic treatment to maintain acceptable growth in pastures with nematode challenge, and the selected resilient lambs extended 23.6 days to receive the drench treatment and gained 4.5 kg more live weight at 6-month age relative to the control animals. This work shed light on the selection of both productivity and improved health status in the context of pathogen challenge (Morris et al., 2010). Other parameters to define resilience may include disease incidence, survival rates, and immune response

(Sassu et al., 2018). Thus, besides disease resistance and tolerance, resilience is another promising characteristic in livestock breeding.

### 1.3.2 Similarities and differences among the concepts of resistance, tolerance, and resilience

The concepts of disease resistance, tolerance, and resilience show some similarities and differences. They all represent the animal's ability to adapt to and interact with environmental stressors and pathogens (König & May, 2019). All of them have the same ultimate goals in animal production, which are maintaining a high-level productive performance for economic profitability and improving animal welfare by alleviating the impact of disease challenges on animals (Chiou et al., 2000; Råberg et al., 2007).

The influence on host-pathogen interactions is where those three definitions diverge most. The focus of resistance is on the host's capacity to prevent the establishment and/or progression of pathogen infection. Contrary to resistance, tolerance emphasizes how the host lessens the harm brought forth by the infection without directly affecting it. Nevertheless, both resistance and tolerance are host defense mechanisms to maintain the health status of animals and are sometimes hard to distinguish. Furthermore, the mechanisms of resistance and tolerance may differentially influence the host-pathogen co-evolution. The evolution of animals with resistance and tolerance has contrasting effects on the epidemiology of infectious diseases (Miller et al., 2006; Roy & Kirchner, 2000). For resistance, hosts are selected for the ability to reduce pathogen burden and this selective pressure may result in pathogen evolution for increased infectivity or virulence, leading to open-ended non-equilibrium evolutionary dynamics (Woolhouse et al., 2002).

In comparison, tolerance does not have a negative effect on the fitness of the pathogen so selecting the host genetic variation for tolerance may result in dampened antagonistic coevolution (Rausher, 2001), which tends to induce commensal co-evolution between the host and pathogen (Best et al., 2008; Guy et al., 2012). A study using the rodent malaria model suggested the existence of genetic variation for both resistance and tolerance while resistance and tolerance were negatively genetically correlated (Lunney & Chen, 2010), which is consistent with the proposed different defensive mechanisms and influence on host-pathogen co-evolution of resistance and tolerance.

Resilience reflects the host's ability to maintain relatively reasonable production which cares more about the productive potential under the challenges of pathogens (Bisset & Morris, 1996). It should be noted that resilience is usually considered to conflate the host defense mechanisms of resistance and tolerance. However, more work is necessary to characterize the individuals that are more resilient to disease than others. In the context of animal breeding objectives, the meanings of those three concepts may be roughly distinguished by the consequence of selection which is to control infection or to maintain productive performance during infection (Guy et al., 2012).

In terms of evaluating those three concepts in a study of animal disease, it is very intuitive and easy to distinguish whether animals are disease resistant or not by assessing the pathogen burden. In the study of tolerance or resilience, the detrimental effect of disease and animal performance are the main measurement objects, respectively, which are closer to the ultimate goals in animal production (protection of economic benefits and animal welfare) relative to the pathogen burden. The pathogen load is not the main parameter for resilience, whereas the productive performance of animals under disease

challenges should be more considered. More tests need to be run such as serum antibody levels for assessing tolerance.

In the genetic analysis of sheep resistance to the parasite by Albers (Albers et al., 1987), they reported that the resistance represented by the measurement of pathogen burden was highly heritable while the heritability of resilience based on depression of productivity was too weak to obtain selection response by mass selection. They concluded that selection on resistance was effective to improve this trait while also potentially increasing the productivity of animals during infection (resilience). The genetic variation of pig disease resistance trait involves multiple genes regulating aspects of both innate and acquired immunity (Dawson et al., 2013; Groenen et al., 2012). Disease tolerance and resistance have a weak genetic correlation and the heritability for tolerance analyzed by host performance is influenced by pathogen burden (Kause et al., 2012; Koch et al., 2017). Therefore, selection for tolerance is challenging in the absence of pathogen burden records (Kause & Ødegård, 2012). However, a recent study in 2017 using estimating breeding values for resilience effectively acquired selection responses for resistance and tolerance to infections predicted by Monte Carlo simulation (Mulder & Rashidi, 2017). Considering that resilience may combine the mechanisms of resistance and tolerance, this study suggests that selection on resilience might be favorable to increase disease resistance and tolerance of animals, and selection on resilience is a more practical approach since pathogen burden data is not always available and the recording is laborious and costly over the time of infection.

Direct strategies target animal resistance/tolerance to specific pathogens may result in increased susceptibility to other diseases (Wilkie & Mallard, 1999). Alternatively, the

indirect and putatively more global approach focuses on resilience identification could provide a measure of economic loss due to multiple disease challenges and hopefully predict the responses to pathogens in general (Putz et al., 2018a). Clear understanding of the differences in the concept of resistance, tolerance, and resilience is beneficial to pinpoint the breeding objective and make corresponding breeding strategies.

### 1.3.3 Comparison of resistance, tolerance, and resilience in pig disease studies

In early times, researchers used to focus only on disease resistance in pig breeding when screening the genetic variations that benefit the pig's health (Warner et al., 1987), while pig tolerance and resilience are being more considered in modern farms (Glass, 2012). As a result of accumulated information and technological advancement, trait selections are no longer restricted to evaluating productive performance, health, or pathogen burden (Vu et al., 2011), but also active immune response (Zhuge et al., 2012). The studies of disease resistance were usually confined to certain pathogens, the corresponding measurement of traits is relatively straightforward to determine the infection severity such as the viral or bacterial counts and parasite density. However, the case is more complicated for tolerance and resilience when determining the traits to be selected or identifying the underlying associated genes (Medzhitov et al., 2012). To define pigs with disease tolerance, besides assessing the pathogen burden, traits reflecting host production or reproduction such as growth rate, feed intake, and litter size are required to evaluate the impact of the pathogen. There are pieces of evidence that animals exhibit genetic variation with regards to resistance and tolerance and screening those

associated mutations could be a strategy to identify the associated functional genes or develop genetic markers for breeding (Lunney & Chen, 2010; Mazé-Guilmo et al., 2014).

Common traits used in the research of swine disease are typically broken down into a number of categories. The first is the characteristics that matter economically, such as daily weight gain. Second, it involves testing for immunological responses, such as the assessments of antibodies, immune cells, or changes in cytokines. Additional measuring parameters, such as body temperature and necropsy score, could be applied to the health state of animals in addition to these two key factors. It is challenging to characterize and select a pig with disease resilience because of its complicated nature which involves various dynamic biological processes ranging from nutrient metabolism, production, immune response, and recovery from infection (Friggens et al., 2017). Indeed, some heritable immune traits have been explored to be associated with pig performance (Bai et al., 2020). However, there is still a lack of an experimental model to evaluate the potential traits to qualify resilience and the correlation with heritable genetic variations. Our collaborators recently reported using daily feed intake and feeding duration data as measurements of disease resilience of finishing pigs in a natural challenge model that is similar to what I used in this thesis study (Putz, et al., 2018b). They found the feed intake and duration data was modestly heritable and showed a genetic correlation with mortality and treatment rate, but those measurements are usually interfered by environmental and social factors independent of diseases and thus suggested to be used to qualify resilience to general farm stressors. This finding was further confirmed by their more recent study which analyzed the genetic parameters of a series of resilience-related performance traits including feed intake and duration on a bigger population of pigs (including the pigs

which died during the experiment or challenge) in the natural challenging model (Cheng et al., 2020). The recent work from our group using the natural disease challenge model demonstrated that the complete blood count data from challenged weaned pigs were heritable and genetically correlated with growth and treatment rates, and thus, it could be a potential measurable phenotype to assess disease resilience (Best et al., 2008).

Although the three concepts all involve host-pathogen interaction and may share biological pathways to be established, in contrast to disease resistance and tolerance which are usually discussed based on a specific pathogen or disease, disease resilience has a more broad meaning including health and productive performance in the context of livestock production.

### 1.3.4 Applications of resistance, tolerance, and resilience in pig disease control

The preceding paragraphs have introduced the definition and comparison of the resistance, tolerance, and resilience in pig disease study. This section mainly gives some examples of how disease-responding phenotypes are applied in pig disease control. For disease resistance, many studies aimed to identify the responsible genes or variations that reduce the pathogen invasion or confer a stronger defensive response for pathogen clearance (Råberg et al., 2007). About 20 years ago, a study using genomic mapping and genotyping identified that the *FUT1* gene associated with the expression of the gut receptor for *E. coli* F18 in pigs and that a single nucleotide polymorphism (SNP) within this gene was associated with the resistance to the bacterial adhesion and colonization in the small intestine (Meijerink et al., 1997, 2000), which provided a potential biomarker to

breed pigs with adhesion resistance to *E. coli* F18. Similarly, another group focused on a locus on pig chromosome 13 which contained the candidate gene encoding the intestine receptor for *E. coli* F4ab/ac, and identified that a polymorphism within this region was associated with the resistance to diarrhea in pigs (Jørgensen et al., 2003). Besides, some researchers characterized the genetic variations that are associated with other porcine pathogenic infections such as PPRSV (Lunney et al., 2011), PCV2 (Opriessnig et al., 2006), and *Haemophilus parasuis* (Blanco et al., 2008), which provides possibilities for genetic selection in breeding programs and the development of biomarkers based on these genetic variations. However, infection of many pathogens may not be sufficiently blocked based on selection of these genetic variations mentioned above to reach exact resistance, especially for pathogens that are prone to mutation and evolve quickly.

Genome editing is a straightforward way to generate disease-resistant pigs on the condition that the molecular pathway of pathogenesis is adequately understood. A study reported in 2016 made efforts to establish PRRSV resistance in pigs by directly silencing the expression of viral entry receptor CD163 using CRISPR-Cas9 technology, and this successfully conferred protection of these pigs to an experimental virulent PRRSV strain (Whitworth et al., 2016). The same group later generated genome-edited pigs that were deficient for aminopeptidase N which is the receptor for transmissible gastroenteritis virus (TGEV), and those pigs were resistant to TGEV infection (Whitworth et al., 2019). African swine fever virus (ASFV) causes high mortality in domestic pigs that is resulted from the cytokine storm during the late-stage infection, which is another big potential risk to the world's pig industry. Rel-like domain-containing protein A (RELA) was known to be involved in the NF-κb cytokine signaling and showed a difference in the sequence

between domestic pigs and warthog which is naturally resilient to the lethal infection of ASFV (Palgrave et al., 2011). A promising attempt was reported to endow domestic pigs with the ability of resilience by substitution of amino acids at the *RELA* locus to its orthologue in the warthogs (Lillico et al., 2016). A more recent study tested the effectiveness of those substitutions at the *RELA* locus on domestic in responding to experimental ASFV intranasal challenge. However, even though the substitutions slightly delayed the onset of clinical signs and viremia, it was still insufficient to confer resilience to the domestic pigs like warthogs to prevent lethality or maintain weight gain (McCleary et al., 2020) suggesting the resilience of warthogs requires other factors besides the mutations at the *RELA* locus. Nevertheless, this genome editing trial provided valuable insights to assist the intervention's development for the devastating impact of ASFV on the pig industry.

Despite the advantage of breeding for disease tolerance in maintaining the high performance of pigs, identification of key genes that are responsible for disease tolerance is likely to be difficult (Råberg et al., 2007). Potential reasons may include the difficulty to quantify the tolerance trait, the polygenic nature, and the poorly-understood biological mechanisms to establish tolerance. Evidence for the observed different detrimental effects of PRRSV infection in lines and breeds of pigs suggests that there might be genetic variation related to disease tolerance that remains to be identified (Elmore, 2007). Moreover, the experience from the studies of animals other than pigs indicates the difficulty of selecting disease resistance and tolerance to multiple pathogens considering the complexity of the infectious diseases in actual pig farming.

Resilience is likely to be an ideal breeding aim for both pig health and production in

modern farms. Including the studies from our group and collaborators using the natural disease challenge model to assess the resilience phenotypes and heritability as mentioned above, selecting disease resilience in pigs is drawing more attention in recent years. A study published in 2019 suggested that the routinely collected medication records in pig herds such as the treatment of medications could be used to define disease resilience (Guy et al., 2019), and the number of treatments was also included in my study to define the phenotype of resilience. Another report in 2019 developed a potential phenotyping criterion by using the outcome data of PRRSV vaccination to differentiate susceptible and resilient female piglets to PRRSV before transferring to the PRRSV-positive farm, and the reproductive data were recorded to evaluate the heritability (Abella Falcó et al., 2019). Based on the genome-wide analysis of reproductive traits in response to a PRRSV outbreak by our collaborators (Serão et al., 2014), this group found that the genotypes of *SGK1* and *TAP1* genes influenced the reproductive performance and could be used in breeding sows resilient to PRRSV (Laplana et al., 2020). Our collaborators recently described that natural antibodies titers and total IgG levels in the blood of young healthy pigs were genetically associated with disease resilience assessed in our natural disease challenging model and could be used to predict disease resilience in commercial farms (Dekkers et al., 2019). Nevertheless, further studies are necessary and other types of biomarkers need to be screened to predict disease resilience under the high-health conditions of modern pig farms.

With the progress of the knowledge on pig disease-causing pathogens and the development of modern research techniques, selecting pigs with disease resistance, tolerance, or resilience are crucial breeding aims to control disease, maintain pig

production, and improve pig health and welfare. In particular, the development of high-throughput omics (e.g. genomics, epigenomics, transcriptomics, proteomics, and metabolomics) and the trend of integrating multi-omics approaches remarkably spearheaded the research in terms of identifying disease-causing genes or biomarkers for disease response (Suravajhala et al., 2016). In the following sections, I will primarily introduce transcriptomics and metabolomics which I applied to explore biomarkers to predict disease resilience in this thesis.

## 1.4 Transcriptomics

The work of mapping and sequencing the human genome started the study of genomics which is the earliest omics discipline and focuses on the entire genome but not on single genes. With the progress of new technologies, researchers were able to obtain a large amount of information from a tissue or cell sample, and the term "omics" was gradually introduced to generalize the studies to comprehensively and globally measure a set of biological molecules in a high-throughput fashion (Micheel et al., 2012). As briefly mentioned in the last section, omics is further classified as many disciplines regarding the target type of molecules. This section is mainly reviewing the history of transcriptomics and how researchers apply transcriptomics techniques to study health science including the studies of diseases in livestock.

### 1.4.1 Concept and history of transcriptomics

The concept of the transcriptome is the complete set of RNA transcripts (species and quantity) in a single or a population of cells at a specific developmental stage or under a

specific condition. Transcriptomics is the subject to study the transcriptome aiming to investigate and characterize different transcript species (e.g. mRNAs, non-coding RNAs, and small RNAs), the transcriptional structure of genes, and the change of transcripts expression levels in a biological process or response to certain conditions (Wang et al., 2009).

The first study to look into a set of transcripts was reported in 1979 when researchers characterized the abundance of transcripts associated with silk moth development from a cDNA library (Sim et al., 1979). The advent of automated DNA sequencing technology and bioinformatics made it more efficient to acquire and analyze enormous sequence data which significantly led to the progress of transcriptome characterization (Jiménez-Chillarón et al., 2014). The first investigation of a collection of human transcripts was published in 1991 but only 609 mRNA sequences were captured from the human brain (Adams, 1991). Milestone research was published in 1997 which analyzed 60,633 transcripts from 4,665 genes of yeast using serial analysis of gene expression, and this study opened the gate to genome-wide eukaryotic gene expression (Velculescu et al., 1997). With the development of the two key techniques, DNA microarray and RNA-sequencing (RNA-Seq), the transcriptomics studies entered a real high-throughput stage. In the year of 2008, researchers were able to investigate the transcripts from 16,000 genes and globally analyze the alternative splicing events in human tissues (Pan et al., 2008). By the year 2015, the further advance in sequencing technology made it feasible to accurately characterize the gene expression difference in multiple human populations (Melé et al., 2015). Meanwhile, transcriptomics approaches were widely used in different research areas including biology, medicine, and clinical and

pharmaceutical studies (Qi et al., 2011). In the mid-2010s, the seminal application of RNA-Seq in single cells brought transcriptomics studies to a new stage (Kolodziejczyk et al., 2015).

## 1.4.2 Techniques to study transcriptomics

The development of transcriptomics has been characterized by the invention and improvement of technologies, and the innovation of new techniques tended to make the old techniques totally abandoned (Lowe et al., 2017). In the beginning, from the 1980s to 1990s, researchers used the low-throughput Sanger sequencing to determine the gene transcription from the cDNA library using short specific transcript fragments called expressed sequence tags (EST) (Marra et al., n.d.1998). Although EST sequencing was regarded as the most efficient way to rapidly identify specific transcripts at that time, the limit of the throughput and cost of Sanger sequencing made it difficult to acquire a complete transcriptome. To quantify the amount of specific transcripts, hybridization assays such as northern blot (Alwine et al., 1977), and RT-qPCR (Becker-André & Hahlbrock, 1989) were widely used. Tag-based methods were the relatively high-throughput techniques including serial analysis of gene expression (SAGE), cap analysis of gene expression (CAGE), and massively parallel signature sequencing (MPSS), which were able to quantify transcripts from specific genes but were not able to differentiate transcript variants generated from alternative splicing events (Pertea, 2012). Similar to the EST approach, those assays were also based on Sanger-sequencing which was not feasible for large-scale analysis.

Hybridization-based microarray is one of the main contemporary techniques in

transcriptomics which was developed around the mid-1990s (Schena et al., 1995). Basically, a microarray contains a large-scale collection of short complementary nucleotide probes attached to a solid surface to measure the abundance of specific transcripts (cDNA) through hybridization (Pozhitkov et al., 2007). The quantification of gene expression is determined by detecting the hybridized specific transcripts labeled with fluorophore, silver, or chemiluminescence (Chaudhari, 2016).

Micro-array is the first cost-effective high-throughput assay that is able to measure thousands of transcripts in a single assay and distinguish and quantify spliced isoforms of transcripts. However, there are some disadvantages of microarray that limit its development. Micro-array usually generates highly noisy signals because of different target hybridization strengths and cross-hybridization, and microarray has an inherent limit which is that it can only measure known genes (sequence and exon-intron boundaries) but is not applicable to identify new gene transcripts (Suravajhala et al., 2016).

The other major transcriptomics technique is RNA-Seq which was developed during the 2000s (Jørgensen et al., 2003). As indicated by its literal meaning, RNA-Seq is a technique to sequence all the transcripts in a cDNA library in a high-throughput manner and quantify the number of counts for each transcript by following computational analysis of the sequencing result. The high capability of RNA-Seq to resolve the transcriptome is dependent on the innovation of high-throughput sequencing technologies. RNA-Seq uses massive parallel sequencing, also known as next-generation sequencing (NGS) relative to Sanger-sequencing, to sequence the cDNA library from one end or both of each molecule. For example, the invention of Solexa/Illumina technologies is a

milestone that makes RNA-Seq available to capture $10^9$ transcripts which can cover the entire human transcriptome (Abella Falcó et al., 2019; Jørgensen et al., 2003). The basic protocol for an RNA-Seq assay includes RNA isolation, preparation of the cDNA library, high-throughput sequencing, and computational analysis. The cDNA library is processed by generating fragmented short sequences and undergoes clonal amplification. The typical length of a cDNA fragment is around 150 bp, but different sequencing methods may require various lengths ranging from 30 bp to over 10,000 bp. In general, RNA-Seq is compatible with most NGS platforms that are based on spatially-segregated complementary DNA synthesis and in massively parallel sequencing (Anderson & Schrijver, 2010). The sequenced short sequences are then reconstructed by aligning to a reference genome and counted using computational methods (Jørgensen et al., 2003). The principles of RNA-Seq allow it to detect known and unknown transcript sequences in base-pair resolution and determine the gene expression level in a much larger range but with a much lower background noise compared with hybridization-based methods. Those advantages of RNA-Seq make it available to not only precisely measure the transcript expression level, but also widely be used to identify genetic variations such as gene polymorphism and splicing variants. Traditional RNA-Seq uses the RNA purified from tissues or a mixed population of cells, and this is not feasible to acquire the transcriptional profiles from sub-populations or even individual cells. With the increasing demands to understand the distinct roles of different types of cells or different cell clones, the single-cell RNA-Seq method was developed in the last decade which makes it possible to simultaneously determine the transcriptional profiles of hundreds to thousands of individual cells (Kanter & Kalisky, 2015). To overcome the limit of short sequencing

length from traditional RNA-Seq, long-read RNA Seq technologies were developed in the past few years, making it possible to resolve full-length transcripts and directly sequence the RNA molecules. Two companies (Pacific BioSciences and Oxford Nanopore Technologies) are leading this field with their distinct platforms. These evolutionary technologies are making great contributions to not only the improvement of gene annotation but also studies related to gene isoforms, RNA modification and non-coding RNAs. Researchers in this new-generation sequencing field are actively working on improving the sequencing accuracy and experiment throughput to extend its applications (Hu et al., 2021; L. Zhao et al., 2019).

### 1.4.3 Application of transcriptomics in pig health related research

As briefly mentioned above, transcriptomic approaches are broadly applied to health science studies. For disease diagnosis and profiling, for example, transcriptome sequencing has been used to explore novel gene fusions in cancer for diagnosis and screening purposes (Maher et al., 2009). Micro-arrays and RNA-seq are used to generate mRNA profiling as a prognostic tool (Byron et al., 2016). In addition to detecting the change in gene expression, RNA-Seq can be used as a diagnostic tool in conjunction with genomic analysis to discover mutations and splice-altering variations that are linked to the rare Mendelian diseases (Cummings et al., 2017). Allele-specific expression (ASE) is known to associate with a predisposition to disease and RNA-seq can be used to analyze ASE which is a potential biomarker to discover pathogenically relevant genetic variant (Pan et al., 2008).

In animal studies, many researchers were focusing on exploring transcriptome differences for favorable production traits (Meijerink et al., 1997), such as the study of muscle development in cattle (Cummings et al., 2017) and pigs (Huang et al., 2008), and adiposity in broilers (Bourneuf et al., 2006). There is increasing attention on using transcriptomic approaches as a tool to study health-related traits in livestock. For example, McCabe and colleagues applied RNA-Seq to profile the liver transcriptome of lactating dairy cows with different severity of negative energy balance (McCabe et al., 2012). A study comparing the transcriptome of the fat, liver, and muscle tissues between Jeju native pigs and Berkshire pigs revealed differentially expressed genes (DEGs) not only related to body growth but also immune response pathways (Ghosh et al., 2015). In addition, a whole blood transcriptomics study on growing pigs with divergent genetic selection for residual food intake showed DEGs associated with immune capacity and defense mechanisms (Jégou et al., 2016).

Transcriptomics strategies provide efficient ways to help understand host-pathogen interaction which can be mainly divided into two categories: characterization of pathogens with different virulence or response to the host defense; and hosts exhibiting differential responses to invading pathogens. On the side of pathogens, Yonghou Jiang's group developed a microarray-based approach to identify and differentiate various PCV1 and PCV2 clinical strains (Jiang et al., 2010). Moreover, a pan viral DNA microarray called Virochip was reported to be able to successfully detect multiple common pig pathogens including PRRSV, PCV2, PRCV, and SIV in different types of clinical samples (T. L. Nicholson et al., 2011). A study using the Dual RNA-Seq approach on *Salmonella enterica* serovar Typhimurium and human cells identified a pathogen small

RNA involved in regulating invasion-associated effectors and virulence genes, and some infection-specific non-coding transcripts in the host cells (Westermann et al., 2016). On the side of the host, Dirk Werling's lab generated the transcriptome profiles of myeloid immune cells in pigs challenged by PCV2, which provides information to understand PCV2 pathogenesis (Mavrommatis et al., 2014). Agnieszka Podolska and others profiled the microRNAs in the lung tissue of pigs infected by *Actinobacillus pleuropneumoniae* and identified some candidate microRNAs that are involved in the host response to this pathogen (Podolska et al., 2012). There were batches of studies on livestock aiming to characterize the different responses of the host to pathogens through transcriptomic strategies, and those results combined with genetic analysis may provide a theoretical basis for selective breeding focusing on disease resistance and tolerance. Kadarmideen and his colleague reported using microarray to identify biomarkers for parasite resistance in sheep using differential gene expression (DE) analysis (Kadarmideen et al., 2011). Kapetanovic et al. applied micro-array to assess the response of macrophages to lipopolysaccharide in different breeds of pigs and found some breed-specific DEGs (Kapetanovic et al., 2013). Yunliang Jiang's group investigated the cause of breed-dependent different PCV2 susceptibilities between Yorkshire x Landrace pigs and Laiwu pigs using RNA-Seq and suggested that the transcript level of *SERPNA1* gene may play an important role in the resistance to PCV2 of Laiwu pigs (Y. Li et al., 2016).

Aside from the rapid development of transcriptomics, other omics such as genomics, epigenomics, proteomics, and metabolomics have also entered the high-throughput omics era and have been applied to livestock production and health studies. These omics technologies concentrated on various types of molecules and levels of regulation in the

gene expression process. Researchers are always looking for the best biomarkers to predict production performance and disease risk, and there is a growing trend to use multi-omics approaches on the same sample to fully exploit this biological information for a more holistic improvement of animal production and health (Suravajhala et al., 2016). In the next part, I will introduce metabolomics which aims to acquire the global metabolite profiles of a sample, and in this thesis, I also applied metabolomics technology to explore potential blood metabolites to predict the disease resilience trait.

## 1.5 Metabolomics

The metabolome, like the transcriptome, is an important part of omics that focuses on studying the components of all small-molecule metabolites in biological samples under certain physiological conditions. By using this method, researchers can study the correlation between the amount or type of metabolites and the biological phenotypes of interest. It is also an important alternative method when transcriptome and proteome are not enough to understand the related mechanism behind the research.

In biological processes, there is a big gap between genotype and phenotype, which includes steps from DNA transcription, and protein translation, to product generation of cell activities (metabolites) (Fiehn, 2002). However, the expression of a particular phenotype from a genotype is not a simple linear pipeline. The existence of multiple feedback loops, such as from metabolites to proteins or transcripts, makes the regulation network more complicated for a specific phenotype. Furthermore, the cell activity network can be extended to the interactions with neighboring metabolites (Johnson et al., 2016). Nevertheless, metabolites are directly produced or involved in cell metabolic

activities, and metabolites are considered the closest molecules contributing to modulating the phenotype (Kasture et al., 2012). Thus, metabolomics is broadly applied to discover particular metabolites as biomarkers in clinics and industries (Guijas et al., 2018; Putri et al., 2013). Integrating the data from metabolomics with that from transcriptomics and proteomics is a trend to understand the genotype-to-phenotype gap in system biology.

### 1.5.1 Concept and history of metabolomics

Most metabolites are low-molecular-weight molecules (typically less than 1.5 kDa in humans) that play roles as substrates, intermediates, or products in metabolic processes (Wishart et al., 2007). And the function can be classified as “primary” or “secondary” dependent on their direct participation in energy storage and metabolism, or involvement in various biological pathways such as cell-to-cell communication, and antibiotic activity (Vinayavekhin & Saghatelian, 2010). The complete profile of metabolites in a biological sample constitutes its "metabolome". Analogous to the transcriptome and the proteome, the metabolome is a dynamic profile of a single cell, a tissue, or even a whole organism. As an omics of studying metabolites, metabolomics is previously defined as "the systematic analysis of metabolites structure, concentration and pathways, and metabolite interactions within and between organs under environmental influences" (German et al., 2005). The aims of metabolomics studies are acquiring comprehensive information on the metabolome in the context of a physiological and pathological state, and biological interventions (e.g. drug treatment, gene manipulation) (Cummings et al., 2017).

The history of metabolomics follows the progress of analytic technologies. In the

late 1940s, the research from Roger Williams firstly opened the gate of "metabolic profile" and he revealed that the metabolic components of urine and saliva were associated with schizophrenia using paper chromatography (Gates & Sweeley, 1978). With more advanced analytic technologies applied, researchers were able to precisely determine the metabolic profile in biological samples. In the late 1960s, Horning et al. first used gas chromatography-mass spectrometry (GC-MS) to determine the steroid hormones and their metabolites in human urine samples, and he introduced the term "steroid profile" which is considered the prototype of metabolomics (Horning et al., 1968). After a few years, Hoult and colleagues measured the concentrations of metabolites from intact muscle tissue using phosphate nuclear magnetic resonance (NMR) spectroscopy (Hoult et al., 1974). In 1984, a more sensitive NMR, proton NMR ($^{1}$H NMR), was applied to determine metabolites in serum, plasma, and urine samples of patients with diabetes (Nicholson et al., 1984).

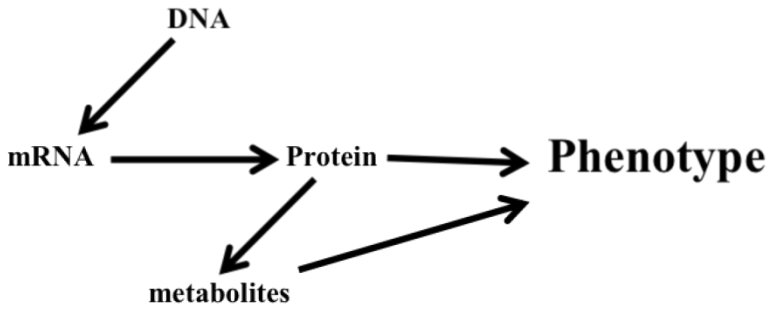



**Figure 1.2** The network of molecular factors affecting phenotype.

Solid black arrows indicate the trajectory of how genetic information is translated into phenotype.

In 1995, researchers used liquid chromatography-mass spectrometry (LC-MS) to study lipid composition in the cerebrospinal fluid of sleep-deprived cats (Cravatt et al., 1995). Afterward, several other combined methods and new techniques such as Fourier-transform infrared spectroscopy (FT-IR) and capillary electrophoresis (CE) was gradually applied in studies of metabolites (Lenz & Wilson, 2007). However, most of the studies introduced were focusing on specific types of metabolites or specific conditions. In 2005, Siuzdak's group released an online database called "METLIN" which is the first integrated tandem mass spectrometry (MS) database in human metabolomics that includes structural information of the known endogenous metabolites, drugs, and drug metabolites (Smith et al., 2005). The public source of METLIN is still updating and growing, making it a useful tool for metabolite identification and largely facilitating the progress of metabolomics. In 2007, three years after the Human Metabolome Project was launched, the Human Metabolome Database (HMDB) was released. HMDB is currently the largest comprehensive database for human metabolome that contains information on metabolites acquired from both literature and experimental data of various techniques (Wishart et al., 2007). As introduced above, most techniques aiming to measure metabolites in biological samples require chromatography to pre-separate the ingredients. The step makes metabolites profiling a high-throughput measurement is the application of methods allowing real-time analysis of metabolic products without chromatographic separation in advance. For example, Winkler et al. used highly sensitive proton-transfer-reaction mass spectrometry to monitor the metabolic products in human breath gas after ingestion of isotope-labeled ethanol (Winkler et al., 2013). Link et al. determined the real-time metabolome change from starving to growing of *Escherichia coli* by direct

injection of samples to a high-resolution time-of-flight MS (Link et al., 2015). The real-time measurement of metabolome is big progress in metabolomics, which makes it available to precisely monitor the change of metabolism since metabolite concentration is dynamic and may change very quickly spontaneously or in response to a particular stimulus.

## 1.5.2 Techniques to study metabolomics

Currently, the most broadly used techniques in metabolome profiling are NMR and MS-based methods which have the potential for the detection, quantification, and identification of metabolites in biofluids or tissues. Some NMR and MS-based methods require chromatographic separations and are coupled with gas chromatography (GC) or liquid chromatography (LC) platforms. Some NMR and MS platforms support the direct loading of samples without pre-separation steps including 1H NMR spectroscopy, direct infusion MS, and FT−IR which are considered relatively high-throughput assays. However, it is currently impossible to use a single technique to identify all metabolites in a sample because of the complexity of the metabolome (Cummings et al., 2017). I will introduce NMR, particularly 1H NMR spectroscopy and frequently-used methods for statistical analysis, in this section since they are the methods that I performed in my thesis research.

NMR and MS are complementary technologies in metabolomics studies and are used for different purposes due to fundamental methodological differences. NMR detects specific molecules at relatively high concentrations while preventing instrument contamination, resulting in high instrument reproducibility. MS detects, separates, and

enriches signals with mass differences from all analytes, producing a large amount of spectroscopic data. Therefore, MS is widely used in studies focusing on detailed characterization of physiology-relevant metabolic specificity of individuals while NMR is tended to in large-scale studies related to epidemiology and genetics (Dunn et al., 2011).

$^{1}$H NMR spectroscopy is a prevalent technique used in research of biofluids such as toxicity and drug efficacy tests, and biomarker identification for diseases or inherited defects in metabolisms (Podolska et al., 2012). $^{1}$H NMR can provide the highest relative sensitivity among naturally occurring spin-active nuclei since the proton is a component of all organic molecules. Sample preparation is very simple for $^{1}$H NMR and it does not require the preselection of analytic parameters or sample derivatization. However, considering biofluid samples contain a large amount of water, a water suppression procedure that attenuates the signal from the water proton is needed. The measurement of $^{1}$H NMR spectra is quick (a few minutes per sample) and maintains a low signal-to-noise rate for milliliter–level biofluid samples (Podolska et al., 2012). All of these features above make $^{1}$H NMR a relatively high-throughput technique in detecting metabolites for broad purposes.

As mentioned at the beginning of this section, there is not an ideal technique that possesses all the desired advantages because of the high complexity of metabolism. However, combining multiple platforms is a possible way to maximize the content and quality of the metabolite profile of interest. Continuing exploration and improvement of analytic techniques will provide strong power for the development of metabolomics.

There are many methods for analyzing the metabolite data obtained from the measuring platform. Unsupervised learning analysis is usually used to classify or extract characteristics only through metabolite data before supervised correlation analysis of metabolite and performance traits or experimental grouping. The most commonly used unsupervised learning methods include principal component analysis (PCA) which effectively reduces the dimensionality of the data set to a few to account for the most significant changes. The goal of the PCA algorithm is to reduce the dimensionality of the data by transforming the original variables into a smaller number of independent variables (called principal components) while retaining most of the information in the original data set (Nyamundanda et al., 2010). For example, Wang's group determined the time-dependent and tissue-specific metabolites change in the context of inflammatory bowel disease using $^{1}$H NMR spectroscopy combined with PCA and orthogonal projection to latent structure-discriminant analysis (OPLS-DA) (Dong et al., 2013). Aside from PCA, the k-means clustering algorithm is another unsupervised learning method. Different from decreasing the number of variables in PCA, the k-means clustering algorithm is used to group multiple observations into a few clusters, which is a common tool in analyzing metabolomics data. The basic principle of the k-means clustering is randomly selecting one subject from the pre-defined K groups, respectively, as the clustering centers, then assigning every other subject to each of the clusters based on the nearest distance to each of the clustering centers, which is finally made up of K clusters (MacQueen, 1967). The similarity of the subjects in a particular cluster can be assessed by coefficients that reflect whether those metabolites are functionally-relevant and have changed in the same manner. For example, Nikas et al. acquired the metabolome profile

of the striatum area of the brain using $^{1}H$ NMR and used the Cophenetic Correlation coefficient to assess the similarity of mice in different models to diagnose Huntington's disease (Nikas & Low, 2011).

Metabolomics data are commonly used in linear models but are affected by multicollinearity. On the other hand, multivariate statistics is an emerging method for high-dimensional related metabolomics data, the most popular method of which is the Projection to Latent Structures (PLS) regression (Abdi, 2010) and its classification version PLS-DA (Discriminant Analysis) (van Velzen et al., 2008). Other data mining methods such as ANOVA-simultaneous component analysis (ASCA) (Smilde et al., 2005) and k-nearest neighbors (KNN) (Shi et al., 2004) have received increasing attention in untargeted metabolomics data analysis. In a univariate approach, variables are analyzed one by one using classical statistical tools (such as Student's t-test, ANOVA, or mixed models), and only those variables with sufficiently small p-values are considered relevant. However, when making multiple comparisons, a correction strategy (such as adjusted p-values or q-values) should be used to reduce false discovery (Kell & Oliver, 2004). For multivariate analysis, the model should always be validated to ensure that the results are universal. In most cases, these types of statistical analyses do not help researchers understand the detailed biological mechanisms behind these associations. Therefore, researchers often combine existing biological function analyses to help understand the underlying logic of the biological process behind these associations. For example, metabolite set enrichment analysis (Persicke et al., 2012) or metabolic pathway analysis (Kankainen et al., 2011) can be used to achieve this point.

After the analysis using the methods introduced above, receiver operating

characteristic (ROC) analysis is a good choice for selecting biomarkers. When determining biomarkers, ROC analysis can be objectively neutral without being influenced by traits/biological mechanisms (Xia et al., 2013).

With the development of scientific technology, although metabolomics analysis techniques are becoming faster, cheaper, more reliable, and more accurate, the simultaneous and unambiguous identification of all metabolites in biological systems remains a challenge. With the establishment of the metabolomics database, a secondary analysis of the published metabolomics data sets in similar populations collected under similar research conditions called “meta-analysis” is performed. “meta-analysis” can improve the accuracy of pattern analysis and biomarker prediction by using the collective power of multiple studies to overcome research bias and small-sizes effects (Tseng et al., 2012).

### 1.5.3 Application of metabolomics in health science studies

The fast-growing knowledge of the importance of metabolites in biological activities makes metabolomics a popular field in understanding the mechanisms of the host in keeping healthy or developing disease. Meanwhile, because the metabolome is considered the final level to influence phenotype and function in system biology, the application-driven nature of metabolomics makes it more broadly applicable in developing biomarkers in various fields to diagnose disease and predict risk in health science and screen favored traits in agriculture (Ghosh et al., 2015). I will give some examples of how researchers apply metabolomics in different fields including livestock studies where my work belongs.

One of the hottest research areas using metabolomics methods is how commensal bacteria interact with the host to maintain homeostasis or influence disease development and treatment (Ursell et al., 2014). Using MS-based techniques, Schultz and Siuzdak's group revealed that gut microbiota has a big impact on the metabolite profile of the host blood and possibly influences the capacity of drug metabolism (Wikoff et al., 2009). Alessio and Ravel's group compared the fecal metabolites from infants having a diet with or without gluten using $^{1}$H NMR. They provided a way to identify potential metabolome biomarkers to predict the onset of celiac autoimmune disease in infants by combining genomic, microbiome, and metabolomics analysis (Sellitto et al., 2012).

Metabolomics is also increasingly applied in medical research, especially the health issues closely related to metabolism such as cardiovascular diseases and diabetes. Barbas's group examined the metabolic state of stable carotid atherosclerosis by analyzing the plasma metabolome using GC-MS and $^{1}$H NMR together. They observed that several metabolites that are strongly associated with insulin resistance were significantly changed in the patients compared with the healthy control, suggesting those metabolites could be potential biomarkers for diagnosis purposes (Teul et al., 2009). Suhre et al. applied multi-platform metabolomics techniques to compare the plasma metabolome of individuals with or without diabetes, and they found metabolites in the pathways related to kidney dysfunction, lipid metabolism, and interaction with the gut microflora were perturbed in diabetic individuals. Their work sheds light on using metabolic biomarkers to diagnose and predict the progression of diabetes and its complications (Suhre et al., 2010). Cancer cells are transformed cells and behave differentially in metabolism pathways compared to healthy cells (Muñoz-Pinedo et al.,

2012). By analyzing the tissue and plasma samples tracked by $^{13}C$-isotopomer using $^{1}H$ NMR in combination with GC-MS, the study by Fan et al. on lung cancer demonstrated the altered anabolic metabolism of cancer cells which improves the understanding of how cancer cells fulfill a high energy demand (Fan et al., 2009). Metabolomics is also widely applied to identify biomarkers to detect and diagnose cancers from various types of samples such as serum, urine, and saliva (Armitage & Barbas, 2014).

Another big aspect that metabolomics has been applied in is livestock study and industry. It appears to be a novel direction in livestock breeding that uses metabolites as a special intermediate phenotype (metabotype) only or in combination with conventional approaches (e.g. genetic background) to link with the preferred phenotypes of economic relevance. One of the main aims of recent studies is to use the data acquired using metabolomics approaches to evaluate and predict reproductive and productive traits (e.g. fertility, growth rate, slaughter weight, milk quality) (Fontanesi, 2016; Goldansaz et al., 2017). The San Cristobal group used $^{1}H$ NMR to generate fingerprints of plasma metabolome from three main European breeds of growing pigs with which to predict production phenotypes. They found that some economic phenotypes including lean meat percentage and average daily feed intake were well-predicted by the corresponding metabolomics data (Rohart et al., 2012). Melzer et al. used GC-MS to analyze the metabolites in milk samples from Holstein cows and found that milk traits were correlated with and could be predicted by particular metabolites. They also reported that the metabotype (metabolites) were less affected by the influencing factors such as farm and sire compared with traditional milk traits, suggesting the metabolite patterns of milk are more reliable phenotypes to predict milk performance (Melzer et al., 2013). With

more and more attention on metabotype in livestock breeding, it is not surprising that genome-wide association studies with metabotypes (mGWAS) emerged, which helps researchers to better understand how genetic variation and environmental factors influence metabolic pathways (Abdi, 2010). One of the typical studies was that Buitenhuis et al. calculated a genomic relationship matrix combining the identified milk metabolites by $^{1}$H NMR and the genotypes of Holstein cows. They found that some milk metabolites have high heritability and they also detected quantitative trait locus (QTL) for some milk metabolites (Buitenhuis et al., 2013). This study uncovered the relationships between genetic background and milk metabolic profiles and provides a theoretical foundation for using metabotypes in livestock breeding.

Aside from the applications in assessing and predicting favorable economic traits, metabolomics is also a useful approach to assessing animal health status and diagnosing diseases as it is applied to human health. The Gronwald group investigated the possibility of using milk metabolites to predict the risk of ketosis which is a common metabolic disease in dairy cows. They performed NMR on milk and plasma samples from Holstein cows and observed that glycerophosphocholine level and glycerophosphocholine to phosphocholine ratio in milk were correlated with a low ketosis incidence, implying those factors could be a prognostic biomarker for ketosis (Klein et al., 2012). Saleem and colleagues studied the effect of a high-grain diet on the rumen health of cows by examining the compounds in rumen fluid using a combination of $^{1}$H NMR, GC-MS, and direct flow injection MS/MS. They observed that some toxic and pro-inflammatory metabolites were enriched in the cows from high-grain-diet groups and those cows also exhibited perturbed amino acid components in the rumen (Saleem et al., 2012).

Although metabolomics is gaining increasing attention in livestock studies, some trends of imbalanced applications can be seen in different research categories. One of the main gaps is that most of the metabolomics studies were related to bovine. In addition, despite the mined data indicating animal health is one of the main research areas applying metabolomics (van Velzen et al., 2008), there still lacks studies on some key health issues that might benefit from metabolomics study, for example, the pathogenesis and prediction of mastitis in dairy cows (Mudaliar et al., 2017) and PRRS in pigs (Schroyen et al., 2016). It is expected to see more applications of metabolomics approaches in other livestock species in the future to help answer fundamental and practical questions with regard to livestock production, reproduction, nutrition, and health.

## 1.6 Research focus and objectives

As stated above, infectious disease outbreaks have caused the modern pig industry to suffer significant financial losses, and the infection frequently consists of a complex of pathogens like PRDC. There are a number of reasons why selective breeding for disease resistance might not be the best way to protect animal health and welfare. First of all, traditional breeding programs focusing on high productive performance have resulted in greater susceptibility for pigs to pathogen challenge, as this selection index theory permits improvement in traits that are negatively associated (Guy et al., 2012). In addition, disease resistance is usually specific to one single type of pathogen, which makes it difficult to effectively control the diseases caused by multiple types of pathogens. Moreover, the currently known key genes and mechanisms that are involved in establishing disease resistance are not enough to guide its application in animal

breeding (Guy et al., 2012; S. Zhao et al., 2012). Matching the goal of modern pig production to maintain high productivity and improve animal welfare, our lab and collaborators are targeting disease resilience aiming to better manage the potential threatening polymicrobial disease challenge while maintaining a sustainable production rate. This is also beneficial to circumvent the disadvantage of current disease treatment by reducing the treatment cost and lowering the risk of antimicrobial resistance (Best et al., 2008; Dawson et al., 2013). We are focusing on testing new phenotypes that can be used to easily and accurately differentiate pigs with disease resilience, and at the same time, exploring potential molecular biomarkers by using multi-omics approaches to predict disease resilience in advance to disease challenge. The overall hypothesis of this study is that pig disease resilience can be predicted by bio-molecules expression level in the peripheral blood before exposure to pathogens.

Based on the introduction about the three types of biomolecules (APPs, gene transcripts, metabolites) above, this study included three main sections. Considering that APPs are important components in the innate immune response and the blood level of some APPs have a dramatic change in responding to infection, the first section of my project is examining specific APPs in the blood in advance to pathogen challenge to test their potential in predicting disease resilience. With the technology progress and the commercial application of high-throughput methods in omics study, the second and third sections of my study were testing whether there were potential biomarkers for disease resilience by extracting the information from the transcriptome of whole blood and plasma metabolome using RNA-Seq and NMR, respectively. The objective of my study was to test whether it is feasible and what specific APPs, gene transcripts, and serum

metabolites can be used as biomarkers to predict pig disease resilience.

# CHAPTER 2: Investigation of using acute phase proteins as biomarkers to predict pig disease resilience

## 2.1 Background

Bishop and Woolliams summarized the modern concept of disease resilience as a trait of an animal with an ongoing infection without diminishing productivity (Bishop & Woolliams, 2014). Given that disease resilience is an immune-related trait that cannot be detected until getting a disease, it is necessary to explore markers to identify higher resilience potential for breeding purposes before pigs get infected. APPs are proteins whose levels change immediately after infection or exposure to other stressors, and this process with concentration change of APPs is called the acute phase reaction (APR). Infection or inflammatory response can lead to the release of APPs from liver hepatocytes into circulation (Baumann & Gauldie, 1994). It has been shown that the secretion of APP is mediated by the pro-inflammatory cytokines including IL-1, IL,-6, and TNF-α (Yoshizaki, 2011). After synthesis, APPs participate in the innate immune response and play important roles in restoring homeostasis and facilitating healing. As introduced above, because of the rapid responding concentration change of APPs (even prior to the manifestation of some symptoms) and the conserved activation process of APR in mammals, APPs are widely used as biomarkers in the diagnosis of infectious disease and inflammation in different types of animals (Cray et al., 2009; Murata et al., 2004). In addition, APPs have also been used to assess the productive performance and herd health of pigs (Clapperton et al., 2005; Piñeiro et al., 2007; van den Berg et al.,

2007). Considering that resilience is related to both health status and productive performance, I hypothesized that APPs could be potential biomarkers to predict pig disease resilience.

There have been studies suggesting that some APPs were associated with infection or inflammatory lesions in pigs (Eckersall, n.d.; Saini & Webert, 1991). Haptoglobin (HP) is one of the strongly reacting APPs in most mammals. It can bind free hemoglobin in plasma generated from hemolysis to neutralize and limit the resulting oxidative damage to organs, and it is also thought to exert a bacteriostatic effect by restricting iron for bacterial growth (Petersen et al., 2004; Putnam, 1975). Besides, HP may play an important role in supporting the proliferation and development of adaptive immune cells (Huntoon et al., 2008). C-reactive protein (CRP) is another strongly reacting APP in humans and pigs and it was reported to function as an opsonin to bind bacteria and free DNA in the innate immune system (Ng et al., 2004; Steel & Whitehead, 1994). CRP was also thought to differentially regulate the classical and the alternative pathway of complement activation during the innate immune response (Suankratay et al., 1998). An early study using turpentine stimulation (which induces a sterile inflammatory lesion) in pigs revealed that the serum concentrations of HP and CRP peaked at 48 hours after treatment and increased by 2.6 fold and 8 fold, respectively, which suggested HP and CRP may be potential markers for the detection of infection or inflammation (Eckersall et al., 1996). Consistently, a study using experimental infection of *A. pleuropneumoniae* in pigs showed that HP and CRP were highly induced and peaked after 48 hours of inoculation with an over 20-fold and around 6-fold increase, respectively (Heegaard et al., 1998). Infection of Myho and PRRSV in pigs was also reported to elevate the serum

concentration of HP (Asai et al., 1999; Magnusson et al., 1999). Considering the quick inducible response and association with host immune response, HP and CRP may have the potential to indicate pig health and immune fitness, and thereby reflect pig resilience. α1-acid glycoprotein (AGP) is an APP found in all mammals which can bind and inhibit the pro-inflammatory activity of lipopolysaccharide and may have broad functions in regulating immune response and inflammation following bacterial infection (Cray et al., 2009; Hochepied et al., 2003; Huang et al., 2012). Even though AGP can be induced by infection and inflammation, its responsive increase in serum is species-dependent (Bteich, 2019). Studies on pigs suggested that the serum concentration of AGP is not remarkably changed during a typical APR compared with other APPs (Eckersall et al., 1996; Petersen et al., 2004). Thus, I included AGP as a non-inducible APP in the pathogen challenge study.

In this chapter, I will focus on those three APPs (HP, CRP, and AGP) to explore whether APP can function as biomarkers to predict disease resilience in pigs under the stress of pathogens in a natural disease challenge model as stated below.

## 2.2 Methods

The University of Alberta Animal Care and Use Committee for Livestock approved the experiments detailed in this thesis (AUP00002227). The pig handling and treatment were in accordance with the guidelines from the Canadian Council on Animal Care (CCAC) (Abul Abbas et al., 2019). All the animal experiments and sample collection were carried out at the Centre for Research in Animal Science Deschambault (CRSAD) for the quarantine and acclimatization unit and the Centre de développement du porc du

Québec inc. (CDPQ) for the test station in Deschambault in Quebec City, Canada. CDPQ had full oversight on the project along with the herd veterinarians.

### 2.2.1 Natural Disease Challenge Model

This natural disease challenge model was designed to imitate a severe health challenge as may be experienced in some commercial pig farms when infectious diseases spread. Our natural disease challenge model recruited healthy, crossbred (Landrace x Yorkshire) barrows (~21 days of age) from seven different multiplier farms which were all enrolled by members of PigGen Canada. Each batch was supplied by a single breeding company. All of these qualified farms were free of, or under the effective control of, the diseases associated with infections including PEDV, TGEV, PRRSV, PCV2, SIV, *Pasteurella multocida, Haemophilus parasuis, Brachyspira hyodysenteriae,* Mhyo, APP, *Staphylococcus hyicus*, and *Sarcoptes scabiei var suis,* confirmed by the veterinarian for at least six months. Until the work of this thesis, this model was established with fourteen batches of pigs (n=893 in total) which were further denominated as cycle 1 for batches 1 – 7 and cycle 2 for batches 8 – 14. For the disease challenge, those healthy weaned pigs (60-75pigs/batch) were transported from their multiplier herds every three weeks to the Centre de developpement du porc du Quebec inc. (CDPQ) test station in Deschambault (Quebec City, Canada).

For the batches in cycle 1, multiplier pigs were initially housed in a nursery unit of the test station for 3 weeks after arriving at CDPQ for quarantine and acclimatization. Then, multiplier pigs were transferred to the test unit for natural disease challenge by introducing the different pathogens with seeder pigs from commercial farms

demonstrating the diseases. The first 4 batches of healthy multiplier pigs were challenged in this way with four groups of seeder pigs (67 pigs in total). Pathogens, including three strains of PRRSV, two strains of SIV, Mhyo, *Haemophilus parasuis*, *Brachyspira hampsonii*, *Salmonella enterica* Serovar typhimurium, and *Streptococcus suis*, *Cystoisospora suis*, and *Ascaris suum*, were confirmed in the seeder pigs within the first week of arrival. Among the 67 seeder pigs, the first three groups (37 pigs in total) of 42-day old pigs were selected from the commercial farms where that had been confirmed positive for PRRSV infection, while the rest were finisher stage pigs (28 in total) from commercial farms with detection of enzootic pneumonia caused by Mhyo. Most healthy multiplier pigs from the first 4 batches were housed with the seeder pigs in separate pens next to each other allowing direct nose-to-nose contact. Only 7-8 multiplier pigs from batch 4 were arranged with 2 seeder pigs in the same pen to induce cross-infection. After confirming the establishment of diseases in the 4 batches of multiplier pigs, they worked as "new seeder pigs" to challenge the next batches of newly arrived healthy multiplier pigs in the grower-finisher unit, and this continuous challenging pattern applied to all the following healthy multiplier pigs without introducing any more unhealthy commercial pigs (Figure 2.1). After 6 weeks from arrival at CDPQ, the multiplier pigs were transferred to the grower-finisher unit in the test station until slaughtering age or weight.

For the healthy quarantine unit, the test unit of the late nursery stage, and the test unit of the grower-finisher stage, respectively, there were approximately 4, 7, and 13 pigs per pen for each stage. It should be noted that the nursery and test units were located in the same building during cycle 1, separated by a hallway and under tight biosecurity supervision (Figure 2.1). However, biosecurity practices failed to prevent pathogen

dissemination from the test unit to the newly arrived multiplier pigs. Thus, for the cycles following cycle 1, a nursery unit (CRSAD) located 1 km away from the test unit was established to keep it free of infection during the nursery stage. Besides, feed medication during the grower-finisher phase was pulsed and vaccination for PCV2 (Ingelvac CircoFLEX®) two weeks after entry was applied to the multiplier pigs to reduce mortality and treatment frequency upon challenge.

This model allowed us to assess the defined phenotypes and genetic variations associated with disease resilience in an environment simulating real commercial situations including infections by multiple pathogens and cross-transmission among pigs. The most important reason for using this model is that we aimed to use the animal data and samples collected before the pathogen challenge to predict the health status and productive performance of pigs under pathogen challenge. Thus, the samples used in this experiment were collected within the first week of the pig's arrival from the breeding company at the healthy quarantine nursery to avoid contamination in pigs before the challenging step.

Besides cycles 1 and 2 described above, another 4 cycles of multiplier pigs (1850 pigs in total) were introduced in this natural disease challenge model following the protocol for Cycle 2. Seven batches (26-32) of pigs from Cycle 4 and Cycle 5 were selected for the metabolome analysis experiment including a total of 460 individuals (described in Chapter 4). Four batches (8, 9, 10, 12) of pigs in total of 209 pigs from Cycle 7 were selected to validate the findings of transcriptome analysis (described in Chapter 3).

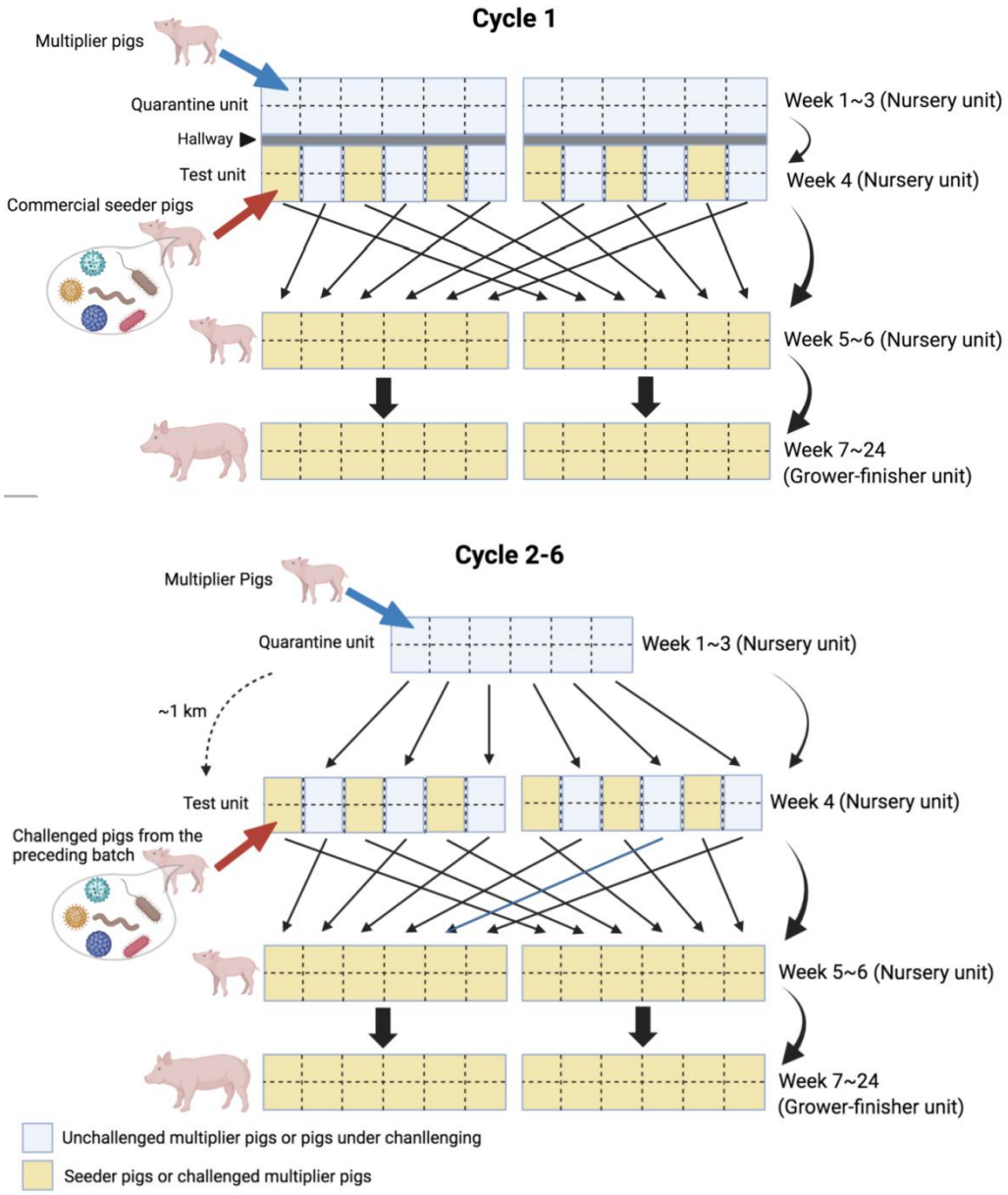



**Figure 2.1** Schematic of pig arrangement in the test station of the natural disease challenge model.

Multiplier pigs were challenged after three weeks of arrival at the test station by allowing nose-to-nose contact with commercial seeder pigs (cycle 1) or the preceding batch of challenged pigs (cycles 2-6) to predict the health status and performance (phenotypes) of pigs under pathogen challenge. Thus, the samples used in this experiment were collected within the first week of the pig's arrival from the breeding company at the healthy quarantine nursery to avoid contamination in pigs before the challenging step.

## 2.2.2 Tracking of pig health status and performance traits

Random blood samples from multiplier pigs (16 to 20 in total, 4 pigs per pen) after 4 weeks and 6 weeks post challenging were collected to confirm infection of specific pathogens including PRRSV, SIV, Mhyo, and *Actinobacillus pleuropneumoniae*, using ELISA. The level of mortality, clinical symptoms, and treatments was carefully monitored and recorded weekly by the veterinary advisory team during the whole experiment. Pigs identified with disease exhibiting clinical symptoms were treated individually by injectable medication directed and approved by the veterinarians. Group medication treatment through watering (nursery stage) or feeding (grower-finisher stage), and even amendment of challenging pattern (e.g. suspending nose-nose contact) was applied for the consideration of animal welfare when an increased mortality rate was observed (e.g. a cutoff of 8% for nursery stage). The phenotype data "treatment frequency" indicates the number of treatments that an individual pig received during the experiment and involves the injection of any drug. It does not take into account group treatment through medicated water or feed.

Clinical indications, disease severity, and treatment response including those with intolerable suffering from illness and non-response to treatment—were taken into consideration while deciding whether to carry out euthanasia (defined as humane end points for the model). Certain circumstances, such as a fractured or dislocated joint, non-weight-bearing lameness that persisted after seven days of treatment, and acute peripheral cyanosis coupled with stomach breathing, also warranted performing an emergency euthanasia procedure without the consent of veterinarians. All the humane interventions followed the regulations established by the Animal Protection Committee and an annual

approval certificate with authorization number 15PO283 issued by Comité de Protection des Animaux.

It should be pointed out that only the cases of death due to infectious disease were counted for mortality in this study. However, the death that resulted indirectly from infections, such as traumatic injuries, herniation, sarcoptic mange, and death caused by blood sampling, was excluded from the data analysis.

Performance traits in this study include age, the weight at weaning, nursery, and grower-finisher recorded every three weeks, slaughter age and weight, backfat and muscle thickness at slaughter, carcass weight, and lean yield. The daily feed intake of nursery pigs was recorded manually by pen while for each individual grower-finisher, daily feed intake amount and frequency were automatically archived specifically to their electronic ear tag using the Insentec IVOG system. Pigs without symptoms and reaching the targeted off-test live weight (130 kg) or the age of 24 weeks were marketed. Death age and weight of pigs including the euthanized ones were also recorded. As briefly mentioned above, daily mortality and treatment rate and corresponding causes were carefully recorded for assessment by the veterinary advisory team on a weekly basis.

## 2.2.3 Blood sample collection and preparation

The blood samples were collected from multiplier pigs in the natural disease challenge model described above at around 26, 54, and 82 days of age (Table 2.1). Bleeding was performed by the veterinarian to collect the blood in Vacutainer® blood collection tubes (BD, New Jersey, USA) containing Ethylene Diamine Triacetic Acid (EDTA). Blood samples were then processed to overnight shipping on ice to the

University of Alberta for the following preparation within 24 to 48 h upon sampling.

To determine the complete blood count (CBC), the blood samples were homogenized and then loaded on the ADVIA® 2120i Hematology System (Siemens Healthineers, Erlangen, Germany). The CBC data used in this study include (1) the exact cell number count ($10^3$/μL): white blood cells (WBC), red blood cells (RBC), lymphocytes (Lym), monocytes (Mono), neutrophils (Neu), eosinophils (Eos), basophils (Bas), platelets; and (2) the relative percentage of Lym, Mono, Neu, Eos, and Bas in total WBCs. For plasma preparation, EDTA tubes were centrifuged at 1,500 rpm for 10 minutes at 4°C to separate the plasma from blood cells. The plasma samples were selected to determine APP concentration from cycle1 batch3, due to its appropriate morbidity and mortality which will be explained below. An ELISA kit (Life Diagnostics, West Goshen, USA) was then used to measure the concentration of the three APPs (HP, CRP, AGP). The 10X Diluent and 20X Wash Solutions from the ELISA kit were diluted to the 1X working concentration using sterile distilled water. According to the physiological concentration range of the APP proteins, plasma samples were diluted by 1600, 500, and 2000 fold for HP, CRP, and AGP testing, respectively, with 1x Diluent solution before the ELISA assay which will be described below.

## 2.2.4 Pig sample selection for APP study

The plasma samples were chosen after the collection of all samples from cycle1 pigs. According to the overall mortality record of the 7 batches in cycle 1 (Table 2.2) batches 4, 5, and 6 had a high death rate of around 50% compared to other batches, which indicates an overwhelming response of those pigs to the disease possibly due to over-challenging.

In contrast, the mortality rate was less than 5% in batch 1, indicating the challenge was too subtle. Both extreme cases are not appropriate to be used for APP determination. From this point, batches 2, 3, and 7 were considered candidates.

Further inspection of the sickness and mortality rate of those 3 batches of pigs at the times of each bleeding (Table 2.3), showed there were 70% sick pigs in batch 2 at the first time of bleeding while batches 3 and 7 only had 1 and 5 sick pigs respectively. Batch 2 was excluded since the high sick rate was observed at first bleeding only after 1 week of arrival in the test station even though the pigs were maintained in quarantine units. For the two candidate batches left, batch 7 maintained low sick rate ( ≦ 10%) within the whole experiment while batch 3 pigs with intermediate death (17.65%) and sick (1.69%-39.22%) rate appeared to be more appropriate to assess different responses in the natural disease challenge model. Batch 1 had continuous mild sick rates (≦10%) indicating insufficient disease pressure. Therefore, to better differentiate pig responses to pathogens based on their state of health, batch 3 from cycle 1 was selected for the APP analysis.

**Table 2.1** Blood sampling for each batch of test pigs in the natural disease challenge model.

The pig age listed was approximate for the whole batch. Old and new quarantine unit was described above that are differentiated by the distance from the test unit. EDTA indicates tubes with EDTA (BD Vacutainer® blood collection tubes, New Jersey, USA) to prevent blood coagulation for plasma APP measurement. Tempus indicates Tempus™ Blood RNA tube (Thermo Fisher Scientific Inc., Wilmington, USA) for RNA extraction which will be described in the next chapter. N/A indicates sample collection was not applied in this study.

<table>
<tr><th>Bleeding</th><th>Time after arrival (days)</th><th>Pig age (day±2)</th><th>Pig location (Cycle 1)</th><th>Pig location (Cycle 2-6)</th><th>Pen type</th><th>Sample tube</th><th>Sampled pigs</th></tr>
<tr><td>1</td><td>6</td><td>26</td><td rowspan="2">Quarantine unit (old)</td><td rowspan="2">Quarantine unit (new)</td><td rowspan="3">Nursery</td><td>EDTA,Tempus</td><td rowspan="4">Whole batch</td></tr>
<tr><td>2</td><td>20</td><td>40</td><td>N/A</td></tr>
<tr><td>3</td><td>34</td><td>54</td><td rowspan="2">Test unit</td><td rowspan="2">Test unit</td><td>EDTA,Tempus</td></tr>
<tr><td>4</td><td>62</td><td>82</td><td>Grower-finisher</td><td>EDTA,Tempus</td></tr>
</table>

**Table 2.2** Mortality rate of the 7 batches of test pigs at the end of cycle 1

| Batch | Total number | Death | Mortality rate % |
|---|---|---|---|
| 1 | 56 | 2 | 3.57 |
| 2 | 69 | 16 | 23.19 |
| 3 | 60 | 18 | 30.00 |
| 4 | 76 | 40 | 52.63 |
| 5 | 60 | 32 | 53.33 |
| 6 | 59 | 29 | 49.15 |
| 7 | 55 | 9 | 16.36 |

**Table 2.3** Mortality and sick rate of the 7 batches of test pigs from cycle 1 at each time of bleeding

| | 1st bleeding | | | | | |
|---|---|---|---|---|---|---|
| Batch | Date | Health | Sick | Sick rate | Death | Mortality rate |
| 1 | 15/11/19 | 55 | 0 | 0.00 | 0 | 0.00 |
| 2 | 15/12/7 | 40 | 28 | 70.00 | 1 | 1.47 |
| 3 | 15/12/29 | 59 | 1 | 1.69 | 0 | 0.00 |
| 4 | 16/1/19 | 73 | 3 | 4.11 | 0 | 0.00 |
| 5 | 16/2/9 | 59 | 1 | 1.69 | 0 | 0.00 |
| 6 | 16/3/1 | 36 | 23 | 63.89 | 0 | 0.00 |
| 7 | 16/3/22 | 50 | 5 | 10.00 | 0 | 0.00 |
| | 2nd bleeding | | | | | |
| 1 | 15/12/15 | 49 | 5 | 9.26 | 1 | 1.85 |
| 2 | 16/1/5 | 37 | 31 | 45.59 | 0 | 0.00 |
| 3 | 16/1/26 | 47 | 13 | 21.67 | 0 | 0.00 |
| 4 | 16/2/16 | 58 | 14 | 19.44 | 3 | 4.17 |
| 5 | 16/3/8 | 49 | 8 | 14.04 | 3 | 5.26 |
| 6 | 16/3/29 | 18 | 34 | 65.38 | 7 | 13.46 |
| 7 | 16/4/19 | 48 | 3 | 5.88 | 4 | 7.84 |
| | 3rd bleeding | | | | | |
| 1 | 16/1/12 | 47 | 7 | 12.96 | 0 | 0.00 |
| 2 | 16/2/2 | 34 | 26 | 43.33 | 8 | 13.33 |
| 3 | 16/2/23 | 31 | 20 | 39.22 | 9 | 17.65 |
| 4 | 16/3/5 | 20 | 26 | 56.52 | 27 | 58.70 |
| 5 | 16/4/5 | 22 | 10 | 31.25 | 25 | 78.13 |
| 6 | 16/4/26 | 15 | 18 | 54.55 | 19 | 57.58 |
| 7 | 16/5/7 | 45 | 5 | 10.00 | 1 | 2.00 |

## 2.2.5 ELISA assay for pig HP, CRP, and AGP

The ELISA assay used in this chapter was performed following the manufacturer's instructions with minor modifications. Briefly, APP standard protein from the kit was reconstituted by adding 1 ml of the 1x diluent and mixed gently until dissolved and would be good to be used within 4 hours of reconstitution. Working concentrations of standard APP proteins including 150, 75, 37.5, 18.75, 9.38, and 4.67 ng/ml were prepared by serial dilution using the 1x diluent.

Standards or diluted samples (100 µl/well) in duplicate were added to the coated strip wells (96 well plates) secured in the plate holder and then incubated on an orbital micro-plate shaker at 150 rpm at room temperature for 45 minutes. The supernatant was removed and the wells were washed for 5 times with 400µl 1x wash solution using a plate washer and then struck sharply onto absorbent paper or paper towel to remove all residual liquid. The entire wash procedure was performed as quickly as possible after incubation. HRP-conjugated antibody was dispensed into each well (100 µl) and the plate was incubated on an orbital micro-plate shaker at 150 rpm at room temperature for 45 minutes. After incubation, wells were washed as detailed above. Then, 100 µl of TMB reagent was added into each well and further incubated on an orbital micro-plate shaker at 150 rpm at room temperature for 20 minutes. The reaction was stopped by adding 100µl of Stop solution to each well and gently mixing the plate to make sure that the blue color was totally changed into yellow. The ELISA result was determined by measuring the absorbance at 450 nm with a microtiter plate reader within 15 minutes after the reaction was stopped.

The standard curve was constructed by plotting the absorbance values of the standards versus the concentrations. The actual concentrations of plasma APPs were calculated by plugging the absorbance into the standard curve and then multiplying the dilution factor. The samples were re-diluted appropriately and the ELISA assay was repeated when the A450 values of a tested sample fall out of the range of the standard curve.

### 2.2.6 Statistical analysis

One-way ANOVA was used to compare the APP concentrations in different groups. Significance is labeled in following plots for the comparison between two specific groups where the P-value was less than 0.05 (*=P<0.05, **=P<0.01, ***=P<0.001, ****=P<0.0001). Spearman correlation analysis was performed to analyze the correlation of APP concentrations or concentration changes with series of pig health and productive traits.

## 2.3 Results

### 2.3.1 Comparison of APP plasma concentration among the selected six groups of pigs in the natural disease challenge model.

Among the 60 pigs from batch 3/cycle 1, the pigs were further selected to form 6 groups (Table 2.4) whose plasma samples were used in the APP study. Group 1 and Group 2 pigs all stayed healthy during the whole experiment while having the lowest and

highest body weight at the end of the experiment, respectively. The first two groups represent the resistant phenotype with differential productive performance. Group 3 was found sick at the early time of the disease challenge (week 4 of arrival) but recovered quickly and stayed healthy during the following period of the experiment, which represents a quick resolving phenotype. Group 4 were pigs found sick also at the early point (week 4 of arrival) and still suffered the disease afterward till the end or even died quickly. Group 4 was thus considered the susceptible phenotype. The rest of the pigs from batch 3 were detected with clinical signs at a relatively later stage of the experiment, and the ones that stayed alive and maintained relatively high productive performance were assigned to Group 5 whereas the ones died and died at the late stage of the experiment were designated as Group 6. Of note, Group 5 is more similar matching the criteria of disease resilient but Group 6 was the worst type of pigs in actual farming cases that develop disease and die at the relatively later stage of pig production and cause dramatic economic loss.

Blood samples were collected from the multiplier pigs following the described plan in Table 2.1, and the ELISA assay was performed to determine the plasma concentration of AGP, HP, and CRP from the grouped pigs (groups 1-6) listed in Table 2.4. While there was no significant difference among the six groups of pigs for HP and CRP plasma concentration before challenging, AGP was found to present a significantly higher concentration in Group 6 than most of the other groups except Group 1 (Figure 2.2A). In contrast, the “resilient” Group 5 showed the lowest AGP level. No significant difference between the three APPs was detected at the early stage (Blood 3) post-disease challenge and for the later stage (Blood 4), a significant difference was only observed for HP

**Table 2.4** Grouping of the pigs from batch 3 of cycle 1 and the weight record during the stay in the test station.

The record of grouped pigs is highlighted by a light grey background and the shadow area indicates where the data is not applicable.

| CDPQ Tag | Weight 1 (lbs, 21d) | Weight 2 (lbs, 42d) | Weight 3 (lbs, 63d) | Weight 4 (lbs, 84d) | Weight 5 (lbs, 105d) | Weight 6 (lbs, 126d) | Weight 7 (lbs, 147d) | Weight 8 (lbs, 168d) | Premature death date | Premature death weight (lbs) | Grouping |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 171 | 7.3 | 12.5 | 19.5 | 21.5 | 31.5 | 47.5 | 58.5 | 75.5 | | | |
| 158 | 8.4 | 15 | 19.5 | 28 | 35 | 45.5 | 59 | 76 | | | |
| 145 | 6.8 | 11 | 16.5 | 24 | 33.5 | 44.5 | 55 | 82.5 | | | Group 1 |
| 172 | 7.9 | 12 | 18.5 | 24 | 29.5 | 45.5 | 61.5 | 82.5 | | | |
| 140 | 6.6 | 12.5 | 18 | 26.5 | 36.5 | 50 | 65.5 | 85.5 | | | |
| 162 | 8.4 | 15.5 | 19 | 30 | 39 | 58 | 65.5 | 91.5 | | | |
| 181 | 7.3 | 12.5 | 19.5 | 25 | 37 | 53 | 70.5 | 93 | | | |
| 168 | 7.3 | 13.5 | 21 | 28.5 | 46 | 62 | 76 | 93 | | | |
| 159 | 7.5 | 13 | 20 | 25.5 | 38.5 | 59 | 73 | 93.5 | | | |
| 165 | 5.7 | 10.5 | 18 | 25.5 | 40 | 62 | 76 | 93.5 | | | |
| 132 | 5.3 | 9.5 | 15 | 24 | 40 | 56 | 74 | 96.5 | | | |
| 160 | 7.9 | 11 | 17 | 28 | 43.5 | 66.5 | 77 | 96.5 | | | |
| 166 | 7.3 | 13 | 18.5 | 26 | 39 | 60 | 74.5 | 97 | | | |
| 183 | 7.4 | 13.5 | 20 | 27.5 | 41 | 57.5 | 76 | 100 | | | |
| 129 | 5.6 | 11.5 | 12.5 | 24.5 | 37.5 | 59 | 72.5 | 102 | | | |
| 154 | 7.8 | 15.5 | 24.5 | 30 | 43 | 61.5 | 77.5 | 102 | | | |
| 156 | 8.6 | 14.5 | 21.5 | 31.5 | 42.5 | 69 | 83 | 103 | | | |
| 176 | 7.7 | 14 | 20.5 | 28.5 | 46 | 63.5 | 81.5 | 103.5 | | | |
| 157 | 7.8 | 14.5 | 22.5 | 32.5 | 47.5 | 66 | 80.5 | 103.5 | | | Group 2 |
| 142 | 6.8 | 12.5 | 19 | 30 | 45.5 | 60 | 81.5 | 104.5 | | | |
| 185 | 7.5 | 11.5 | 20 | 30 | 48.5 | 68 | 82 | 104.5 | | | |
| | | | | | | | | | | | |
| 128 | 4.7 | 9.5 | 14 | 22 | 34 | 51.5 | 63 | 85.5 | | | |
| 138 | 5.6 | 10 | 14 | 22 | 36 | 53.5 | 66 | 92 | | | |
| 177 | 8.3 | 14 | 21 | 26 | 38.5 | 55.5 | 71.5 | 97.5 | | | Group 3 |
| 147 | 6.6 | 12 | 20.5 | 31.5 | 52 | 69.5 | 87.5 | 109 | | | |
| 155 | 7.5 | 13 | 19.5 | 25.5 | 32 | 47 | 62 | 88.5 | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 141 | 5.8 | 10.5 | 14.5 | 18.5 | 27.5 | 43 | 55.5 | 74 | | | Group 4 |
| 163 | 5.8 | 9 | 15.5 | 20.5 | 23.5 | 36.5 | 46 | 66 | | | |
| 143 | 6.3 | 11 | 17.5 | 16.5 | | | | | 22-Feb-16 | 16 | |
| 151 | 6.2 | 10.5 | 16.5 | 20 | | | | | 27-Feb-16 | 16 | |
| 180 | 7 | 10.5 | 18 | | | | | | 17-Feb-16 | 16 | |
| 184 | 7.8 | 10.5 | 14.5 | 17.5 | 24.5 | 31.5 | 42.5 | | 26-Apr-16 | 50 | |
| | | | | | | | | | | | |
| 161 | 8.8 | 14 | 22.5 | 27 | 38 | 59 | 76 | 96.5 | | | Group 5 |
| 135 | 5.2 | 10 | 16.5 | 24 | 40 | 56.5 | 72.5 | 94 | | | |
| 174 | 6.8 | 11.5 | 18.5 | 25 | 36 | 55.5 | 70.5 | 93.5 | | | |
| 150 | 7.3 | 11.5 | 21 | 26 | 38.5 | 57 | 65.5 | 88 | | | |
| 175 | 7 | 12.5 | 18 | 20 | 32.5 | 52 | 65 | 85 | | | |
| 144 | 6.3 | 11.5 | 19 | 24 | 31 | 41.5 | 55 | 84 | | | |
| 170 | 7.6 | 16.5 | 20 | 18.5 | 31 | 46.5 | 55 | 74.5 | | | |
| 179 | 7.8 | 13 | 21 | 22.5 | 30 | 40.5 | 51.5 | 72.5 | | | |
| 131 | 5.7 | 10.5 | 20.5 | 28.5 | 29 | 34.5 | 40 | 62 | | | |
| 178 | 8.6 | 14 | 21 | 20.5 | 22.5 | 34 | 39.5 | 60 | | | |
| 136 | 5.7 | 12 | 16 | 16.5 | 21 | 31.5 | 39 | 51 | | | |
| 146 | 7.7 | 14.5 | 17 | | | | | | 4-Feb-16 | 13 | |
| 149 | 7.6 | 12.5 | 15.5 | | | | | | 4-Feb-16 | 13 | |
| 130 | 6 | 12 | 18.5 | | | | | | 16-Feb-16 | 16 | |
| 173 | 6.6 | 12 | 16 | | | | | | 17-Feb-16 | 12 | |
| 137 | 6 | 12.5 | 18.5 | 17 | | | | | 21-Feb-16 | 16 | |
| 153 | 6.3 | 12.5 | 17 | 13 | | | | | 21-Feb-16 | 15 | |
| 148 | 7.9 | 13.5 | 18.5 | 17 | | | | | 24-Feb-16 | 15 | Group 6 |
| 152 | 7.2 | 13.5 | 19.5 | 24 | | | | | 25-Feb-16 | 20.5 | |
| 182 | 8.4 | 14.5 | 22.5 | 26 | | | | | 29-Feb-16 | 26 | |
| 164 | 8 | 12.5 | 20 | 18 | | | | | 1-Mar-16 | 20 | |
| 134 | 6.7 | 13 | 20 | 23 | 17.5 | | | | 15-Mar-16 | 17.5 | |

between Groups 3 and 6. Therefore, the plasma concentration of HP and CRP before and after the challenge seems hard to reflect the feature of any of the six groups of pigs in the context of this natural disease challenge model. Interestingly, the AGP plasma level stayed relatively lower in the "resilience" group compared to the rest of the groups across the three measurements.

Heatmaps were generated and shown in Figure 2.2B to better describe the trend of concentration change for each APP in the six groups. It is clear that all groups' plasma levels of HP and CRP rose sharply in response to the disease challenge, which is consistent with earlier research showing that these two proteins functioned as positive APPs in pigs' APR. In contrast, AGP was maintained reasonably stable in plasma with little fluctuation compared to the other two APPs evaluated in this experiment. To further look into the APP response in individual pigs instead of pooling results across the three time points, curves were plotted depicting specific APP changes in each pig of the six groups and are shown in Figure 2.2 C. Similar to showing the trend in the heatmaps that CRP and HP increased dramatically as with the disease challenge, however, in contrast to most pigs in each group, some pigs were detected with reducing concentration from the early stage of challenging to the later stage. Interestingly, this rise-then-drop trend across the three bleedings was not found in all the pigs from the "susceptible" group (Group 4), implicating either a non-sustained immune response that cannot efficiently control infection or an overwhelming immune response that largely impaired host homeostasis. Nevertheless, it is still difficult to differentiate the two production-favoring groups (Group 2 and Group 5) from the others.

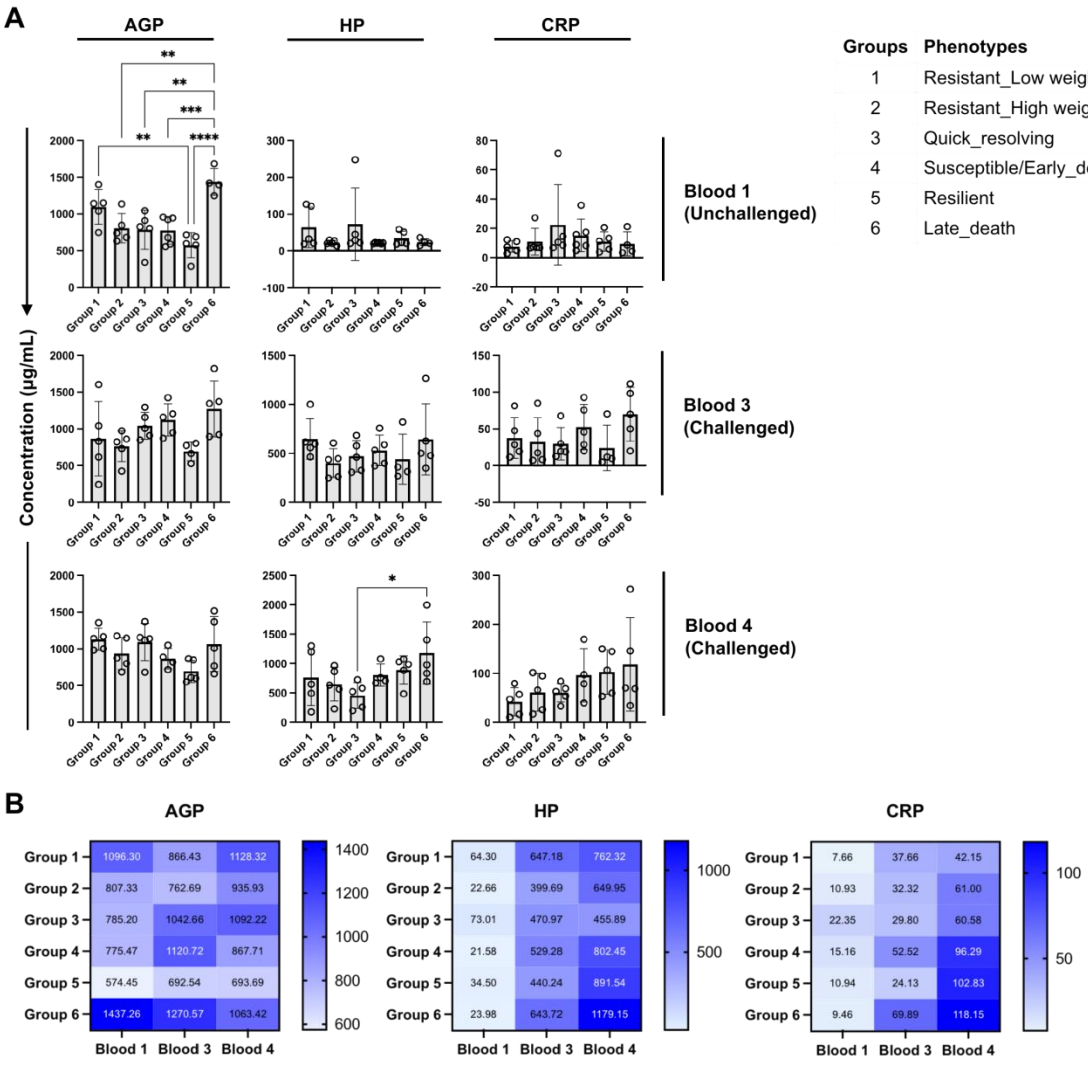
Figure 2.2 illustrates the comparison of APP concentrations side-by-side among the

three blood samples and the six groups of pigs, it is also worthwhile to compare the time course fold change of each APP in different groups. For this purpose, three ratios of APP concentration were calculated to show the change of APP from before-challenging to early-challenged (B3/B1), before-challenging to late-challenged (B4/B1), and early-challenged to late-challenged (B4/B3) as listed in Figure 2.3. The fold changes of AGP were all around 1 in all the six groups, indicating its plasma concentration is independent of our challenging impact. The two positive APPs, HP and CRP, responded rapidly to the initial disease challenge with HP (~19.4 fold) showing a greater changing fold than CRP (~5.2 fold). At a later stage after challenging, Groups 2 and 5 still showed an increase of CRP for around 5 to 10 fold from the early challenging stage while HP concentration was only raised slightly afterward. The resilience group was found to have the smallest changing fold of HP responding to the initial challenge (B3/B1) but there was no significance detected compared to the rest of the groups (Figure 2.3). It is interesting that the Group 6 showed the most dramatic responding level of HP and reached significance when compared with Group 3 for B4/B1, suggesting that HP level changed differently between those two groups in responding to the disease challenge. This phenomenon was similarly detected for the change of CRP. However, in contrast to other groups, the CRP level in the resilience group did not strongly respond to the initial challenge but greatly up-regulated at the later stage in our model. This exclusive changing pattern found in the resilience group (Group 5) may indicate a relatively dull response of CRP to acute infections. Therefore, I successfully recapitulated the response of AGP, HP, and CRP in our disease model as in previous studies. There was a unique responding pattern of CRP in the resilience group, making CRP a better potential marker to differentiate the

economy-favorable pigs than HP and AGP.

### 2.3.2 Correlation of APP plasma concentration with pig health and productive traits

In real farms, how livestock defend against a complex disease challenge is primarily based on their immune system. To varying degrees, livestock genetic variations affect their defensive responses to diseases such as resistant or susceptible. Meanwhile, the host's ability to control the disease has an impact on livestock productivity, either directly or indirectly. Thus, precise characterization of disease resilience relies on both health and productive traits. As the results showed above that the three APPs did respond differentially to our disease challenge model among the six groups of pigs, it is meaningful to dissect the associated health and productive traits that we used to define the pig groups to characterize how each of these traits correlates with the APP concentration or its changing amplitude. For this purpose, the Spearman correlation test was performed on the data of nine single measurements of APPs (3 bleeding for 3 APPs) and a nine-fold change of APPs (B3/B1, B4/B1, B4/B3 for 3 APPs) collected from the six groups of pigs with a series of health and productive records.


A
AGP
HP
CRP
Concentration (µg/mL)
Blood 1 (Unchallenged)
Blood 3 (Challenged)
Blood 4 (Challenged)
Group 1
Group 2
Group 3
Group 4
Group 5
Group 6
Groups Phenotypes
1 Resistant_Low weight
2 Resistant_High weight
3 Quick_resolving
4 Susceptible/Early_death
5 Resilient
6 Late_death
B
AGP
HP
CRP
Group 1 1096.30 866.43 1128.32
Group 2 807.33 762.69 935.93
Group 3 785.20 1042.66 1092.22
Group 4 775.47 1120.72 867.71
Group 5 574.45 692.54 693.69
Group 6 1437.26 1270.57 1063.42
Group 1 64.30 647.18 762.32
Group 2 22.66 399.69 649.95
Group 3 73.01 470.97 455.89
Group 4 21.58 529.28 802.45
Group 5 34.50 440.24 891.54
Group 6 23.98 643.72 1179.15
Group 1 7.66 37.66 42.15
Group 2 10.93 32.32 61.00
Group 3 22.35 29.80 60.58
Group 4 15.16 52.52 96.29
Group 5 10.94 24.13 102.83
Group 6 9.46 69.89 118.15
Blood 1 Blood 3 Blood 4

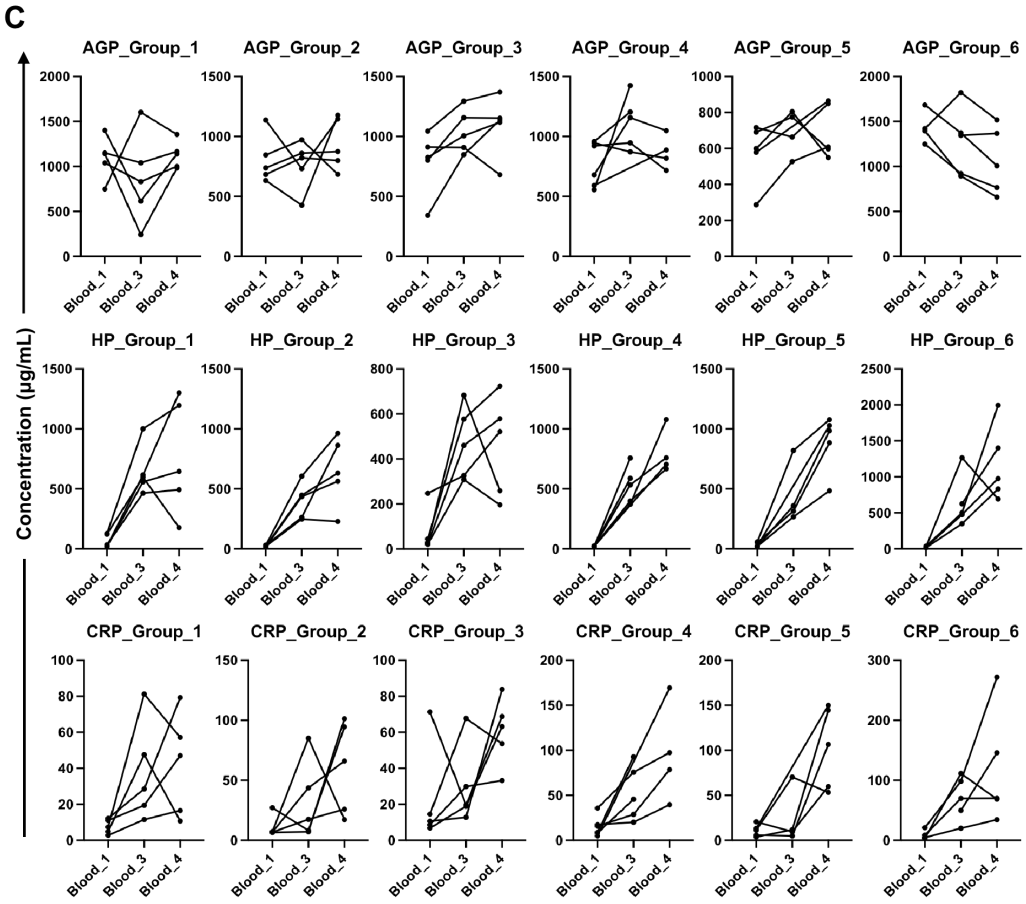



**Figure 2.2** Comparison of plasma AGP, HP, and CRP concentration determined by ELISA among the defined six groups of pigs in batch 3 of cycle 1 (Table 2.4) in the natural disease challenge model.

(A) Statistical comparison of the three APPs' plasma concentrations in the six groups at the time of each bleeding. Phenotypes of each pig group defined by the health and production records are listed on the right (B) Heatmap showing time course change of the plasma concentration for the three APPs before and after challenging. The numbers listed inside of the plots are the mean values of APP concentrations pooled from each group of pigs. (C) Tracking of APP concentration for each pig in the groups from the three times of bleeding during the experiment.

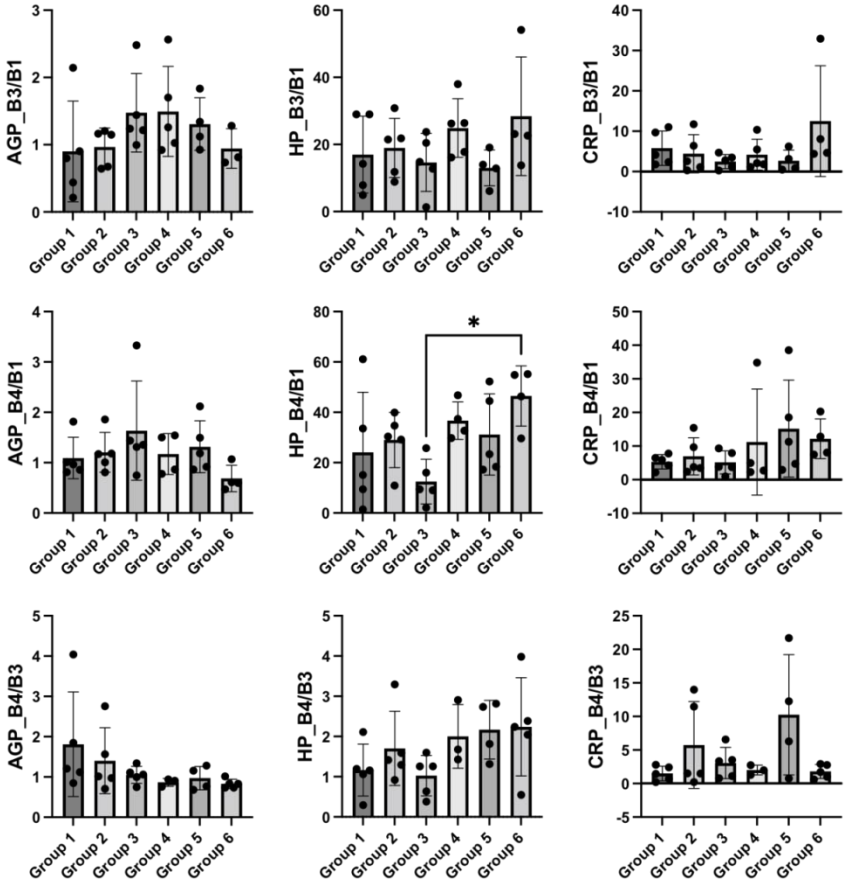
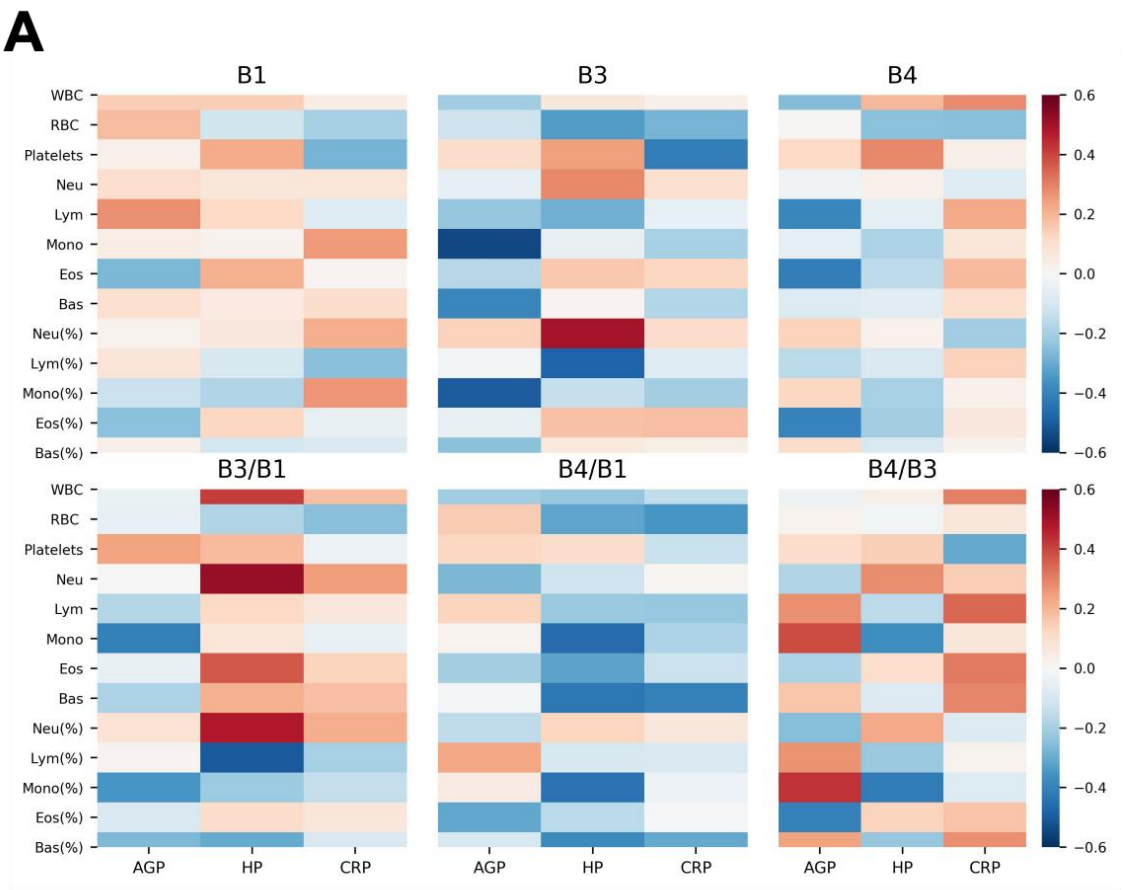



**Figure 2.3** Fold change of plasma concentration between samples from each bleeding for the three APPs.

The ratio of B3 (Blood 3)/B1 (Blood 1), B4 (Blood4)/B1, and B4/B3 were calculated for each pig and then pooled based on the grouping shown in Table 2.4.

Peripheral blood features are a critical readout of the host's health status in responding to infections. We focused on the CBC data that were acquired from the three times bleeding and represented the peripheral blood cell concentration and composition at the same time when the APP concentrations were determined. As shown in Figure 2.4 A, AGP plasma concentration was generally positively correlated with most blood cells except eosinophils before challenging while switching to negatively correlated upon disease challenge. In comparison, HP and CRP showed various and distinct correlations with different types of blood cells. Of note, a strong positive correlation between HP and Neu(%) was observed early after the disease challenge but faded afterward, indicating an acute response of neutrophil proliferation in our model during early infection. This correlation was not dramatic for CRP probably due to less abundant change of plasma concentration in response to challenge compared with HP. Interestingly, CRP showed a slower induction than HP in our model (Figure 2.3), this is reflected by its stronger correlation with blood cell change from B4/B3 than that from B3/B1 whereas HP has the reverse trend. Among the plots shown in Figure 2.4 B, a strong positive correlation between HP and WBC from B3/B1 suggests a quick mobilization of the innate immune system during early response while the positive correlation between CRP and Lym from B4/B3 may indicate the proliferation of adaptive immune response which happens at a relatively later stage of infection. The broad variations observed for these three APPs in terms of their association with different types of blood cells indicate they play distinct roles before and after infection is established.


A
B1
B3
B4
B3/B1
B4/B1
B4/B3
WBC
RBC
Platelets
Neu
Lym
Mono
Eos
Bas
Neu(%)
Lym(%)
Mono(%)
Eos(%)
Bas(%)
AGP
HP
CRP
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

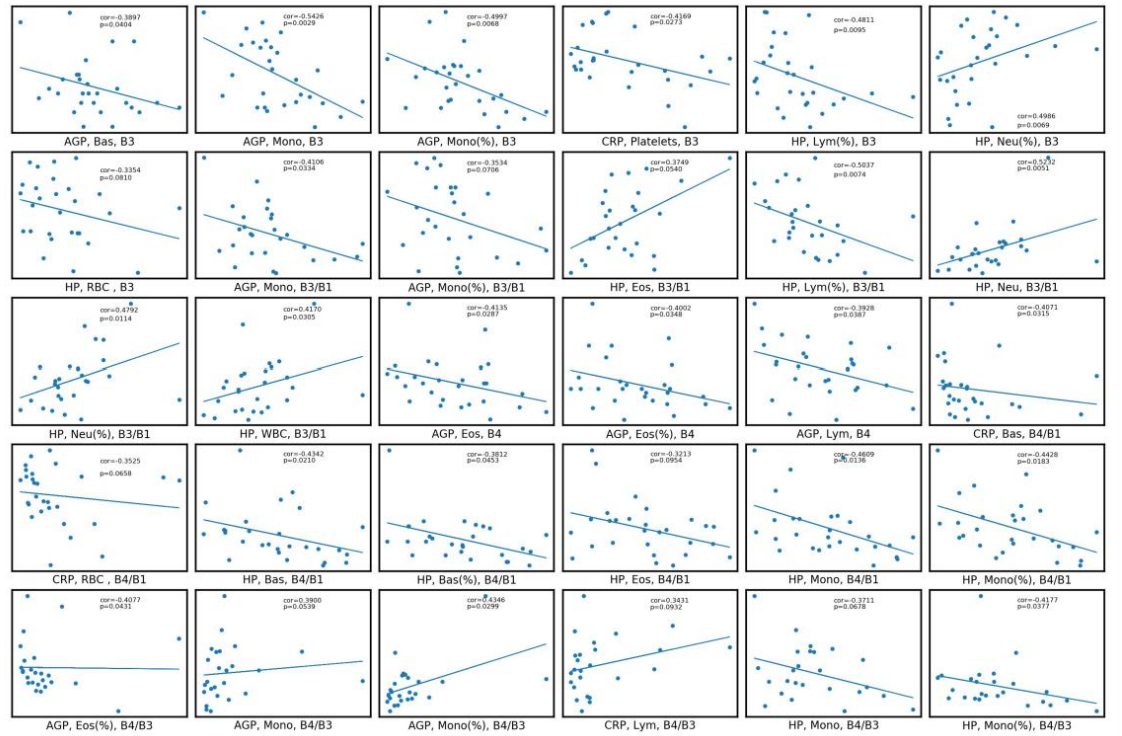

B
AGP, Bas, B3
AGP, Mono, B3
AGP, Mono(%), B3
CRP, Platelets, B3
HP, Lym(%), B3
HP, Neu(%), B3
HP, RBC , B3
AGP, Mono, B3/B1
AGP, Mono(%), B3/B1
HP, Eos, B3/B1
HP, Lym(%), B3/B1
HP, Neu, B3/B1
HP, Neu(%), B3/B1
HP, WBC, B3/B1
AGP, Eos, B4
AGP, Eos(%), B4
AGP, Lym, B4
CRP, Bas, B4/B1
CRP, RBC , B4/B1
HP, Bas, B4/B1
HP, Bas(%), B4/B1
HP, Eos, B4/B1
HP, Mono, B4/B1
HP, Mono(%), B4/B1
AGP, Eos(%), B4/B3
AGP, Mono, B4/B3
AGP, Mono(%), B4/B3
CRP, Lym, B4/B3
HP, Mono, B4/B3
HP, Mono(%), B4/B3

**Figure 2.4** Spearman correlation analysis of CBC data with APP concentration from each bleeding and fold change between blood sampling.

(A) Heatmap of correlation coefficient (denoted by "cor") for all the groups of CBC/APP tested. Abbreviations of cell types are listed in 2.2.3 (B) List of dot plots for the correlation which has a P-value less than 0.1. P-value and cor value are shown inside of each plot (see also Table 2.5).

**Table 2.5** List of correlation coefficient values from Figure 2.4A.

Statistically significant correlations are labeled (*=P<0.05, **=P<0.01).

| Correlation | AGP | HP | CRP | AGP | HP | CRP | AGP | HP | CRP |
|---|---|---|---|---|---|---|---|---|---|
| Samples | B1 | B1 | B1 | B3 | B3 | B3 | B4 | B4 | B4 |
| WBC | 0.15 | 0.15 | 0.04 | -0.21 | 0.07 | 0.03 | -0.26 | 0.20 | 0.28 |
| RBC | 0.19 | -0.12 | -0.20 | -0.12 | -0.34* | -0.28 | 0.01 | -0.24 | -0.25 |
| Platelets | 0.03 | 0.22 | -0.28 | 0.11 | 0.25 | -0.42** | 0.12 | 0.29 | 0.03 |
| Neu | 0.10 | 0.07 | 0.07 | -0.05 | 0.29 | 0.10 | -0.02 | 0.03 | -0.07 |
| Lym | 0.27 | 0.12 | -0.07 | -0.23 | -0.29 | -0.05 | -0.39** | -0.06 | 0.23 |
| Mono | 0.04 | 0.02 | 0.26 | -0.54** | -0.05 | -0.20 | -0.05 | -0.19 | 0.07 |
| Eos | -0.27 | 0.21 | 0.01 | -0.17 | 0.16 | 0.13 | -0.41** | -0.16 | 0.19 |
| Bas | 0.10 | 0.05 | 0.10 | -0.39** | 0.02 | -0.17 | -0.08 | -0.07 | 0.10 |
| Neu(%) | 0.02 | 0.06 | 0.22 | 0.14 | 0.50** | 0.11 | 0.14 | 0.02 | -0.21 |
| Lym(%) | 0.08 | -0.10 | -0.25 | -0.02 | -0.48** | -0.07 | -0.16 | -0.09 | 0.14 |
| Mono(%) | -0.13 | -0.18 | 0.27 | -0.50** | -0.14 | -0.20 | 0.13 | -0.19 | 0.03 |
| Eos(%) | -0.24 | 0.13 | -0.04 | -0.04 | 0.18 | 0.18 | -0.40** | -0.20 | 0.06 |
| Bas(%) | 0.03 | -0.11 | -0.09 | -0.25 | 0.06 | 0.03 | 0.10 | -0.09 | 0.02 |

| Correlation | AGP | HP | CRP | AGP | HP | CRP | AGP | HP | CRP |
|---|---|---|---|---|---|---|---|---|---|
| Samples | B3/B1 | B3/B1 | B3/B1 | B4/B1 | B4/B1 | B4/B1 | B4/B3 | B4/B3 | B4/B3 |
| WBC | -0.04 | 0.42** | 0.18 | -0.21 | -0.23 | -0.15 | -0.04 | 0.03 | 0.30 |
| RBC | -0.05 | -0.18 | -0.25 | 0.15 | -0.32 | -0.35* | 0.02 | -0.02 | 0.07 |
| Platelets | 0.24 | 0.19 | -0.04 | 0.12 | 0.10 | -0.14 | 0.11 | 0.14 | -0.31 |
| Neu | 0.00 | 0.52** | 0.25 | -0.27 | -0.12 | 0.01 | -0.18 | 0.28 | 0.15 |
| Lym | -0.17 | 0.12 | 0.06 | 0.13 | -0.22 | -0.23 | 0.27 | -0.16 | 0.34* |
| Mono | -0.41** | 0.07 | -0.04 | 0.02 | -0.46** | -0.18 | 0.39* | -0.37* | 0.07 |
| Eos | -0.04 | 0.37* | 0.14 | -0.20 | -0.32* | -0.13 | -0.19 | 0.10 | 0.31 |
| Bas | -0.19 | 0.21 | 0.18 | -0.01 | -0.43** | -0.41** | 0.17 | -0.08 | 0.29 |
| Neu(%) | 0.08 | 0.48** | 0.22 | -0.15 | 0.13 | 0.07 | -0.25 | 0.23 | -0.08 |
| Lym(%) | 0.01 | -0.50** | -0.19 | 0.23 | -0.09 | -0.09 | 0.27 | -0.22 | 0.02 |
| Mono(%) | -0.35* | -0.22 | -0.14 | 0.05 | -0.44** | -0.03 | 0.43** | -0.42** | -0.08 |
| Eos(%) | -0.09 | 0.11 | 0.07 | -0.31 | -0.16 | -0.01 | -0.41** | 0.13 | 0.17 |
| Bas(%) | -0.26 | -0.30 | -0.09 | -0.10 | -0.38** | -0.31 | 0.24 | -0.23 | 0.27 |

The first treatment age and treatment frequency are not only important health indicators but also components of farm expenses that should not be neglected. Correlation analysis with the APP concentrations across the time range of three blood samples demonstrated generally complementary trends of correlation between the first treatment age and treatment frequency, particularly after infection was established (Figure 2.5A). At the early stage of the challenge, either absolute concentrations or fold changes of all the three APPs showed a negative or positive correlation with the first treatment age or treatment frequency, respectively. Considering it is more favorable for pig health and farm economy with a higher age of first treatment but a lower treatment frequency, these correlations provide great evidence that the elevation of three APPs in plasma may work as sensitive markers indicating early infections. Intriguingly, a positive correlation was observed for CRP B4/B3 but not HP with the first treatment age, and correspondingly a negative correlation with the treatment frequency, implying CRP concentration at the later stage of challenging may reflect a stronger adaptive immunity. Although the concentration of AGP did not change markedly after challenging, significant negative correlations were observed for both AGP B3 and AGP B3/B1 with the first treatment age (Figure 2.5 B). Combined with the variable level of AGP in B3 among tested pigs (Figure 2.2C), the real-time concentration of AGP early after the challenge may predict future pig health status under pathogen-challenging pressure.

Improving production performance is the key part of the ultimate goal for pig selection and breeding. It would be attractive to look into whether APP concentration could provide some evidence for selecting pig production-related traits. I started with the correlation analysis between the APP concentrations with the pig weight trait here

expressed by ADG. Although there was not a very strong correlation when looking at the ADG around the three bleeding times with the corresponding concentrations of the three APPs, pathogen challenging definitely changed the direction of the correlation trend from global slight positive correlation to negative (Figure 2.6 A). This observation might be explained by the increased APP concentration resulting from acute infections whereas the challenging stress slowed down the weight gain. When the analysis was performed on pigs classified as dead and alive during the experiment, it is intriguing that only the APP concentration or concentration change in pigs that died before the end of the experiment showed a strong correlation with their ADG. Among the three APPs tested, CRP is the only one that showed a significant negative correlation with ADG before challenging in the dead pig group (Figure 2.6 B), implying CRP concentration at the steady state could be a valuable marker when analyzing in combination with pig growth. In addition, the significant positive correlation between HP B3/B1 values and dead pig ADGs may point to aberrant host response to the disease challenge, which then caused a sharp loss of growth rate with disease going on as demonstrated by the reverse correlation of the ADG with HP B4 or HP B4/B1.

Another critical aspect to evaluate pig production performance and health status is pig feeding. I then analyzed the correlation between APP data and feed-related traits including feed intake (FI), average feed intake (AFI), and feed conversion ratio (FCR) (Figure 2.7). Similar to the trend with pig weight traits, the APP concentration did not provide strong hints to assess pig appetite (FI or AFI) when analyzing all the tested pigs together. A subtle trend for the relationship between FI and APP concentrations was an overall negative correlation as shown in Figure 2.7 B. Of note, more apparent correlations

were observed between FCR and APP data in the death pig group than in the live and total, which is reminiscent of the results in Figure 2.6 when analyzing weight-related traits. The three plots in Figure 2.7 B that show a significant correlation between APP data and FCR_total might be resulting from the more number of live pigs in the experiment which maintained a relatively stable FCR value. Unexpectedly, given that a low FCR and a higher ADG would be more favorable for a farm regarding profitability, the general correlation trends with APP data between FCR and ADG in dead pigs are not in contrast to each other. This could be explained by the different recording times for pig weight and FI with the latter one starting late in grower-finisher pigs whose weight might have more fluctuation compared to that across the whole experiment time range. Nevertheless, it is apparent that pigs that stayed alive across the experiment have relatively stable FCR compared to the dead individuals, highlighting the importance of monitoring FI and weight in the pig industry to assess pig health status.

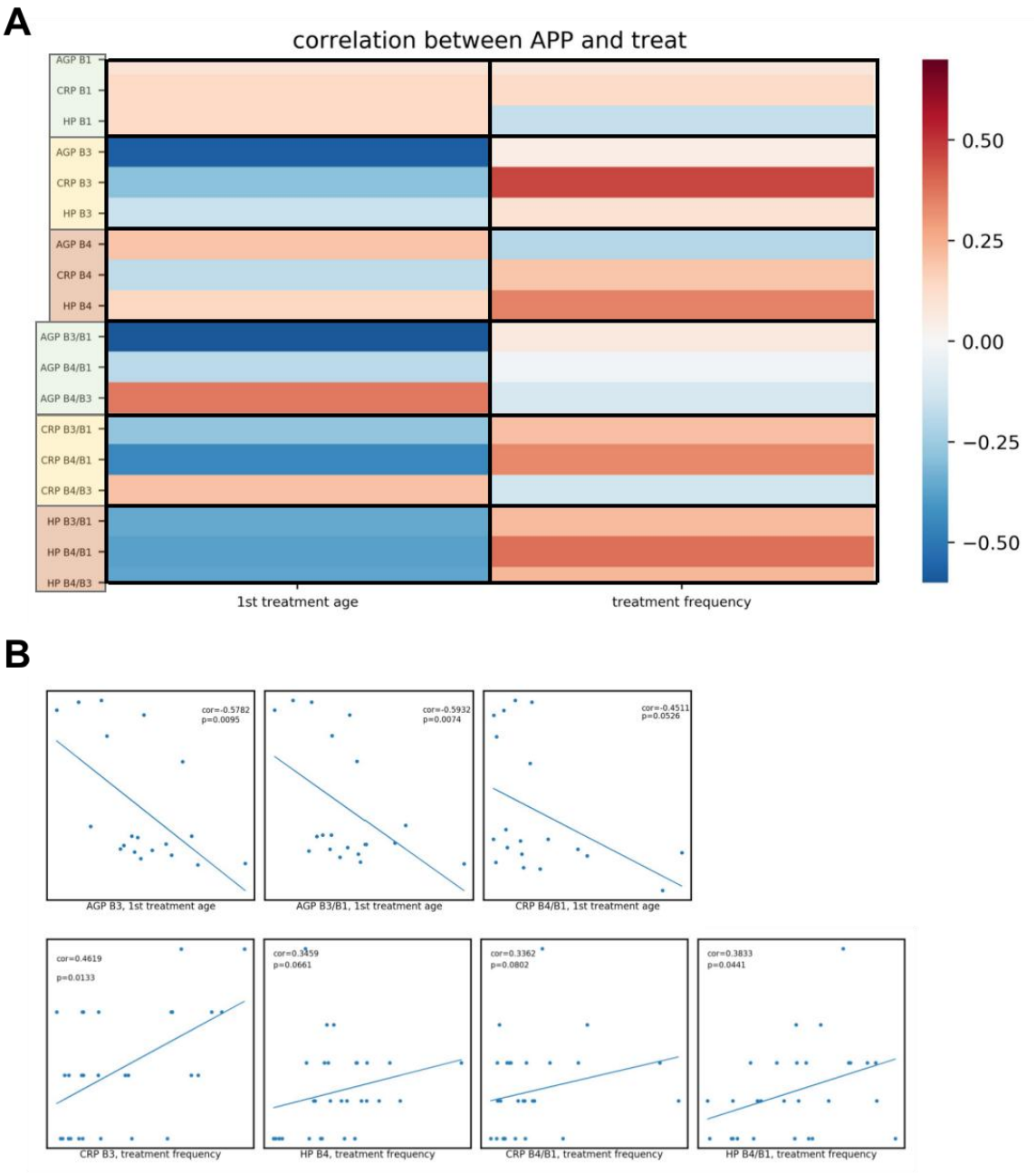



**Figure 2.5** Spearman correlation analysis of treatment records with APP concentration from each bleeding and fold change between blood sampling.

(A) Heatmap of the correlation coefficient for the first treatment age and treatment frequency with the APP data. Note that only pigs who received clinical treatment are included in the correlation analysis for the first treatment age. (B) List of dot plots for the correlation which has a P-value less than 0.1. P-value and cor value are shown inside of each plot.

**Table 2.6** List of correlation coefficient values from Figure 2.5A.

Statistically significant correlations are labeled (*=P<0.05, **=P<0.01).

| Correlation | 1st treatment age | treatment frequency |
|---|---|---|
| **AGP B1** | 0.10 | 0.08 |
| **CRP B1** | 0.14 | 0.13 |
| **HP B1** | 0.14 | -0.16 |
| **AGP B3** | -0.58** | 0.05 |
| **CRP B3** | -0.29 | 0.46** |
| **HP B3** | -0.15 | 0.11 |
| **AGP B4** | 0.20 | -0.19 |
| **CRP B4** | -0.18 | 0.19 |
| **HP B4** | 0.15 | 0.35* |
| **AGP B3/B1** | -0.59** | 0.07 |
| **AGP B4/B1** | -0.19 | -0.02 |
| **AGP B4/B3** | 0.37 | -0.12 |
| **CRP B3/B1** | -0.28 | 0.21 |
| **CRP B4/B1** | -0.45* | 0.34* |
| **CRP B4/B3** | 0.21 | -0.13 |
| **HP B3/B1** | -0.36 | 0.22 |
| **HP B4/B1** | -0.38 | 0.38** |
| **HP B4/B3** | -0.37 | 0.23 |

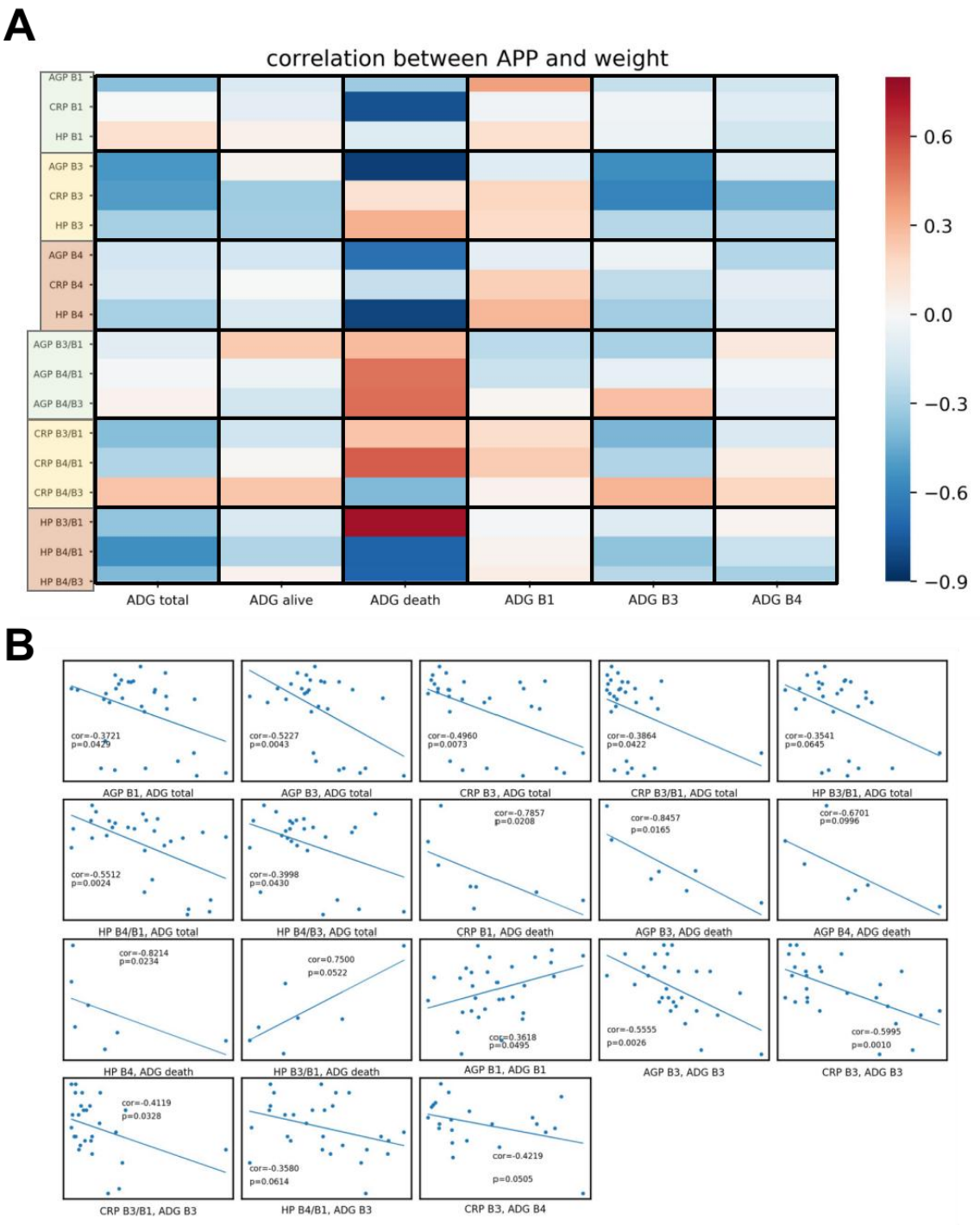



**Figure 2.6** Spearman correlation analysis of pig weight records with APP concentration from each bleeding and fold change between blood sampling.

(A) Heatmap of the correlation coefficient for the pig weight traits represented by the average daily gain (ADG) with the APP data. ADG values were calculated using the equation: (Weight on day B – (minus) Weight on day A)/(days from A to B). ADG total indicates the ADG of all the tested pigs from the start of the experiment to the slaughter

or death age, and the ADG alive and death separate the pigs based on whether they stayed alive until the slaughter age. ADG B1, ADG B3, and ADG B4 means the ADG around the three bleeding times (Time range for ADG B1, B3, and B4 are day 10 to 20, day 37 to 58, and day 58-79, respectively, from the start of the experiment). (B) List of dot plots for the correlation which has a P-value less than 0.1. P-value and cor value are shown inside of each plot.

**Table 2.7** List of correlation coefficient values from Figure 2.6A.

Statistically significant correlations are labeled (*=P<0.05, **=P<0.01).

| Correlation | ADG total | ADG alive | ADG death | ADG B1 | ADG B3 | ADG B4 |
|---|---|---|---|---|---|---|
| **AGP B1** | -0.37** | -0.14 | -0.33 | 0.36** | -0.21 | -0.17 |
| **CRP B1** | -0.01 | -0.09 | -0.79** | -0.04 | -0.04 | -0.11 |
| **HP B1** | 0.14 | 0.05 | -0.12 | 0.15 | -0.05 | -0.18 |
| **AGP B3** | -0.52** | 0.03 | -0.85** | -0.10 | -0.56** | -0.14 |
| **CRP B3** | -0.50** | -0.32 | 0.14 | 0.19 | -0.6** | -0.42* |
| **HP B3** | -0.30 | -0.31 | 0.32 | 0.18 | -0.25 | -0.24 |
| **AGP B4** | -0.17 | -0.16 | -0.67* | -0.09 | -0.05 | -0.27 |
| **CRP B4** | -0.13 | -0.01 | -0.21 | 0.21 | -0.23 | -0.10 |
| **HP B4** | -0.30 | -0.14 | -0.82** | 0.30 | -0.31 | -0.14 |
| **AGP B3/B1** | -0.10 | 0.23 | 0.29 | -0.24 | -0.29 | 0.10 |
| **AGP B4/B1** | -0.02 | -0.04 | 0.49 | -0.20 | -0.07 | -0.03 |
| **AGP B4/B3** | 0.04 | -0.17 | 0.50 | 0.02 | 0.28 | -0.09 |
| **CRP B3/B1** | -0.39** | -0.18 | 0.25 | 0.16 | -0.41** | -0.13 |
| **CRP B4/B1** | -0.27 | 0.01 | 0.54 | 0.22 | -0.27 | 0.06 |
| **CRP B4/B3** | 0.26 | 0.24 | -0.40 | 0.04 | 0.31 | 0.19 |
| **HP B3/B1** | -0.35* | -0.14 | 0.75* | -0.02 | -0.11 | 0.03 |
| **HP B4/B1** | -0.55** | -0.27 | -0.72 | 0.03 | -0.36* | -0.20 |
| **HP B4/B3** | -0.40** | 0.04 | -0.72 | 0.07 | -0.25 | -0.30 |

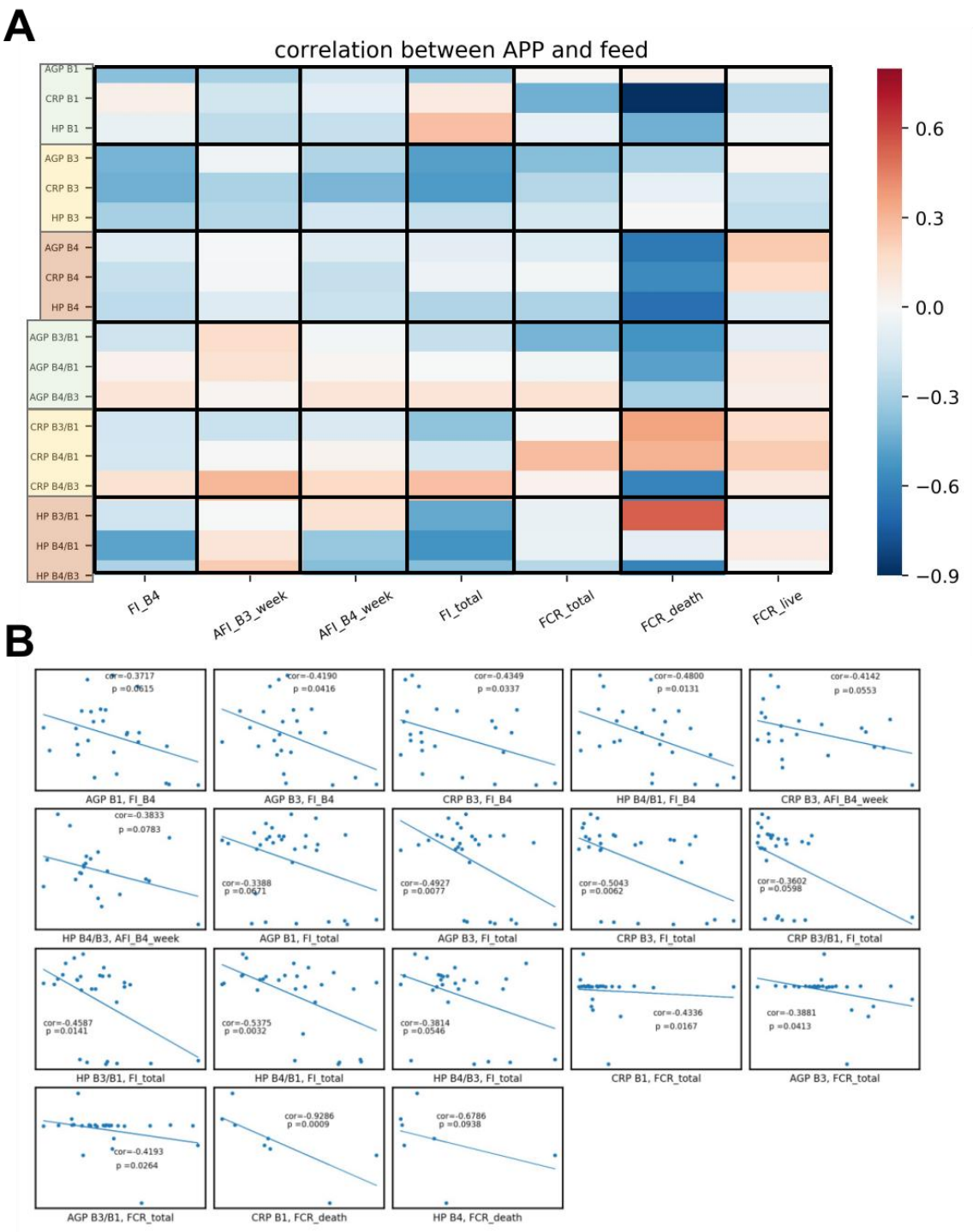



**Figure 2.7** Spearman correlation analysis of pig feed records with APP concentration from each bleeding and fold change between blood sampling.

(A) Heatmap of the correlation coefficient for the pig feed traits represented by the FI, AFI, and FCR with the APP data. Note that all the feed records are only available for grower-finisher pigs. FI_B4 indicates the FI for the day of bleeding 4. FI_total indicates

the total FI before death or slaughter. AFI values were averaged FI within a specific time range. AFI_B3_week and AFI_B4_week mean the AFI for the following week of bleeding 3 and 4, respectively. FCR values were calculated using the equation: (Total FI from day A to day B)/(weight of day B - weight of day A). FCR_total indicates the FCR of all the tested pigs across the start of the experiment to the slaughter or death age, and the FCR alive and death separate the pigs based on whether they stayed alive until the slaughter age. (B) List of dot plots for the correlation which has a P-value less than 0.1. P-value and cor value are shown inside of each plot.

**Table 2.8** List of correlation coefficient values from Figure 2.7A.

Statistically significant correlations are labeled (*=P<0.05, **=P<0.01).

| Correlation | FI_B4 | AFI_B3_week | AFI_B4_week | FI_total | FCR_total | FCR_death | FCR_live |
|---|---|---|---|---|---|---|---|
| **AGP B1** | -0.37* | -0.29 | -0.16 | -0.34* | 0.01 | 0.05 | 0.02 |
| **CRP B1** | 0.06 | -0.17 | -0.08 | 0.08 | -0.43** | -0.93** | -0.24 |
| **HP B1** | -0.07 | -0.23 | -0.21 | 0.28 | -0.08 | -0.43 | -0.04 |
| **AGP B3** | -0.42** | -0.04 | -0.27 | -0.49** | -0.39** | -0.29 | 0.03 |
| **CRP B3** | -0.43** | -0.29 | -0.41* | -0.5** | -0.26 | -0.07 | -0.19 |
| **HP B3** | -0.3 | -0.26 | -0.17 | -0.20 | -0.16 | 0.00 | -0.22 |
| **AGP B4** | -0.11 | -0.02 | -0.11 | -0.09 | -0.12 | -0.64 | 0.23 |
| **CRP B4** | -0.21 | -0.02 | -0.21 | -0.06 | -0.02 | -0.57 | 0.17 |
| **HP B4** | -0.23 | -0.11 | -0.20 | -0.27 | -0.28 | -0.68* | -0.13 |
| **AGP B3/B1** | -0.19 | 0.16 | -0.03 | -0.22 | -0.42** | -0.54 | -0.09 |
| **AGP B4/B1** | 0.04 | 0.13 | 0.02 | -0.01 | -0.02 | -0.49 | 0.08 |
| **AGP B4/B3** | 0.11 | 0.03 | 0.12 | 0.12 | 0.13 | -0.3 | 0.06 |
| **CRP B3/B1** | -0.17 | -0.20 | -0.14 | -0.36* | 0.01 | 0.36 | 0.17 |
| **CRP B4/B1** | -0.16 | 0.01 | 0.04 | -0.16 | 0.28 | 0.31 | 0.23 |
| **CRP B4/B3** | 0.12 | 0.30 | 0.18 | 0.28 | 0.04 | -0.6 | 0.10 |
| **HP B3/B1** | -0.18 | -0.01 | 0.13 | -0.46** | -0.07 | 0.54 | -0.07 |
| **HP B4/B1** | -0.48** | 0.11 | -0.34 | -0.54** | -0.06 | -0.09 | 0.08 |
| **HP B4/B3** | -0.28 | 0.22 | -0.38* | -0.38* | -0.25 | -0.60 | 0.00 |

## 2.4 Discussion

In this chapter, I measured the plasma concentration of AGP, HP, and CRP from 31 pigs which were further grouped into Group 1-resistant low production, Group 2-resistant high production, Group 3-quick resolving, Group 4-susceptible, Group 5-resilience, and Group 6-late death based on their health and production performance. In our natural disease challenge model, a sharp increase in concentrations of HP and CRP was detected while a change in concentrations of AGP was negligible in response to the stress of complex pathogens. The resultant data is consistent with previous reports on the trend of change in those three APPs in the context of other APRs. However, except for the concentration of AGP in the Resilient group was significantly lower than others before challenge, CRP and HP concentrations at steady state was inapplicable to identify resilience pigs. This is unexpected as AGP was relatively inert to pathogen challenge in our system and originally considered as a negative control in my experiment design. Even though the plasma level is insensitive to pathogens, AGP was believed to play an immune-modulatory role, especially in the immunosuppression activity to T cells (Chiu et al., 1977) . This inhibitory role was presumably through the interaction of AGP with membrane surface proteins of T cells that are associated with cell proliferation pathways (Cheresh et al., 1984; Pos et al., 1990; J. Wu et al., 1999). In addition, AGP might suppress lymphocyte proliferation in response to LPS, suggesting its inhibitory role in the immune system during bacterial infection. The reported inhibitory role of AGP to lymphocytes is consistent with our findings on the lowest AGP concentration in the resilient group, which may result from faster lymphocyte proliferation in response to pathogen challenge to better control infections. The low concentration of AGP was also

able to stimulate mononuclear cell proliferation (Cheresh et al., 1984), implying that resilience pigs may have stronger pathogen sensing and antigen presentation. In the context of an inhibitory activity to immune response, it remains unclear why the AGP concentration in the susceptible pigs was only slightly higher than the resilience group and comparable with that of Group 2 resistant pigs. The resilience pigs may arguably have lower AGP plasma concentration at the steady state to maintain a highly sensitive immunity, but the lower AGP alone might not be reliable to predict pig resilience. In addition, a recent study demonstrated the role of AGP in potentiation of TLR-2 but inhibition of the TLR-4 signaling pathway (activated by differential gram-negative bacterial endotoxins) (Sumanth et al., 2019). Another recent study on bovine AGP uncovered its antibacterial activity to disrupt the biofilm of *Staphylococcus aureus* in milk. It is however unclear whether AGP functions analogously in plasma (Meroni et al., 2019). All the progress not only facilitates our understanding of AGP activity in response to pathogens but also complicates the role of AGP in a context-dependent manner. With regards to the relationship of AGP with animal production other than its immuno-modulatory functions, research by Sun et al. elaborated on the mechanism of AGP in regulating energy intake (Sun et al., 2016). AGP was able to bind the hypothalamic leptin receptor to induce satiety signals in mice. However, compelling results reported more recently showed that AGP lacked activity to influence feed intake in ruminants (Gregg et al., 2019; McGuckin et al., 2020). It seems that AGP may play species-specific roles in influencing animal satiety signals. However, consistent with the mice study mentioned above, a study published last year demonstrated a negative correlation between plasma AGP concentration and feed intake in postpartum dairy cows (Brown et al., 2021). Thus,

AGP is possibly playing a conserved role in influencing animal feeding in certain contexts. In spite of the distinction from an infection-driven APR, these findings are consistent with our data shown in Figure 2.6B that AGP concentrations before and after the disease challenge both exhibited a strong negative correlation with feed intake.

In our model, the dramatic change in HP and CRP concentrations after the challenge suggested that those two APPs can be applied to evaluate pig health status, and in turn to select and exclude unhealthy pigs before integration into an uninfected herd. A previous study using a co-infection model with SIV and *Pasteurella multocida* showed a ~ 3-fold increase in HP and a ~ 10-fold increase in CRP 3 days and 2 days post-infection, respectively (Pomorska-Mól et al., 2013). Data from the current study demonstrated a sharper change in up-regulation of HP compared to that of CRP, which is reverse to the SIV and *Pasteurella multocida* co-infection model. The co-infection model also revealed a quick increase in APP concentration after experimental inoculation for two to three days and a gradual reversion back to the original level within 10 days. The APP concentration shown in that study, however, did not maintain at the high level for several weeks as shown in this chapter (Figure 2.2). These different response patterns were presumably due to the complexity of infection models. The co-infection model mentioned above was an experimental intranasal inoculation of the pathogens which was artificial, while the natural infection method in this thesis more closely represents practical infection kinetics. In that co-infection model, the positive correlation was found between serum concentration of HP and lung scores. It would be significant to include the clinical score in our model in the future analysis to correlate APP with disease severity. APP concentrations could be beneficially used as parameters to monitor ongoing infections

and disease severity. Of note, a recent study, using attenuated Aujeszky vaccine in growing pigs, measured serum HP levels in a challenging model and the resulting data demonstrated that resilient pigs had a minor increment of HP (Laghouaouta et al., 2021). The authors proposed the increment of HP in response to challenges as a novel indicator of pig resilience. Importantly, combined with their other recent works, HP response in their system was found to be genetically controlled and used to identify associated genomic regions with substantial variability in the population (Laghouaouta et al., 2022). This finding is in concordance with my data shown in Figure 2.3 that HP B3/B1 was the lowest in the resilience group. Since in an independent work, another group also proposed the potential of HP as a biomarker for pig resilience, HP appears to be a promising potential biomarker in improving pig resilience through selective breeding for the pigs with relatively small change of HP plasma concentration early after infection. It would also be intriguing to adapt their methodology to quantify resilience in our model which focuses on infectious disease to validate my findings and explore additional indicators for pig resilience.

One notable finding in this Chapter is the different kinetics of APPs in our system. In my hands, CRP revealed a stronger secondary increase post-disease challenge compared to HP (Figure 2.3). This is even more noteworthy since the resilience group exhibited the strongest up-regulation of CRP from the early to the later stage of the disease challenge. CRP proteins have been known to exist at least two distinct conformations, the native pentamers and the monomers. The two CRP forms may bind to receptors and lipid rafts by two distinct modes. The two forms of CRP may thus perform their biological function differently in the pro-inflammatory or anti-inflammatory

contexts (Sproston & Ashworth, 2018; Y. Wu et al., 2015). Even though many studies support CRPs as a biomarker for acute-inflammation (Cardoso et al., 2015; Hu et al., 2017; Memar et al., 2019), CRPs have also been shown to up-regulate chronic inflammation (Luan & Yao, 2018; Pepys & Hirschfield, 2003). Thus, there might exist two explanations for the notable secondary up-regulation, especially for the resilience group, found in the current study: 1. Resilient pigs were initially resistant to some specific pathogens but become infected by those pathogens at a later stage and stimulated the secondary CRP increase; 2. Resilient pigs were initially invaded by some pathogens but were not able to clear them. The Resilient pigs, due to the residual pathogens, then developed chronic inflammation at a later stage. To investigate these assumptions, it would be informative to keep tracking the particular pathogen infections in pigs during the challenge experiment, and the resultant data will be beneficial to understand how resilient pigs differentially interact with pathogens compared to resistant or susceptible pigs. During the pandemic of COVID-19, several attempts have been made to study whether CRP blood level was associated with SARS-Cov-2 response. CRP is known to be synthesized by hepatocytes stimulated by IL-6 which is the main driver of cytokine storm in COVID-19 patients (Zhang et al., 2020). Despite differential sampling time and cohort, those studies independently reported a positive correlation of a high level of CRP with the severity and mortality of COVID-19. These studies collectively suggested that CRP blood level could be applicable to predict the disease severity and guide therapeutic options (Ahnach et al., 2020; Potempa et al., 2020). Interestingly, one of these studies found that the serum CRP determined on admission was positively correlated with neutrophil count but negatively correlated with lymphocyte count. This data is consistent

with my data shown in Figure 2.4 for the early CRP concentration after the disease challenge. The work by myself and others provide insight into the potential use of CRP in the pig industry as an early marker to predict disease development to prevent the onset of severe pathology and specify healthcare strategies to improve animal welfare.

In summary, APPs have various biological functions and may play crucial roles in pathogenesis or immune response to infectious diseases. The results in this Chapter shed light on the potential of using plasma concentrations of specific APPs as parameters to recognize resilient pigs. The APP levels or kinetics are closely correlated with health and production traits to assess pig resilience. APPs can be readily and quickly measured, and definitely have the potential as biomarkers to indicate and even predict pig health status in the pig industry. However, pig resilience is a multi-factor-driven phenotype; it may be difficult and inaccurate to use the readout from a single marker to predict or evaluate pig resilience. In this regard, it would be necessary to use a larger cohort of pigs (this also applies of course to my transcriptome and metabolome studies) and measure multiple pig APPs including the ones not determined in this study such as Serum amyloid A (SAA) and Pig major acute-phase protein (Pig-MAP) in our natural disease challenge model, aiming to explore more candidates that are informative for selecting and breeding resilient pigs. Meanwhile, the APP data in conjunction with other types of data including genetic, epigenetic, and metabolites could provide a more powerful tool to precisely define and predict pig resilience.

## 2.5 Summary

Using the natural disease challenge model, I investigated the feasibility of APPs serving as biomarkers to assess or predict pig resilience phenotype. I analyzed the plasma concentration of HP, CRP, and AGP in 6 groups of pigs (Group 1-resistant low production, Group 2-resistant high production, Group 3-quick resolving, Group 4-susceptible, Group 5-resilience, Group 6-late death) that were classified based on their health and production performance from our experiment model. Blood samples were collected at three different stages (B1, B3, and B4) with the first sample two weeks before challenging, the second and third were sampled two and five weeks post challenging, respectively. Plasma samples were isolated and APP concentrations were determined by ELISA. Consistent with the previous literature mentioned in the background section, HP and CRP had generally quick and sharp responses to the experimental disease challenge while AGP only showed slight fluctuations across the whole experiment. When comparing the APP concentrations among the six groups of pigs in steady state, the late-death group was found to have the highest AGP concentration whereas the resilience group had the lowest. However, there was no significant difference for HP or CRP before the disease challenge. For the blood samples post challenging, there was no significant difference for concentrations of the AGP and CRP among the six groups, only HP showed a significant difference between the quick resolving and late death group in the third blood samples. Similarly in the calculations of APP fold changes, the only significant difference observed was between the quick resolving and the late-death group for the HP change from B1 to B4. As shown in Figure 2.3, even though each fold change of CRP concentration could not significantly

differentiate the resilience group from others, there was a unique trend for the resilience group that started with a subtle increase in the early infection but a sharp increase during the later stage. Generating a larger cohort in the future would be helpful to validate this unique kinetic of CRP in resilience pigs.

As it was difficult to predict resilience using concentration or concentration change of a specific APP using a small cohort, also considering resilience is a multi-factors-dependent trait, I did correlation analysis trying to explore some evidence using APP data to assess pig health and production performance which are key factors for assessing resilience. The strong positive correlation between HP with Neu and WBC during early infection implies HP concentration is coupled with a quick mobilization of the innate immune system while the positive correlation between the increase of CRP from early to late stage of infection and Lym implies that CRP is more actively induced along with the activation of adaptive immunity. This could suggest HP may work as more sensitive indicator of innate immunity activation while CRP may more sensitively indicate adaptive immunity. For the correlation analysis of APP data with the record of treatment, weight, and feed intake, the results are highlighted and listed in Table 2.5. Although APP data could not significantly predict pig health and production performance before disease challenge, it is promising that both HP and CRP concentrations in response to early infection are significantly correlated with multiple health and production traits.

To conclude the findings in this chapter, APPs have the potential to predict or assess pig health and production status before or during disease challenge. Combinatory analysis of APP plasma concentration in our natural disease model may provide more constructive information to identify and select pigs that have resilience traits.

**Table 2.9** Summary of significant correlations found between APP data and trait pairs (production/health) in this study.

Positive and negative correlations that have a P-value<0.1 are labeled in the table by "+" or "-", respectively. Corresponding correlation coefficient values are listed in Figure 2.5-2.7.

| Favored Correlation | +/+ or -/- | +/- or -/+ | +/+ or -/- | +/- or -/+ | +/- or -/+ | +/+ or -/- |
|---|---|---|---|---|---|---|
| Traitss<br>APPs | ADG/<br>1st treatment age | ADG/<br>treatment frequency | FI/<br>1st treatment age | FI/<br>treatment frequency | FCR/<br>1st treatment age | FCR/<br>treatment frequency |
| AGP B3 | **-/-** | | **-/-** | | **-/+** | |
| CRP B3 | | **-/+** | | **-/+** | | |
| HP B4/B1 | | **-/+** | | **-/+** | | |

# CHAPTER 3: Investigation of the whole blood transcriptome of healthy pigs to screen for biomarkers associated with disease resilience

## 3.1 Background

In Chapter 2, I found that the concentrations of specific APPs in plasma were able to provide valuable information on pig health status and were potentially applicable to predict disease resilience in pigs. Interestingly, APP concentrations were also correlated with CBC data in our challenge model: Haptoglobin (HP) was positively correlated with white blood cells (WBC) early after infection while C-Reactive Protein (CRP) was positively correlated with lymphocytes (Lym) during a later stage. Although I tested several specific APPs for their potential to predict disease resilience, an untargeted approach was needed to more broadly search for early indicators of disease resilience and better understand how it is established. Given that the peripheral blood cells are crucial in defending against pathogen infection and are common parameters used in disease diagnosis to reflect host immune response, I hypothesized that the pig response to disease challenge can be represented and even predicted by the variations found in the blood cells during a healthy state. Thus, I undertook the task of extracting molecular information, particularly the transcriptome data, from pig peripheral blood before disease challenge. The underlying molecular information, in conjunction with our natural disease challenge model, were exploited to examine whether a specific, or cluster of, gene expressions can be used to predict disease resilience of pigs.

Transcriptomic approaches have been used to characterize the host immune response to some specific pathogen infections in pigs. By comparing the gene expression between healthy and sick pigs, researchers were able to identify key genes that are indicators or mediators of pathogenesis of the diseases. Examples include but are not limited to the infection of *Haemophilus parasuis* (Melnikow et al., 2005), *Salmonella enterica* (Wang et al., 2007), and PRRSV (Kommadath et al., 2017). Some studies have investigated the variations in porcine gene expression in pigs with differential responses to infection. Huang et al. identified distinct regulatory pathways in pigs, classified as 'high' or 'low' shedders of *Salmonella* (T.-H. Huang et al., 2011). Different gene expression profiles have also been observed in the lung of pigs which showed different susceptibilities to Glässer's disease caused by *Haemophilus parasuis* (Wilkinson et al., 2010). Gene expression profiles in the lung, lymph node, and blood of pigs were divided into more or less susceptible to PRRSV infection (Arceo et al., 2012). Circulating immune cells in peripheral blood are exposed to sites of injury, infectious agents, and stress hormones, and play critical roles in immune surveillance and clearance of invading pathogens. Mobilization and activation of immune cells in blood in responding to infectious diseases can indeed induce differential gene expression. A study on *Salmonella* investigated the peripheral blood transcriptome profiles of pigs before and after pathogen inoculation, and identified co-expressed gene modules associated with fecal *Salmonella* shedding by comparing low and persistent shedders. Intriguingly, the differential gene expression with these modules were also significant before the *Salmonella* challenge (Kommadath et al., 2014). It should be noted that some gene modules associated with *Salmonella* shedding are related to innate immunity, suggesting the expression level of those genes in healthy

state are crucial to determine the host response to pathogen post infection. This study provided the rationale that gene expression data before disease challenge could possibly be used to predict pig response to diseases. There are also some human studies which exploited the transcriptome of blood cells to identify gene expression markers to predict development of infectious diseases, e.g. discrimination between pathogen species that produce similar disease symptoms (Ramilo et al., 2007), latent or active disease states (Berry et al., 2010), or individuals responsive or unresponsive to anti-microbial treatment (Bloom et al., 2012).

Genetic variations influence immune-related traits and may further impact disease status and animal production traits (Knap & Bishop, 2000) and many immune-related traits have moderate to high heritability (Flori et al., 2011). Once marker genes are identified to potentially predict pig disease response, it is also important to correlate differential RNA expression to causative DNA polymorphisms for breeding purposes. This is exemplified by much evidence suggesting that immune responses to infectious diseases are associated with genetic variations in livestock. Boddicker identified a major QTL in the pig genome that was responsible for variation in susceptibility to PRRSV infection (Boddicker et al., 2012). Subsequently, a mutation in the *GBP5* gene was found to be associated with this phenotype (Schroyen et al., 2016). Another example that combines transcriptomic and genetic variation study reported a QTL on the pig chromosome 12 identified for PCV2 infection and a mutation in the *SYNGR2* gene was found to impact viral replication after infection (Walker et al., 2018). Therefore, with this background, I hypothesized that profiling the blood transcriptome in pigs before disease challenge would be an efficient way to identify marker genes to predict disease resilience

in pigs. These efforts could not only contribute to practical breeding methods by coupling with genetic variations but also facilitate the understanding of the molecular mechanism of this complex phenotype.

In this chapter, I aimed to test whether transcriptomic profiles in healthy pigs can be used to identify marker gene expression patterns for predicting pig response under disease pressure. To this end, blood transcriptome in unchallenged animals was profiled followed by tracking their post-challenge outcomes to disease using our natural disease challenge model. This could provide insights into the potential of using blood transcriptomes as biomarkers to predict disease resilience. I initially sequenced the whole blood transcriptome of 58 pigs before disease challenge and then grouped those pigs into Early_dead, Susceptible, Resistant, and Resilient based on their subsequent health and productivity records in response to the diseases. To validate the finding from the initial 58 pigs and further investigate how gene expression changes in different groups of pigs from before and after exposure to pathogens, another cohort of 209 pigs were subsequently included but with three blood collections from either before or after disease challenge to perform transcriptome analysis.

## 3.2 Methods

### 3.2.1 Grouping of pigs for transcriptome analysis

The pig grouping method for whole blood transcriptome analysis from the natural disease model was elaborated in Chapter 2 and was based on pre-slaughter weight, number of medical treatments, and mortality to assess pig health status and productivity. Pre-slaughter weight was measured before pigs were sent to a slaughterhouse, and it was

positively related to the economic value of pigs. For blood samples used for transcriptome analysis, as illustrated in Table 2.1 from the Chapter 2, “Blood 1” was collected from the initial cohort of 58 animals (cohort 1) from Cycle 2 (batch 8, 9, 10, and 12), while the second cohort (cohort 2, Table 3.2) consisted of samples from “Blood 1, 3, and 4” of Cycle 7 (batches 46-50). Pigs were grouped into “Resistant”, “Resilient”, “Early_dead”, and “Susceptible” using their records introduced above after the natural disease challenge (Table 3.1 and 3.2). Resistant animals weighed 110-145 kg at the age of 184 days, while a cutoff of 120 kg was used to differentiate sick but alive pigs which were split into “Resilient” and “Susceptible” (Figure 3.1). The cutoff of 120 kg was chosen based on two facts : 1. The market pigs in Canada weight around 280 pound (~127 kg) at an age of approximately 5 to 6 months (The British Columbia Society for the Prevention of Cruelty to Animals (BC SPCA) 2023); 2. The average pre-slaughter weights of pigs in Cycle 2 and Cycle 7 are 122.9 kg and 119.6 kg, respectively, and the cutoff of 120 kg nearly divided 50% of pigs above and below the average weight (Figure 3.1A). Resistant pigs did not receive any treatment while pigs in the other groups received various numbers of treatments (Table 3.1, 3.2 and Figure 3.1B). More detailed records for production performance and medication of the grouped pigs are listed in Table 3.1 and 3.2. For the 58 pigs in cohort 1, total RNA was extracted and sequenced from whole blood collected before the disease challenge (Blood 1) as shown in Table 2.1 of Chapter 2. Whereas, for cohort 2 there were 209 individual pigs with different numbers across the three sampling time points: Blood 1 (193 pigs), Blood 3 (176 pigs), and Blood 4 (131 pigs) with a total of 500 samples.

**Table 3.1** Grouping of 58 pigs (cohort 1) for transcriptome analysis according to production index and medical treatment information.

| Groups | Resistant | Resilient | Early_dead | Susceptible |
|---|---|---|---|---|
| Cycle | 2 | 2 | 2 | 2 |
| Batch | 8,9,10,12 | 8,9,10,12 | 8,9,10,12 | 9,10,12 |
| MTN(#) | 0 | 2.79 | 2.5 | 3.1 |
| RTN(#) | 0 | 1-5 | 0-7 | 1-8 |
| MPSW(kg) | 128.44 | 136.89 | NA | 102.55 |
| RPSW(kg) | 107.5-142 | 124.5-148 | NA | 85.5-112.5 |
| pig# | 16 | 14 | 18 | 10 |
| MSA(day) | 184.25 | 183.79 | NA | 187.8 |
| RSA(day) | 162-205 | 181-185 | NA | 181-205 |
| MDA(day) | NA | NA | 115.28 | NA |
| RDA(day) | NA | NA | 51-194 | NA |
| MDW(kg) | NA | NA | 55.22 | NA |
| RDW(kg) | NA | NA | 15-110 | NA |

| Abbr. | |
|---|---|
| MTN | Mean treatment# |
| RTN | Range of treatment# |
| MPSW | Mean pre-slaughter weight |
| RPSW | Range of pre-slaughter weight |
| MSA | Mean slaughter age |
| RSA | Range of slaughter age |
| MDA | Mean death age |
| RDA | Range of death age |
| MDW | Mean death weight |
| RDW | Range of death weight |

**Table 3.2** Grouping of 209 pigs (cohort 2) for transcriptome analysis according to production index and medical treatment information. Detailed numbers of pigs in each group from each blood collection was listed below.

| Groups | Resistant | Resilient | Early_dead | Susceptible |
|---|---|---|---|---|
| Cycle | 7 | 7 | 7 | 7 |
| Batch | 46,47,48,49,50 | 46,47,48,49,50 | 46,47,48,49,50 | 46,47,48,49,50 |
| MTN(#) | 0 | 1.15 | 1.36 | 1.81 |
| RTN(#) | 0 | 1-3 | 0-4 | 1-4 |
| MPSW(kg) | 122.71 | 129.37 | NA | 100.70 |
| RPSW(kg) | 94.5-144 | 120-147 | NA | 55.5-119.5 |
| pig# | 76 | 26 | 80 | 27 |
| MSA(day) | 177.43 | 185.88 | NA | 185.07 |
| RSA(day) | 148-190 | 169-203 | NA | 169-203 |
| MDA(day) | NA | NA | 79.82 | NA |
| RDA(day) | NA | NA | 28-179 | NA |
| MDW(kg) | NA | NA | 23.38 | NA |
| RDW(kg) | NA | NA | 6-115 | NA |

| | Blood1 | Blood3 | Blood4 | Sum |
|---|---|---|---|---|
| Resistant | 70 | 70 | 67 | 207 |
| Resilient | 24 | 25 | 22 | 71 |
| Early_dead | 75 | 56 | 20 | 151 |
| Susceptible | 24 | 25 | 22 | 71 |
| sum | 193 | 176 | 131 | 500 |

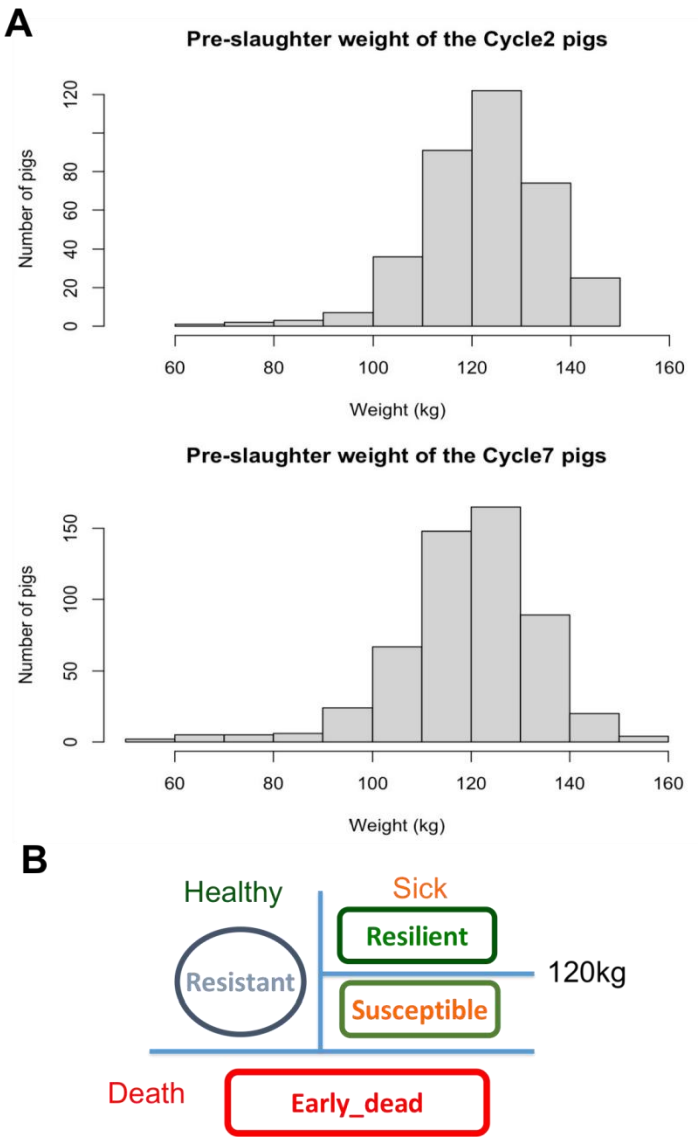



**Figure 3.1 Phenotypic grouping of pigs selected from Cycle 2 and Cycle 7 used for transcriptome analysis.**

(A) Pig pre-slaughter weight distribution in Cycle 2 and Cycle 7. (B) Schematic illustrating the 4 phenotypic groups of pigs.

### 3.2.2 Blood collection

Blood samples (3 mL per pig) were collected in Tempus Blood RNA tubes (Applied Biosystems, Foster City, USA) following the standard procedures for drawing blood from pigs into blood collection tubes containing liquid reagents. Immediately after the Tempus tube was filled, blood was stabilized by shaking the tube vigorously or vortexing the contents for 10 seconds to ensure that the stabilizing reagent makes uniform contact with the sample. Failure to adequately mix the stabilizing reagent with the blood leads to inadequate stabilization of the gene expression profile and the formation of micro clots, which can potentially clog the purification filter. This was observed for some samples which were excluded from this study. Tempus tubes were stored at -20°C until RNA extraction.

### 3.2.3 Whole blood RNA purification

The total RNA was extracted and purified from the whole blood samples using a commercial kit (Norgen Biotech, Thorold, Canada) following the manufacturer's instructions. Briefly, blood samples in Tempus tubes were first mixed using a shaker at 150 rpm for 10 minutes at room temperature. The samples were then directly poured into a clean 50-mL tube. Diluent Solution was added into the 50-mL tube containing the blood to bring the total volume to 12 mL. The cap was replaced on the tube, and the tubes were vortexed vigorously for 30 seconds to ensure proper mixing of the contents and prevent clogging at the purification step. The tubes were centrifuged at 4 °C at 3,000 x g for 30 minutes. The supernatant was carefully decanted. The RNA pellet is transparent and invisible and tubes must be handled carefully to prevent disturbing the RNA pellet. The tubes were inverted on absorbent paper for 1 to 2 minutes. The remaining drops of

liquid were blotted off the rim of the tube with clean absorbent paper. Lysis Solution (600 μL) provided in the kit was added to each tube, and the tubes were vortexed briefly to re-suspend the RNA pellet. Ethanol (95%-100%, 300 μL) was pipetted to the bottom of each tube and vortexed briefly. Re-suspended RNA (600 μL) was transferred into the purification filter column with a collection tube and then centrifuged at 14,000 x g for 1 minute. The pass-through liquid in the collection tube was discarded. The centrifuge step was repeated until all the re-suspended RNA went through the purification filter. Wash Solution (400 μL) was added into the purification filter and then centrifuged at 14,000 x g for 1 minute. The pass-through liquid in the collection tube was discarded. DNaseI working solution (100 uL) was pipetted into the purification filter column and then centrifuged at 14,000 x g for 1 minute. The liquid in the collection tube was re-added to the purification filter column and then incubated at room temperature for 15 minutes. Without any centrifugation, Wash Solution (400 μL) was directly added to the purification filter column and then centrifuged at 14,000 x g for 1 minute. The liquid waste was discarded from the collection tube and then the washing step was repeated one more time with 400 μL Wash Solution. After discarding the liquid waste, another centrifugation at 14,000 x g for 2 minutes was processed to dry the purification filter. To elute captured RNA, the purification filter column was transferred to a new collection tube and 50 μL Elution Solution was added and then centrifuged at 200 x g for 2 minutes following 14,000 x g for 1 minute. The eluted RNA was transferred to a new 1.5 mL tube and placed at −20 °C for short-term storage or −80 °C for long-term storage.

### 3.2.4 Determination of RNA concentration and quality

The RNA concentration was measured using NanoDrop 1000. Agilent 2200 TapeStation was used to evaluate RNA quality. High-Sensitivity RNA ScreenTape and RNA Sample Buffer (Agilent Technologies, Santa Clara, USA) were maintained at room temperature for 30 minutes before use. The RNA samples were diluted if the concentration was > 400 ng/μL based on the detection range of the RNA ScreenTape kit. 5μL Sample Buffer and 1μL RNA sample were added to an 8-strip tube, then vortexed at 2000 rpm for 1 minute. The samples were spun down and then denatured by heating to 72 °C for 3 minutes followed by placing them on ice for 2 minutes. Samples were loaded into the Agilent 2200 TapeStation instrument and caps were carefully removed from the tube strips. The analysis proceeded using the 4200 Tapestation Controller Software. RNA integrity numbers were recorded to represent the RNA sample quality. Factors including sample size, long-distance shipping and storage may decrease cell viability and RNA integrity, the cutoff for the acceptable RNA integrity number for RNA-seq library preparation was >5.0 based on previous studies (Choi et al., 2014; Jo et al., 2016; Sarathkumara et al., 2022).

### 3.2.5 Globin Reduction and qPCR Confirmation

As the RNA samples were purified from the whole blood, globin RNA mainly derived from red blood cells needs to be removed from all RNA samples to ensure adequate sequencing depth of the rest of the transcriptome. The globin reduction (GR) protocol was developed referring to a previous study (Choi et al. 2014) to specifically remove alpha hemoglobin (HBA) and beta hemoglobin (HBB). The porcine-specific oligonucleotides used are shown in Table 3.2:

**Table 3.3** Sequence of the oligos for GR and qPCR.

| **GR Oligos** | **Sequence** |
|---|---|
| SAG-13Re (HBA) | 5'-CCATTTCGCCCACTCTTAGCATCCAC |
| SAG-14Re (HBA) | 5'-GTGCAAGGGGGGGTGCGCAGAGAC |
| SBAH-37Re (HBB) | 5'-AGGGGAACTTAGTGGTACTTGTGGGC |
| SBAH-38Re (HBB) | 5'-GGTTCAGAGGAAAAAGGGCTCCTCCT |
| **qPCR Primers** | **Sequence** |
| SAG-11FO (HBA primer Forward) | 5'-CCCACCACCCCGATGATTTC |
| SAG-11RE (HBA primer Reverse) | 5'-TCAGCGATCAGGAGGTCAGG |
| SBXS-42F (HBB primer Forward) | 5'-CTCCTGGGCAACGTGATAGT |
| SBXS-38RE (HBB primer Reverse) | 5'-GGTTCAGAGGAAAAAGGGCTCCTCCT |
| Su18S-34F (18S Forward) | 5'-GACAAATCGCTCCACCAACT |
| Su18S-34R (18S Reverse) | 5'-CCTGGGGCTTAATTTGACTC |

A 10X GR oligo mix was prepared by combining the four globin-specific oligos to the final concentrations of 7.5μM (SAG-13Re), 7.5μM (SAG-14Re), 30μM (SBAH-37Re), and 30μM (SBAH-38Re), respectively, listed in Table 3.2. Next, 2μL of the 10X GR oligonucleotide mix was added to 3 μg (~7 μL) of total RNA (incubate RNA samples at 70°C for 2 min and keep on ice before use) and 1μL of 10X oligonucleotide hybridization buffer (100 mM Tris–HCl, pH 7.6; 200 mM KCl) to constitute the 10 uL of hybridization mix. This mix was incubated in a thermal cycler at 70°C for 5 minutes and then cooled to 4°C. The RNA-DNA hybrids were digested with 2 unit RNase H (Ambion, Austin, USA) in the reaction buffer (100 mM Tris–HCl, pH 7.6, 20 mM $MgCl_2$, 0.1 mM DTT, SUPERase-in) at 37°C for 10 minutes and cooled to 4°C. The reaction was stopped by the addition of 0.5 M EDTA. The globin-depleted RNA was immediately purified with the RNeasy MinElute Cleanup Kit (Qiagen, Toronto, Canada) according to the manufacturer's instructions. RNA quality of the globin-depleted samples was assessed using an Agilent Bioanalyzer 2100 (Agilent).

To validate the GR efficiency, the mRNA levels of the porcine HBA and HBB transcripts from GR RNA samples were quantified by SYBR Green-based qPCR assay using a StepOneTM Real-Time PCR System (Applied Biosystems). The first strand cDNA reaction was prepared as follows: 250 ng (~6 μL) GR RNA, 5 μL Random Hexamer (50 ng/μL), and 1 μL dNTP (10 mM), and then heated at 65°C for 5 minutes following with adding 4 μL 5x First Strand Buffer, 2 μL DTT (0.1 M), 1μL RNaseOUT (40 U/μL) and kept at 25°C for 2 minutes. Next, 1 μL SuperScript® II reverse transcriptase (Thermo Fisher Scientific) was added and the reaction was incubated at 25°C for 10 minutes, 42°C for 50 minutes, then 70°C for 15 minutes. qPCR was

performed in a total volume of 10 μL per reaction comprising 2 μL of the template (1:10 diluted cDNA), 1 μL of the assay-specific primer mix, 5 μL of the Fast SYBR® Green Master Mix Bulk Pack (Applied Biosystems) and 2 μL of water. The PCR condition used was 95°C for 3 minutes for initial denaturation, 23 cycles of 95°C for 30 seconds, and 60°C for 30 seconds. The primer sequences are listed in Table 3.2.

### 3.2.6 Strand-specific cDNA library construction and sequencing

For the 58 RNA samples from cohort 1, 1 μg of GR RNA was used for strand-specific cDNA library construction using the NEBNext® Ultra Directional RNA Library Prep Kit (Illumina, San Diego, USA) according to the manufacturer's protocol. Briefly, mRNA was enriched using oligo(dT) beads from the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, Ipswich, USA) followed by two rounds of purification, and then fragmented randomly by adding fragmentation buffer. The first strand of cDNA was synthesized using random hexamer primers. The second strand of cDNA was generated by incorporating dUTP in place of dTTP to create blunt-ended cDNA. After a series of terminal repairs, poly-acetylation, and sequencing adaptor ligation, the double-stranded cDNA library was completed after size selection and PCR enrichment. Library preparation for the 500 RNA samples of the 209 pigs from cohort 2 was performed by Novogene Inc (Sacramento, USA). GlobinClear kit (Thermo Fisher Scientific) was used to remove globin RNA and Ribo-Zero kit (Illumina) was used to remove rRNA. The library was prepared using the same kit introduced above from NEB.

The resulting 250-350bp insert libraries were quantified using the Qubit 2.0 fluorometer (Thermo Fisher Scientific) and quantitative PCR. Size distribution was

analyzed using an Agilent 2100 Bioanalyzer (Agilent Technologies). Qualified libraries were sequenced on an Illumina HiSeq 4000 Platform (Illumina) or NovaSeq 6000 (Illumina) using a paired-end 150 run (2×150 bases). 30 million raw reads were generated from each library.

### 3.2.7 Data analysis

Raw single reads were subjected to sequence quality control using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) in the context of reading length, GC content, and repeated sequence to determine if the samples had adequate read quality for downstream analyses.

RNAseq data analysis was performed following a previously published protocol (Pertea et al., 2016) and the workflow is illustrated in Figure 3.2. In brief, filtered reads were aligned to the swine reference genome (Sus_scrofa.Sscrofa11.1) using HISAT2 software (v2.1.0), then assembled and quantified for each sample with StringTie (v1.3.3). After merging transcripts from all samples using StringTie, gffcompare was used to do statistical analysis for the StringTie output with all the transcripts including novel exons, introns, or genes. Meanwhile, after the merging step, StringTie was run one more time to re-estimate the abundance of the merged transcripts and create table counts for the next step. The re-estimation step used the same abundance estimation algorithm as that used in the initial steps. Here, Ballgown (v2.6.0) was used to plot raw data, to do normalization and downstream statistical modeling for differential expression analysis. Batch effect was corrected using the function “removeBatchEffect” from the package limma. For pathway enrichment analysis for the DE genes, the gseGO program from the package clusterProfiler was used.

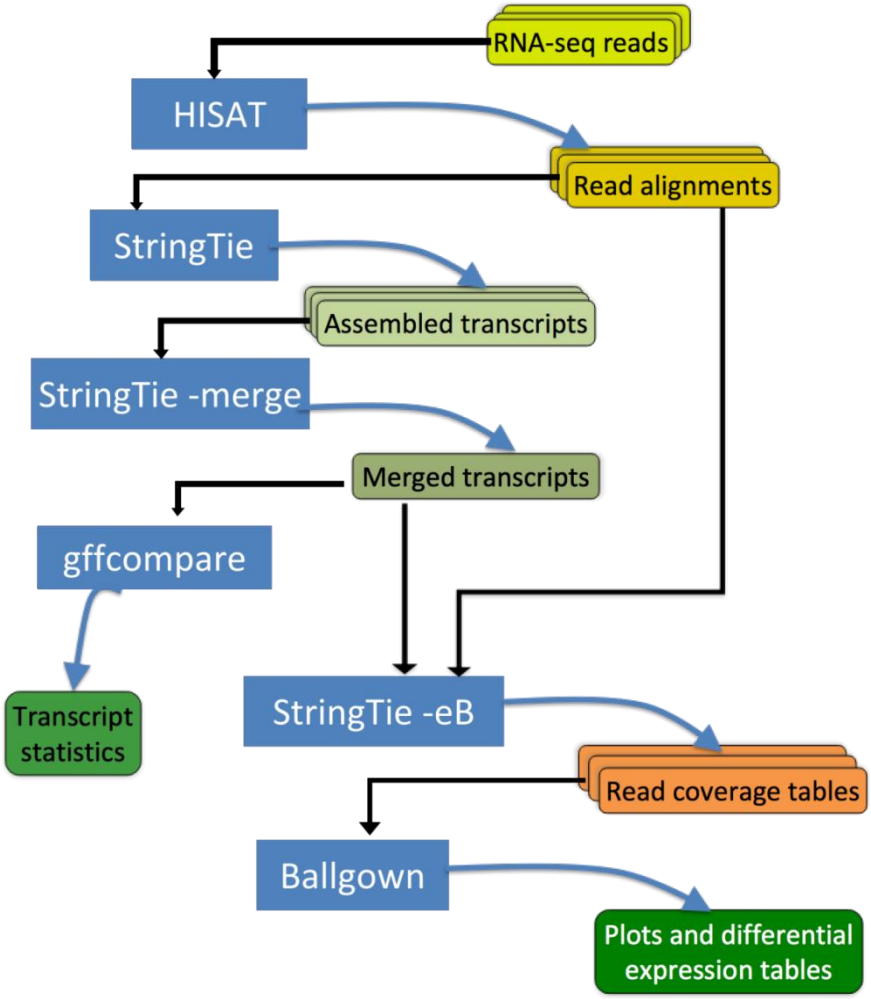



**Figure 3.2** Workflow of RNA-Seq data analysis.

One common issue with RNA-seq data is that some genes may have very few or even zero counts. A common step before gene expression analysis is to filter out those genes. Another approach that has been used for gene expression analysis is to apply a variance filter. Here, all transcripts were removed with variance across samples less than the count of one (*rowVars(texpr(bg)>1*). Then, for differential expression (DE) analysis transcripts and genes that showed statistically significant differences between phenotype groups (Resistant, Resilient, Early_dead, susceptible) were identified using the *stattest* function from Ballgown, which is a statistical test using a standard linear model-based comparison. At this point, the abundance estimates (expressed as FPKM values which stand for fragments per kilobase of transcript per million mapped reads) have already been normalized with respect to library size. I examined both genes and transcripts that were differentially expressed between phenotype groups (Resistant, Resilient, Early_dead, Susceptible) while correcting for any difference in expression due to the batch variable.

eQTL analysis was performed using the "MatrixEQTL" program on candidates with Q-value<0.1 from the DE results. Genotyping results of experimental pigs were acquired using Axiom Pig HD panel SNPs (Applied Biosystems, Waltham, USA) and the reads were mapped to the porcine reference genome Sscrofa11.1. The cutoff used for the minor allele frequency (MAF) was 0.05.

## 3.3 Results

### 3.3.1 Summary of major traits of the four groups of pigs regarding productive performance and medical treatment from the natural disease challenge experiment.

To group the 58 pigs in cohort 1, mean death weight (MDW) or mean pre-slaughter weight (MPSW) were used, respectively, for evaluating the productivity index for the Early_dead group or the remaining three groups of survived pigs (Figure 3.3A). The pigs from the Resilient and Resistant group had relatively higher production compared with the other two groups. Of note, the pigs from the Resilient group maintained the highest MPSW after the disease challenge nevertheless they got sick and received treatment (Figure 3.3A), however, the frequency of medical treatments for pigs from the Resilient group was lower than that from the Early_dead and Susceptible group (Figure 3.3B).

Our collaborators reported recently that the proportion of "off-feed" days during the disease challenge classified from the quantile regression analysis for feed intake could be used to define pig disease resilience phenotype (Putz et al., 2018). To test whether our grouping method is consistent with using this novel trait to differentiate pigs for their responses to disease challenge, I selected the 58 pigs described in Table 3.1 and compared their proportion of "off-feed" days from the data. As shown in Figure 3.4, the Resilience and Resistance groups display the lowest proportion of "off-feed" days among the four groups, implying that the resistant and resilient pigs could maintain regular feed intake when they were under disease challenge stress. This result indicates that our grouping perfectly matches the notion using feed intake-derived "off-feed" days proportion to define disease resilience phenotype.

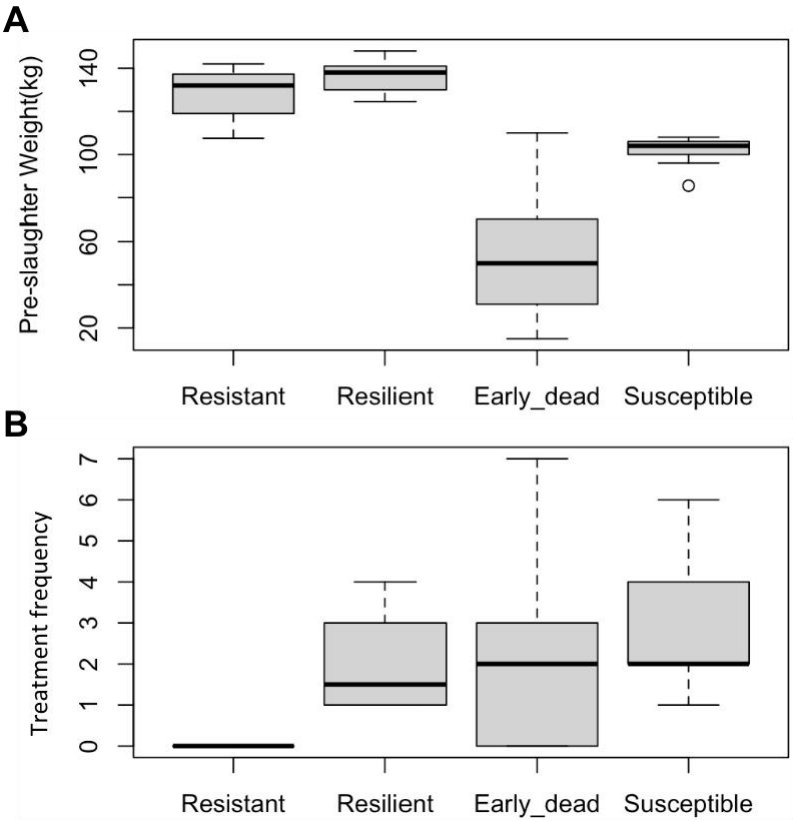



**Figure 3.3** Comparison of the two major grouping parameters including MPSW or MDW (panel A) and treatment frequency (panel B) for the four groups of pigs denoted as "Early_dead", "Resilient", "Resistant" and "Susceptible" in cohort 1.

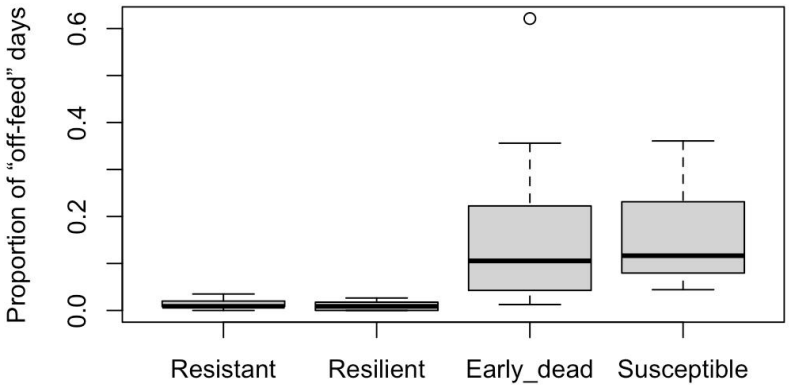



**Figure 3.4** Comparison of "off-feed" days proportion among the "Early_dead", "Resilient", "Resistant" and "Susceptible" groups in cohort 1 pigs.

The proportion values were derived from a 5% quantile regression analysis using the overall feed intake data of the 58 pigs and then aggregating off-feed days within each animal as a proportion.

### 3.3.2 Quality control and mapping of the RNA-Seq data

Next-generation sequencing was conducted to determine the transcriptome of 58 pig whole blood samples collected around 28 days of age before the disease challenge. A total of 305 gigabytes of raw reads in fastq format were obtained from 58 libraries. FastQC tool was used to determine the read quality (Appendix 3.1). As illustrated in Appendix 3.1 B and C, nearly 100% of the reads displayed a Phred score over 30 (Q30: a probability of incorrect base calling of 1 in 1000) across the full length 150 bp, indicating a superior per base quality of all the reads. The alert shown in the section of "Sequence Duplication" (Appendix 3.1A) was not unexpected for RNA-Seq data while it was reasonable that some genes were highly expressed and thus represented in multiple copies from sequencing. The only concern was that some samples had a warning for GC content presented. This was probably due to some highly expressed genes in those samples which caused GC bias resulting from those overrepresented sequences. The alert of overrepresented sequences (including sequencing adapters) was alleviated or eliminated by using filtering and trimming tools. Overall, the RNA-Seq dataset was of good quality.

Based on the mean library size of ~32.4 million base pairs for all 58 samples, the average overall mapping rate was 94.0% which includes 83.8% unique mapped reads and 6.0% multi-mapped reads. The infer_experiment.py in SAMtools was used to check the strandedness and it returned the results for the fraction of reads explained by "1++,1--,2+-,2-+", "1+-,1-+,2++,2--", and "failed to determine" rate was 93.98%, 4.80%, and 1.22%, respectively, confirming a successful strand-specific library construction and read mapping were properly conducted. Among the strand-specifically mapped reads, approximately 50.43% of reads were from the positive strand while 49.57% were from

the negative strand, indicating the genes in whole blood samples were nearly equally expressed from the two DNA strands.

After mapping the reads to the pig reference genome, read counting was processed using the StringTie tool. As shown in Appendix 3.2A and B, a total of 40,465 genes were called with expression. More than half of genes (25,388) only expressed one specific transcript whereas nearly one-third of genes (15,077) expressed more than one specific transcript, indicating frequent alternative splicing activities. The maximum number of transcripts from one gene was 492 (*T cell receptor beta variable gene*). For the distribution of transcript length shown in Appendix 3.2B, the maximum, minimum, and median lengths of detected transcripts are 25,893bp (*PALB2* gene), 11bp (ENSSSCT00000057789), and 2,486bp (*MECP2* gene), respectively, while the mean length of all transcripts is 3,019bp. In addition, 98% of transcripts were shorter than 10kb and among them, 80.7% of transcripts were smaller than 5kb. The FPKM value accounts for sequencing depth and all 58 RNA samples display similar patterns of log2(FPKM) as shown in Appendix 3.2C, indicating comparable sequencing depth and diversity of transcriptome. 99.1% log2(FPKM+1) values were less than 5, indicating that only a small number of genes were expressed at very high levels. Genes with low expression levels could reduce the sensitivity to detect real DE genes among samples (Sha et al., 2015). Thus, the genes with very low expression were removed from DE analysis.

The "GFFcompare" program was used to determine how well the assembled transcripts matched gene annotation information. For the query of 151,368 transcripts in 41,017 loci, there were 100%, 98.9%, and 100% matched exon, intron, and loci, respectively, indicating a high coverage of the transcriptome. Of note, there were

130,317 novel exons, 68,574 novel introns, and 17,456 novel loci discovered as well. However, in the absence of an experimental test of its effectiveness, it is difficult to determine which transcripts are real and which are transcriptional "noise".

### 3.3.3 Global variance analysis for the 58 RNA samples from cohort 1 pigs using principal component analysis (PCA)

To explore whether there was global difference in their transcriptional profiles between each individual group or each blood collection, PCA analysis was performed on the transcriptome counts (FPKM) of the 58 samples from cohort 1. As illustrated in Figure 3.5 A, the distribution of pigs from the four groups was largely overlapping from any blood collection; neither genes or transcript level revealed clear compartmentalization from any specific group to the other three. Therefore, there was no obvious global difference in the pre-challenge transcriptome of the 58 pigs even though the four groups expressed differential phenotypes in the natural disease challenge model.

### 3.3.4 Differential gene expression analysis among the "Early_dead", "Resilient", "Resistant" and "Susceptible" groups.

It is difficult to differentiate the four groups of pigs based on their global transcriptomic profile. DE analysis was therefore conducted, aiming to explore specific genes that are distinctly expressed within each of the four groups. I first compared the gene and transcript expression of one group with those of the other three groups (Figure 3.8). One, three, and one significant DE gene were observed respectively in the Resistant, Resilient, and Early_dead groups, compared to the rest of the groups. Whereas, eight significantly up-regulated genes (including several long-non-coding RNA genes) were

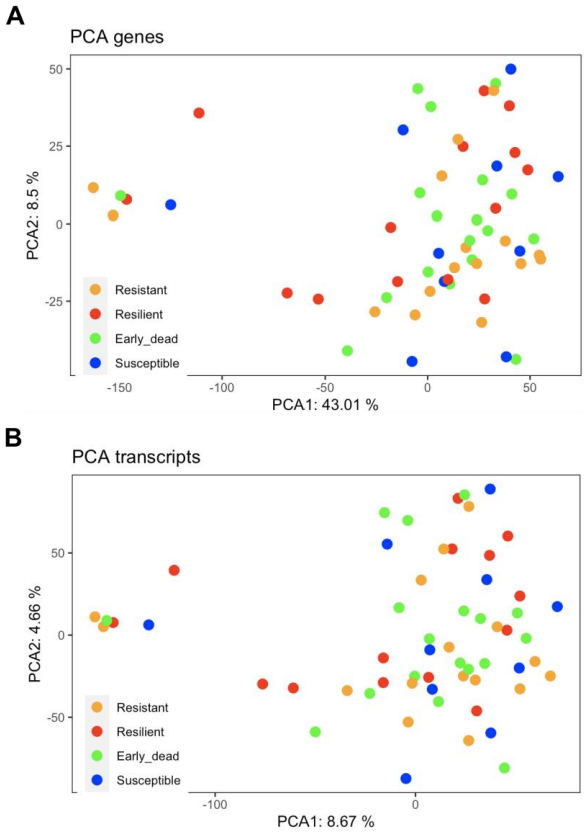



**Figure 3.5** PCA plot of transcriptomic profiles of the 58 pigs based on genes (A) and transcripts level (B).

Each dot represents a pig. The different colors indicate pigs corresponding to the four defined groups. Batch effect was removed.

found in the Susceptible group, only two and three genes were found significantly down-regulated specific for the Resilient and Susceptible groups, respectively.

Notably, there were more DE transcripts than genes identified in each of the four groups specifically compared to the others. The names and numbers of DE genes and transcripts (the 58 pigs and the second bigger cohort described below) are listed in Appendix 3.3. Within the Resistant group, DE genes were found to be involved in cytoskeleton remodeling (*PFN1*, *TES*, *NISCH*), cell signaling transduction (*SMAP2*, *AXIN1*, *MYL9*, *SLA5*, *SPPL2A*, *PTK2B*), transcription factor (*BCL11B*, *ATOH8*), cell metabolism (*MGAT4B*, *PAFAH2*, *NDUFS7*), cell survival and apoptosis (*RFFL*), and chemotaxis (*CCL5*). The DE genes from the Resilient group were associated with immune response (*IL4R*, *IL6ST*, *TMEM106A*), transcription factors (*MXD1*, *PMF1*), cell metabolism (*B3GNT3*, *B4GALT1*, *NDUFS1*, *SLC4A7*), protein modification (*UBE2L6*, *BIRC3*, *NTAN1*), cell signaling transduction (*LYN*, *PRKAR1A*, *PRKACA*), cell survival and apoptosis (*DNAJA3*), and unclear function (*TNRC18*). The DE genes specific for the Early_dead group were related to diverse cellular processes including cell signaling transduction (*CLTC, SNCG, SPPL2A, PLEK*), mRNA splicing (*SRRM2, CTNNBL1*), transcription regulation (*NCOA4, RERE*), protein translation (*EIF3G*), cell metabolism (*CKB*), protein degradation (*LAPTM5, CUL4B*), cell survival and apoptosis (*RNASEL*), cytoskeleton remodeling (*PSTPIP2*), protein transportation (*COPG2, SNAP23*), and microRNA (*ssc-mir-6782*). The Susceptible group were significantly enriched with the two undefined transcripts MSTRG.32803 (q-value=0.02) and MSTRG.21242 (q-value=0.09) and those two transcripts were the only two hits in the DE analysis with a q-value<0.1. The other DE genes found in the Susceptible group were suggested to be

involved in protein modification (*UBR4, RNF151*), nucleic acid binding (*ZCCHC2*), transcription regulation (*KLF17, ILF2*), mRNA splicing (*RBM17*), regulation of secretion (*SV2A*), cell signaling transduction (*ATL3, NFKBIZ, PTP4A1*), structural constituent of the cytoskeleton (*CCDC6*) and cell metabolism (*AGK, GAPDH*). Interestingly, the Susceptible group may have some unique features in mRNA splicing as there were two genes (*KLF17, RNF151*) found with both up-regulated and down-regulated transcripts. So far, at the pig's steady state, it seemed difficult to predict the disease response using the transcription level of specific genes. Nevertheless, I detected multiple DE of group-specific transcripts using a cut-off of P value<0.05 for all four groups of pigs, and two significantly up-regulated transcripts (q-value<0.1) specific for the Susceptible group.

To further characterize differential transcription profiles between each of the groups, I did the DE analysis by comparing pairs of groups (Figure 3.7). As shown in Figure 3.7A, more DE genes were identified when this comparison was performed (rather than one versus the other three) (Figure 3.6A). Some positive immune regulation genes including *CEBPD* (Ko et al., 2015; Spek et al., 2021) and *ATRN* (Duke-Cohan et al., 2000), were up-regulated in the Resistant group compared with the Resilient, Early_dead, and Susceptible groups. Not so many DE genes directly involved in immune defense were detected in the Resistant group which is presumed to have stronger "immunity". The identified genes with immune-priming roles might explain their better control of infection of the Resistant group after being exposed to pathogen challenge. The higher immune-sensing feature of resistant pigs could be supported by some DE transcripts e.g. *CCL5* (encodes a chemokine important in immune defensive and immune cell survival

(Silva et al., 2021; Tyner et al., 2004)) and *LYN* (encodes a kinase important in immunoreceptor signaling of B cells and some other immune cells (Brian & Freedman, 2021)) enriched in the Resistant compared to the Resilient group (Figure 3.7B). In comparison between the Resilient and the Early-dead pigs, the Resilient pigs had higher expression of the *CXCR5* gene which encodes an important chemokine receptor and is regarded as a marker of T follicular cells (Moser, 2015), indicating a more sensitive humoral response potential. The stronger immune response potential of the Resilient group pigs (than the Early_dead ones) is also supported by the higher expression of *IL4R* transcripts. In comparison between the Resilient and the Susceptible group, the Resilient group up-regulated *PPP3CA* which encodes a unit of calcineurin and plays an important role in T cell activation through dephosphorylating the NFAT dependent on calcium (Kilka et al., 2009; Park et al., 2020), implying a higher sensitivity of T cell response. The higher expression of a *TNFSF8* transcript encoding a pro-inflammatory cytokine in resilient pigs (compared with the susceptible ones) also provides a hint for a more active immune response. Meanwhile, the Resilient group has more expression than the Susceptible group for the *BTG1* gene which encodes an anti-proliferative protein, indicating slower proliferation of immune cells. This is consistent with a hypothesis that the immune cells in resilient pigs consume less energy from excessive proliferation and thus have higher productive performance than susceptible pigs. To this end, DE analysis results are consistent with the superior performance of infection control from the Resistant group pigs among all the pigs and the better pathogen control ability and productivity of the Resilient group than the Early_dead group and the Susceptible group.

### 3.3.5 Pathway enrichment analysis for the DE genes

In order to gain deeper insights into the biological implications of the identified transcripts in the DE analysis, Gene Ontology (GO) term enrichment analysis was conducted. It should be noted that this analysis excluded the genes showing both up-regulated and down-regulated transcripts in the DE analysis. As shown in Figure 3.8, three out of the four groups were successfully enriched with group-specific biological pathways except for the Early_dead group. Unexpectedly, even though several pathways associated with gene transcription were down-regulated in the Resistant group comparing to the other groups, only one pathway (transmembrane transport) was up-regulated and no immune-related pathway was enriched. Interestingly, specific for the Resilient group, most of the top down-regulated pathways were involved in cytoskeleton organization (Figure 3.8). During the immune surveillance and activation, the cytoskeleton participates in various pivotal processes including cell migration, phagocytosis, synapse formation, secretion and degranulation (Wickramarachchi et al., 2010). Lower expression of cytoskeleton organization pathways in resilient pigs may compromise their immune defense to some extent whereas decrease energy consumption from immune activities such as cell migration which is dependent on cytoskeleton remodeling and energetically demanding (Guak and Krawczyk, 2020). These findings may shed the light of why resilient pigs cannot efficiently prevent pathogen invasion but could maintain relatively high productivity during infections. Conversely, susceptible pigs up-regulated those cytoskeleton organization pathways, leading to more energy consumption. The Susceptible group also up-regulated pathways of negative regulation of signal transduction and response to stimulus, indicating the immune cells in susceptible

pigs may not respond well to pathogen infection and might be the key reason why they are susceptible. Down-regulation of several metabolic pathways in susceptible pigs may hinder their energy supply from nutrients. Therefore, the enriched specific pathways provide evidence for a less potent function of immune cells in susceptible pigs compared with resilient pigs.

Through comparing each pair of the four groups, I did not see any immune-related pathway that is differentially enriched between resistant and resilient pigs, suggesting there was not dramatic difference in the basal level immunity of pigs from these two groups when they were healthy. However, some immune response pathways were found to up-regulate in resistant pigs compared to the Early_dead or Susceptible groups. This strongly supports the notion that the immune cells in resistant pigs are endogenously more potent than that in the Early_dead or Susceptible group, but not in the Resilient group. It is still difficult to get any insight to explain the higher productive performance in resilient pigs than the others as they up-regulated many proteins in the biosynthesis-related pathways which are also ATP-dependent. Thus, it remains unclear whether the relatively higher productive performance of the resilient pigs resulted from their energy-saving immune response. Last but not least, compared to the Susceptible group, the Early_dead group was found to down-regulate pathways associated with immune cell activation and cell-cell interaction, suggesting the pigs might die from the inefficient control of pathogen infection by the immune system. Taken together, the pathway enrichment analysis suggests a hierarchical baseline immunity in all four groups of pigs tested in their healthy state, ranked by Resistant>Resilient>Susceptible>Early_dead.

It is difficult to identify one or several marker genes' expression patterns to predict a

specific phenotype such as resistant or resilient while it seems feasible to use marker gene expression to exclude some pigs that may have unfavorable phenotypes and thus narrow down the scale of the breeding population. To explore whether it was possible to use gene set information to predict specific pig response to disease challenge, a machine learning approach was used but no favorable model was found to provide high confidence prediction (See discussion in Chapter 5). The large individual variations and small cohort size may cause the failure of identifying promising hits from the RNA-Seq. A larger cohort with fewer batches of pigs is needed to validate our findings.

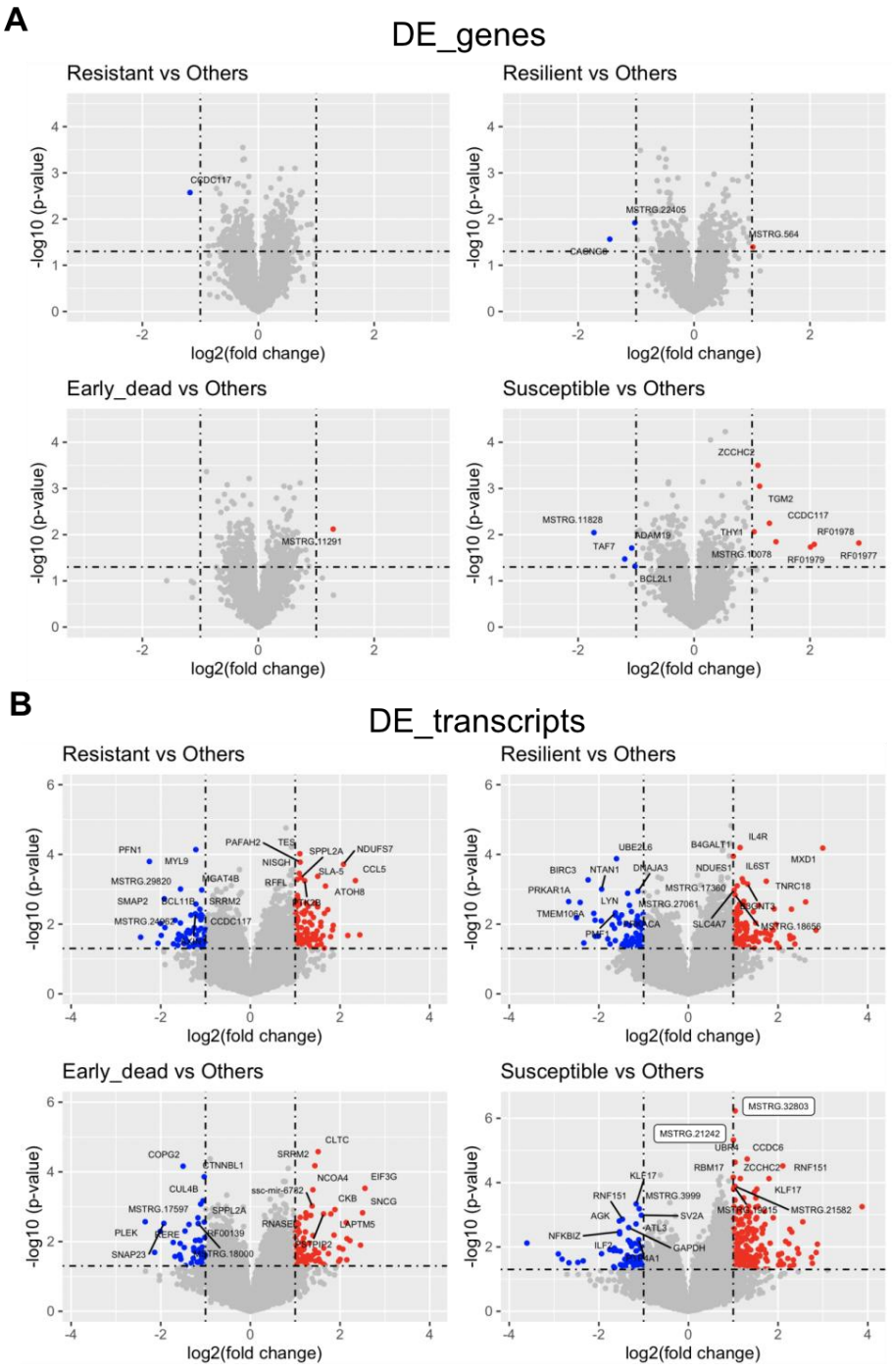



**Figure 3.6** Differential expression of genes (A) and transcripts (B) level comparing one group with the rest of the other groups in cohort 1.

Genes or transcripts with differential expression were highlighted in red (up-regulated) and blue (down-regulated) using a cut-off value of 2 for fold change and 0.05 for P-value. The names of the top 10 genes and transcripts with the lowest P-values were labeled inside each plot.

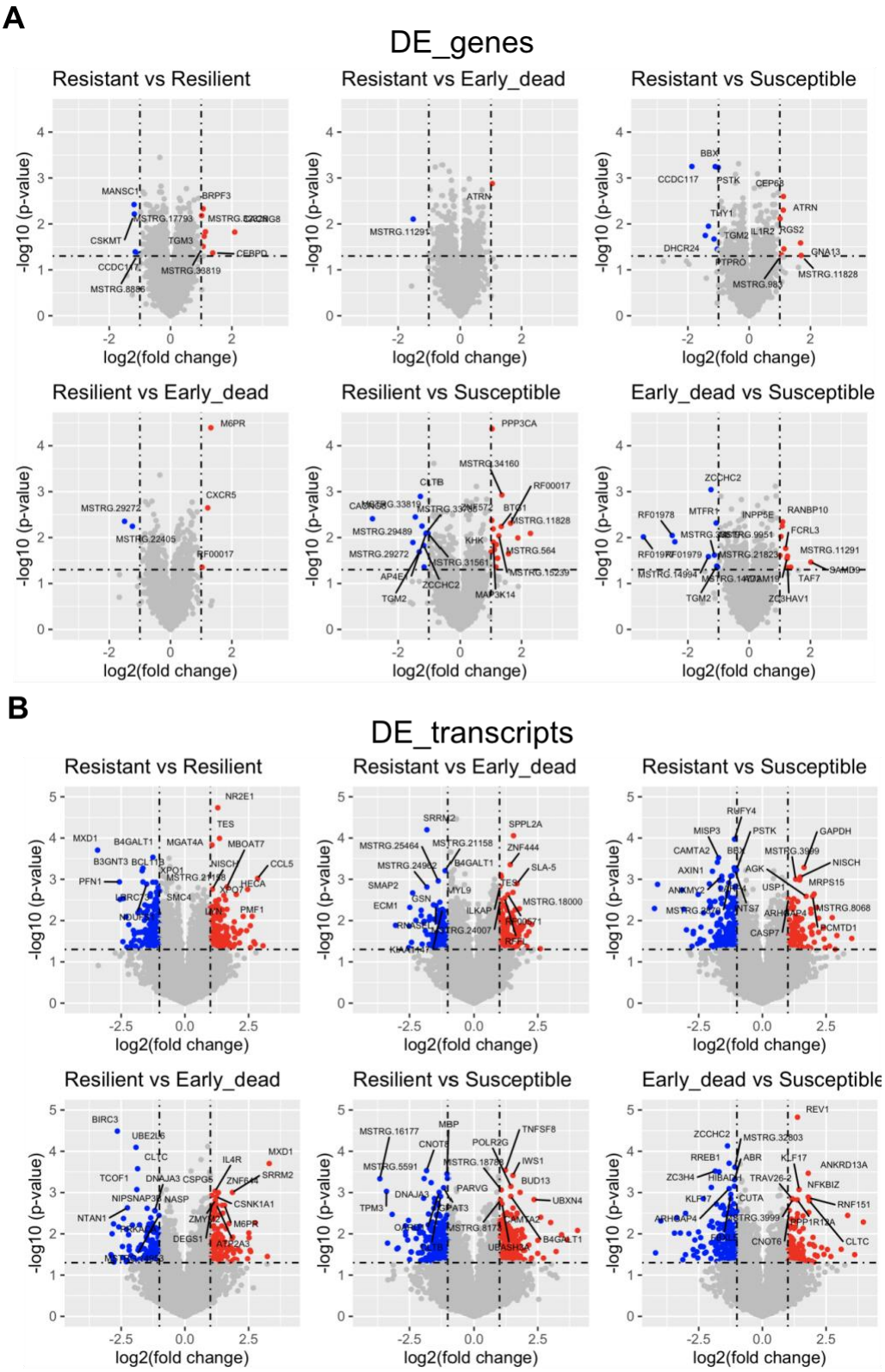
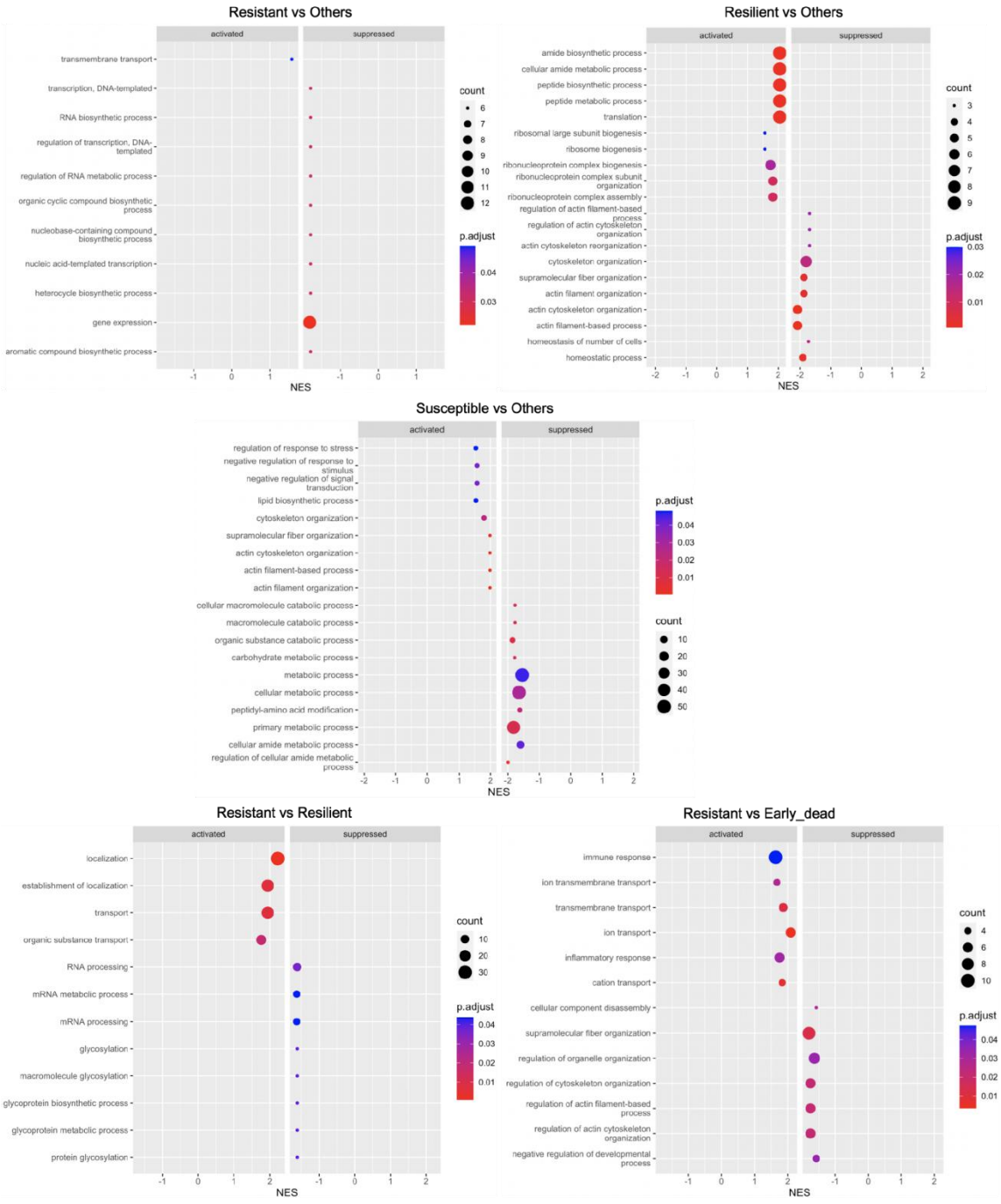



**Figure 3.7** Differential expression of genes (A) and transcripts (B) level comparing two specific groups in cohort 1 pre-challenge samples.

Genes or transcripts with differential expression were highlighted in red (up-regulated) and blue (down-regulated) using a cut-off value of 2 for fold change and 0.05 for P-value. The names of the top 10 genes and transcripts with the lowest P-values were labeled inside each plot.


Resistant vs Others
activated
suppressed
transmembrane transport
transcription, DNA-templated
RNA biosynthetic process
regulation of transcription, DNA-templated
regulation of RNA metabolic process
organic cyclic compound biosynthetic process
nucleobase-containing compound biosynthetic process
nucleic acid-templated transcription
heterocycle biosynthetic process
gene expression
aromatic compound biosynthetic process
count
p.adjust
NES
Resilient vs Others
activated
suppressed
amide biosynthetic process
cellular amide metabolic process
peptide biosynthetic process
peptide metabolic process
translation
ribosomal large subunit biogenesis
ribosome biogenesis
ribonucleoprotein complex biogenesis
ribonucleoprotein complex subunit organization
ribonucleoprotein complex assembly
regulation of actin filament-based process
regulation of actin cytoskeleton organization
actin cytoskeleton reorganization
cytoskeleton organization
supramolecular fiber organization
actin filament organization
actin cytoskeleton organization
actin filament-based process
homeostasis of number of cells
homeostatic process
count
p.adjust
NES
Susceptible vs Others
activated
suppressed
regulation of response to stress
negative regulation of response to stimulus
negative regulation of signal transduction
lipid biosynthetic process
cytoskeleton organization
supramolecular fiber organization
actin cytoskeleton organization
actin filament-based process
actin filament organization
cellular macromolecule catabolic process
macromolecule catabolic process
organic substance catabolic process
carbohydrate metabolic process
metabolic process
cellular metabolic process
peptidyl-amino acid modification
primary metabolic process
cellular amide metabolic process
regulation of cellular amide metabolic process
p.adjust
count
NES
Resistant vs Resilient
activated
suppressed
localization
establishment of localization
transport
organic substance transport
RNA processing
mRNA metabolic process
mRNA processing
glycosylation
macromolecule glycosylation
glycoprotein biosynthetic process
glycoprotein metabolic process
protein glycosylation
count
p.adjust
NES
Resistant vs Early_dead
activated
suppressed
immune response
ion transmembrane transport
transmembrane transport
ion transport
inflammatory response
cation transport
cellular component disassembly
supramolecular fiber organization
regulation of organelle organization
regulation of cytoskeleton organization
regulation of actin filament-based process
regulation of actin cytoskeleton organization
negative regulation of developmental process
count
p.adjust
NES

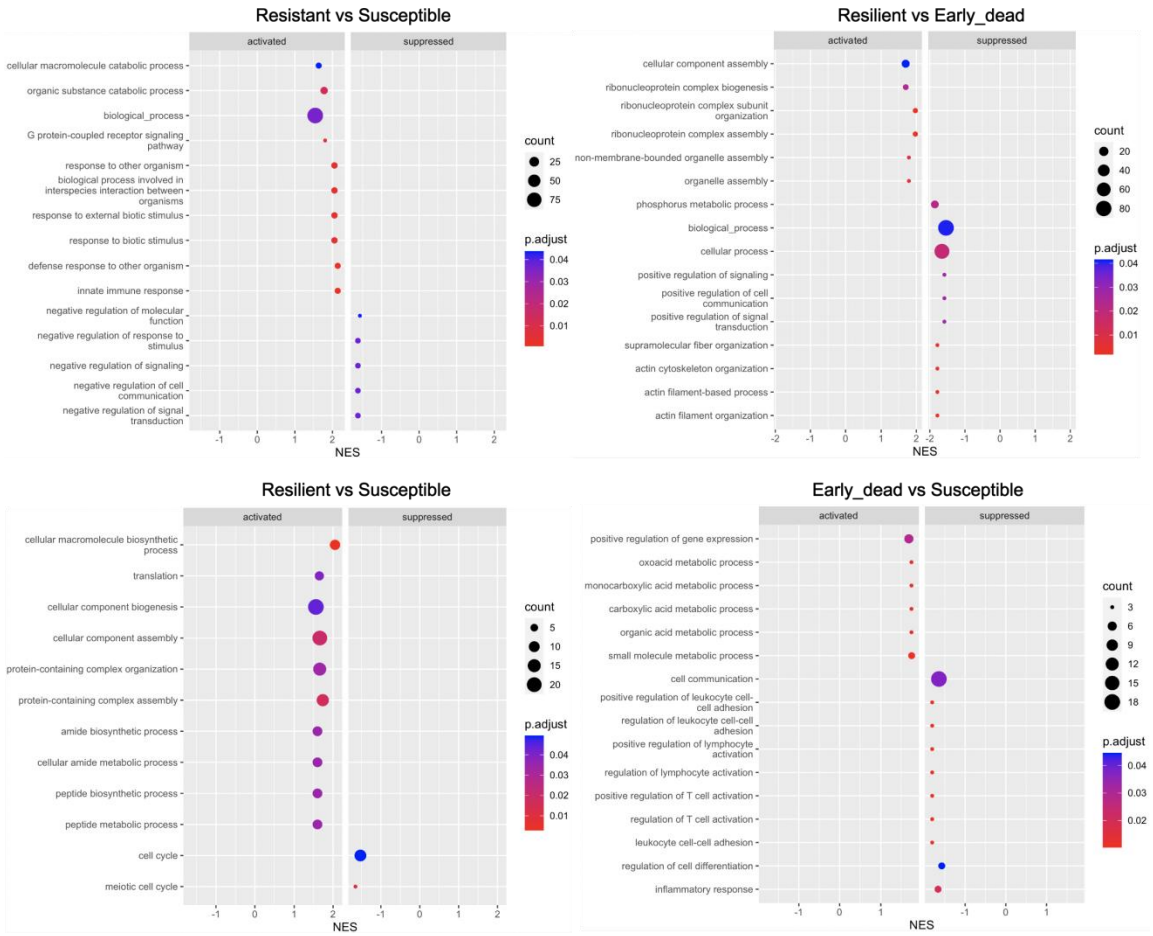



**Figure 3.8** GO term enrichment analysis for comparisons between one specific group and the others, and between two specific groups in cohort 1.

The top 10 (if applicable) up-regulated (left panel) and down-regulated (right panel) GO term groups are listed for each comparison ordered by normalized enrichment scores.

### 3.3.6 eQTL analysis for the transcript MSTRG.32803.53 and MSTRG.21242.14

Only two DE transcripts MSTRG.32803.53 and MSTRG.21242.14 with q-value<0.1 were identified specifically between the Susceptible group with the other groups (Figure 3.8B), and both transcripts were significantly up-regulated in the Susceptible group as shown by the box-and-whiskers plot (Figure 3.9). To better understand these two unique transcripts, I further characterized the genome locations where they were transcribed from (Table 3.4). MSTRG.32803.53 had three exons and the Exon 1 was perfectly aligned to an *IgC* gene which encoded the constant chain of immunoglobulin. MSTRG.21242.14 contained 9 exons and Exon 3 to 8 could be aligned to the signaling lymphocytic activation molecule family member 9 gene (*SLAMF9*).

eQTL analysis was performed for the two aforementioned transcripts in the last paragraph using the genotyping data from 57 pigs (one pig was excluded because of no genotype information) (Figure 3.10). The SNP *Affx-114651911* and *Affx-115202865* associated with the expression of MSTRG.32803.53 and MSTRG.21242.14 are located on chromosomes 12 and 13, respectively. In both cases, the pigs with higher expression of the target transcripts (i.e. a plausibly "susceptible" phenotype) were heterozygous genotypes at the corresponding SNP site and appeared rare in the tested population. Further characterization of the two SNPs revealed that *Affx-114651911* is located in the intron of the gene *ENSSSCG00000017712* and *AP2B1*, and *Affx-115202865* is located in the upstream region of the gene *ENSSSCG00000022500.2* while downstream of the gene *ENSSSCG00000025701.2*, *NCK1*, and *MIR9850*.

### 3.3.7 Validation of RNA-Seq findings from the 58 cohort 1 pigs using pre-challenge blood from another 209 pigs (cohort 2)

As mentioned above, to validate the results derived from the 58 pigs in cohort 1, I repeated the RNA-Seq experiment using another 209 pigs from Cycle 7 (cohort 2). Notably, not only pre-challenge blood samples (Blood 1) were collected, two collections of post-challenge blood samples (Blood 3, 4) were also included in the RNA-Seq to further understand the underlying molecular mechanisms contributing to the differential phenotypic responses of the four groups of pigs.

Grouping of pigs in cohort 2 was totally following the rules for cohort 1 elaborated above. Like the grouping of cohort 1 pigs (Figure 3.3), the Resilient pigs had the highest Pre-slaughter weight representing the outstanding production performance while received least treatment times among the sick pigs (Figure 3.11 A and B). The grouping rationality was tested by plotting the proportion of "off-feed" days for each defined groups (Figure 3.11 C), and the Resilient and Resistant groups were the lowest for this parameter, which was also consistent with what had been shown for cohort 1 (Figure 3.4).

Sequencing of the 500 samples generated 2.2 terabyte raw data which showed comparable library size with cohort 1 for each sample. Likewise, quality check of the new RNA-Seq dataset came with consistent and high quality compared to cohort 1 samples (data not shown). Processing of the data including reads trimming, filtering, mapping, assembling, and counting was exactly following the methods described for the cohort 1 samples.

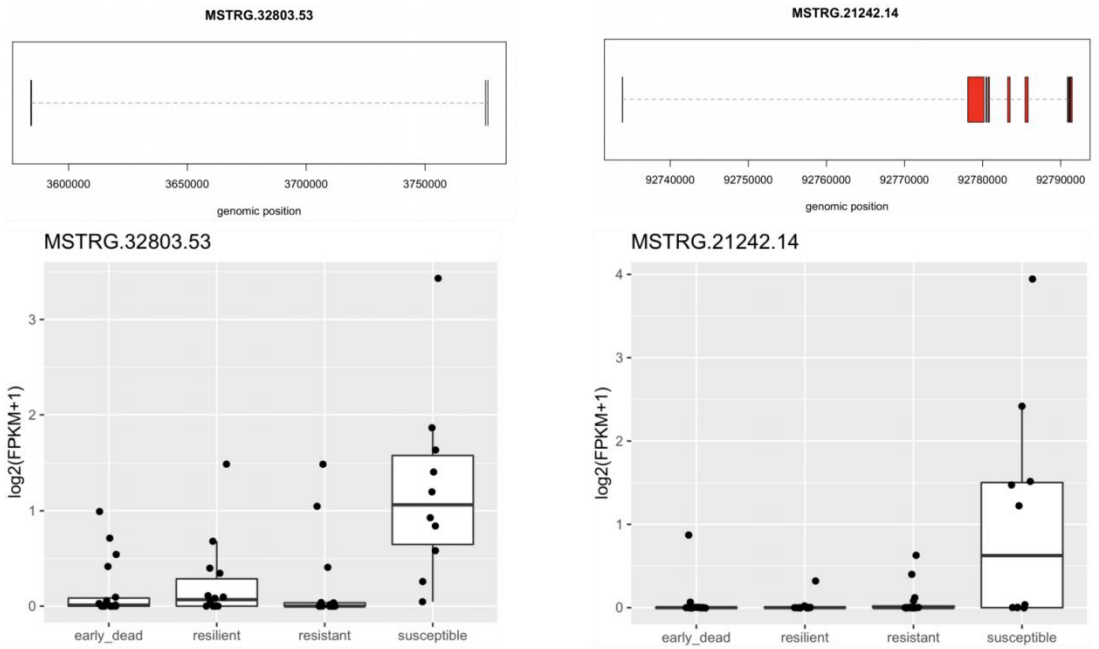



**Figure 3.9** Transcript structure and comparison of FPKM values among the four phenotype groups of pigs for MSTRG.32803.53 and MSTRG.21242.14.

**Table 3.4** Location of genomic loci encoding MSTRG.32803.53 and MSTRG.21242.14

| MSTRG.32803.53 | Chromosome | Start position | End position | Size (bp) |
|---|---|---|---|---|
| Exon1 | AEMK02000452.1 | 3584116 | 3584414 | 299 |
| Exon2 | AEMK02000452.1 | 3775469 | 3775503 | 35 |
| Exon3 | AEMK02000452.1 | 3776442 | 3776451 | 10 |
| **MSTRG. 21242.14** | **Chromsome** | **Start position** | **End position** | **Size (bp)** |
| Exon1 | chr4 | 92733869 | 92733901 | 33 |
| Exon2 | chr4 | 92778097 | 92780173 | 2077 |
| Exon3 | chr4 | 92780439 | 92780524 | 86 |
| Exon4 | chr4 | 92780716 | 92780859 | 144 |
| Exon5 | chr4 | 92783207 | 92783503 | 297 |
| Exon6 | chr4 | 92785440 | 92785814 | 375 |
| Exon7 | chr4 | 92790867 | 92790971 | 105 |
| Exon8 | chr4 | 92791089 | 92791147 | 59 |
| Exon9 | chr4 | 92791229 | 92791456 | 228 |

Note: MSTRG.32803.53 transcript was not mapped on any annotated pig chromosome but mapped on an unplaced assembly AEMK02000452.1.

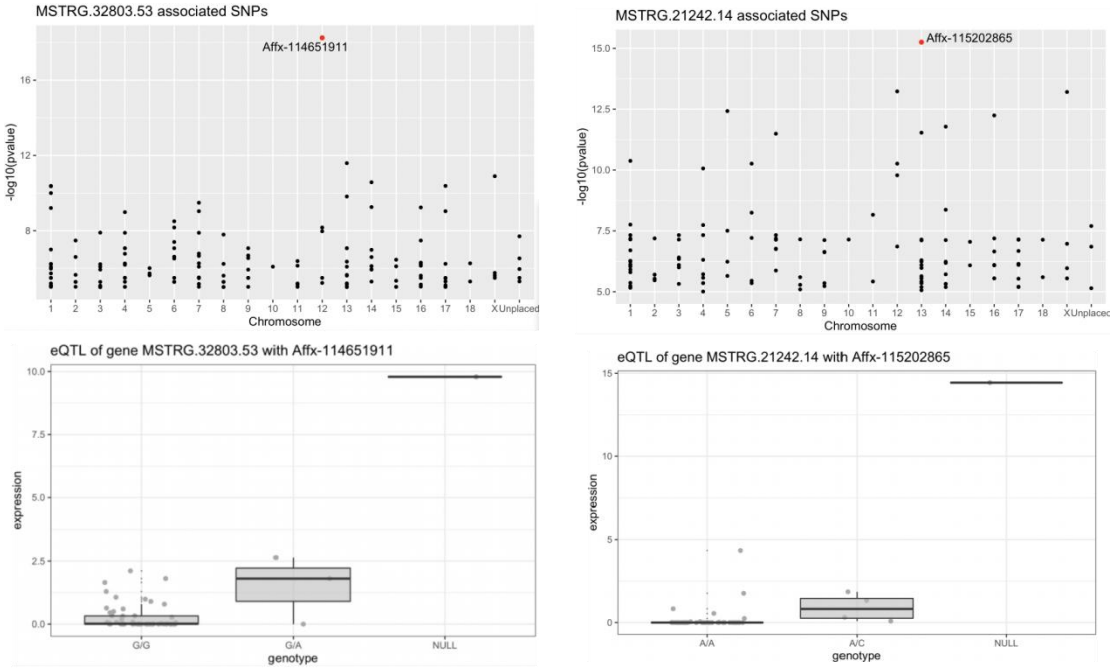



**Figure 3.10** eQTL of transcript MSTRG.32803.53 and MSTRG.21242.14.

The upper panels display the SNPs with FDR<0.05 affiliated to each chromosome, the outstanding SNP was highlighted in red with labeled SNP ID. The lower panels show the expression (FPKM) of the tested transcripts in pigs grouped by their genotypes at the outstanding SNPs. Null indicates the results of genotyping mismatching reference annotation. All the data shown above was derived from 57 pigs used in Chapter 3 with one pig's data not available.

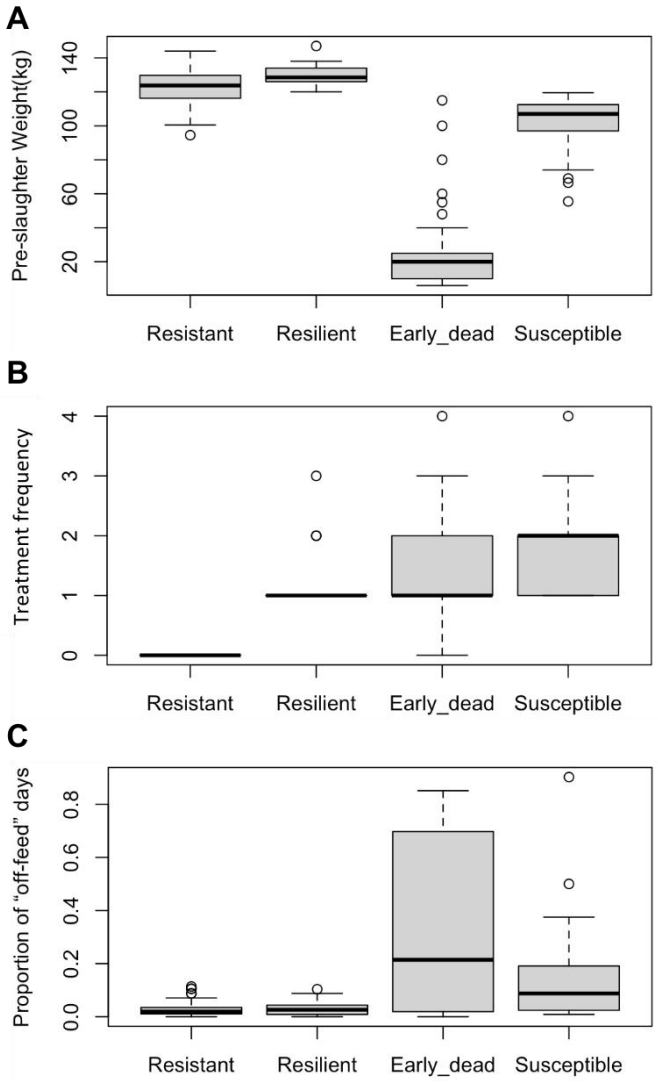



**Figure 3.11** Comparison of the two major grouping parameters (MPSW or MDW in panel A; MTN in panel B) and the "off-feed" proportion (panel C) for the four groups of pigs ("Early_dead", "Resilient", "Resistant" and "Susceptible") in cohort 2.

PCA analysis was performed to explore if there was any global difference in the transcriptomic profile between each group of pigs or each blood collection (Figure 3.12). It is not surprising that the blood transcriptomes of four groups were not well-separated from Blood 1 samples in either gene or transcript level. The post-challenge samples were not able to form unique cluster for any group either (Figure 3.12 A, B). However, more obvious clustering of different blood collections are observed in all the four groups of pigs, particularly in the resilient and resistant group while less clear in the susceptible and Early_dead group (Figure 3.12 C, D). This confirms the efficacy of our natural disease challenge model in inducing significant changes in the transcriptomes of experimental pigs. Moreover, it implies that the Resilient and Resistant group may be more sensitive in responding to pathogen challenge and reflected by gene expression change compared to the Susceptible and Early-dead group. This is consistent with the findings from the cohort 1 pigs suggesting a more potent immune response in the Resilient and Resistant group than the other two.

For DE analysis, the gene or transcript expression was initially compared between one group and the other three groups (Figure 3.13). Surprisingly, no significant gene was observed based on the cohort 1 cutoff (fold change>2, $P<0.05$) for any of the comparisons (Figure 3.13A). The number of DE transcripts was also much fewer than what was shown for cohort 1 (Figure 3.6B and 3.13B). Even though none of the top ten DE transcripts found in cohort 1 shown in Figure 3.6B is recovered in Figure 3.13B, several new interesting transcript hits were identified that were not mentioned above, particularly when comparing the Resilient group with the others. These include some genes involved in protein ubiquitylation (*FBXL18*), signaling transduction (*PIK3C2A*),

transcriptional regulation (*ELK4, CRTC1, RFC1*), ribosome assembly (*RPL32, RPS11*), and cell function and differentiation (*SNCG, SRGN*).

Comparisons between two specific groups did not identify any DE genes except an up-regulated unannotated gene in the Resistant group compared to the Susceptible group (Figure 3.14A). More DE transcripts were found in two-group comparisons (Figure 3.14B) including all the hits displayed in Figure 3.13 B but none overlapped with the top hits from corresponding cohort 1 comparisons (Figure 3.7B). A possible explanation for this inconsistency could be that the genetic and environmental factors may result in variations in the pig cohorts, and most gene expressions are fluctuating except some housekeeping genes in healthy state. Thus, the transcriptome of the cohort pigs contains many "noises" that make it tricky to identify the responsible or correlated genes linked with the post-challenge phenotypes. However, it is still possible that the gene or transcript hits identified from cohort 1 are recovered in cohort 2 but are not ranked as the top hits.

To clarify whether the hits from cohort 1 are "noises" or potential related genes to the post-challenge phenotypes, I generated the up and down-regulated gene lists for all the comparisons (one vs others and one vs one) from cohort 2 and ranked by P-values displayed by heatmaps. Gene or transcript hits from cohort 1 (Figure 3.6 and Figure 3.7) were localized in the corresponding lists if they showed the same changing trend with that from cohort 2 (Figure 3.15 and Figure 3.16). Even though many cohort 1 hits were recovered in cohort 2, most of them were either with high P-value or reverse trend (not displayed). The ones localized close to the low P-value regions (dark color) had larger confidence than those localized far away which could be the "noises" mentioned above.

For example, *SAMD9* was found up-regulated in the Early_dead group compared with the Susceptible group in cohort 1 (Figure 3.7A), and it was also shown with relatively low P-value in the same comparison in cohort 2 (Figure 3.16A), suggesting it could be a real hit. Surprisingly, the *SAMD9*-encoded protein was known as a potent antiviral factor by sensing double-strand nucleic acid and participating in the formation of cytoplasmic granules (J. Liu & McFadden, 2015; Nounamo et al., 2017). This is contradictory to our earlier hypothesis that the pigs from the Susceptible group have stronger immune potential than that from the Early_dead group. Further evidence supporting a stronger immunity of Early_dead pigs is the two up-regulated transcripts *SRGN* and *GIMAP6* shown in Figure 3.14B. *SRGN* encodes a proteoglycan called "Serglycin" which is broadly expressed in immune cells and as reported before, lacking of Serglycin in neutrophils (Niemann et al., 2007), cytotoxic T, and NK cells (Sutton et al., 2016) resulted in a dramatic defect in their target killing, suggesting a crucial role in both innate and adaptive immunity. Increasing numbers of studies uncovered the importance of *GIMAP6* in maintaining immunocompetency. *GIMAP6* is highly expressed in adaptive immune cells and is thought to regulate autophagy and inflammation (Pascall et al., 2018; Shadur et al., 2021; Yao et al., 2022). Based on these findings, it is reasonable to speculate that one possible reason for the death of Early_dead pigs could be an overactive immune response. One more evidence for this speculation is that SAMD9 was believed to act downstream of the proinflammatory TNF-α signaling and contribute to associated inflammation-induced tissue injury (Chefetz et al., 2008).

New DE transcripts resolved from the two-group comparisons include the ones participating in cell metabolism ($CAT$, $MAT2B$, $ACSL1$), signaling transduction ($WDR26$,

*PTPRC, SMG1, GIMAP6*), protein ubiquitylation (UBC), immune response (*IFI44L*), transcription regulation (*STAT6*), and actin binding (*SPTBN1, AVIL*). Even though not as many hits were found to be enriched as in cohort 1 pigs, the DE transcripts enriched from cohort 2 pigs could also provide evidence to infer the differential post-challenge responses of the four groups. For example, compared to the Resistant group, the Resilient group up-regulated a transcript of *CAT* gene which encodes catalase and functions against the oxidative stress by decomposing hydrogen peroxide. It is also notable that this protein was found to promote the growth of many types of immune cells including T and B cells (Takeuchi et al., 1995). Consistently, a *WDR26* transcript was also increased in the Resilient groups, and the protein from this gene also involves in promoting cell growth (Ye et al., 2016) and suppressing the cell stress from hydrogen peroxide (Zhao et al., 2009). This evidence suggests a potential advantage of resilient pigs when they are under oxidative stress due to hydrogen peroxide production during activated immune response to bacteria (Wake Forest University Baptist Medical Center, 2008). Up-regulation of *PTPRC* and *IFI44L* in the Resilient group compared to the Early_dead and Susceptible groups, respectively, further supported the hierarchical baseline immunity model that pigs from the Resilient group outperform the pigs from the other two groups. Based on the fact that no DE gene or transcript related to immune response was identified between the Resilient and Resistant group, and up-regulated immune-related genes in Early_dead pigs than the Susceptible ones from the last paragraph, the baseline immunity model is modified to Resistant≧Resilient>Early_dead>Susceptible.

A *SNCG* transcript was up-regulated in the Resilient group compared to the rest of groups. *SNCG* encodes a protein called γ-Synuclein which is known to inhibit lipid

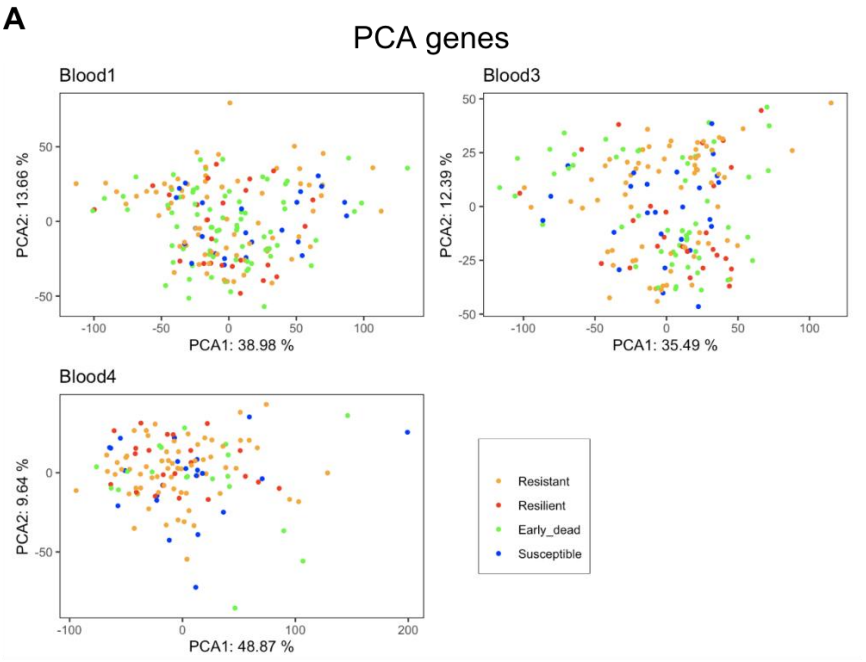
oxidation and energy consumption (Rodríguez-Barrueco et al., 2022). A *CRTC1* transcript, encoded by a gene reported to limit animal appetite and energy consumption, was down-regulated in the Resilient group (Altarejos et al., 2008). These observations match our hypothesis that resilient pigs gain more weight than others may involve consuming less energy.

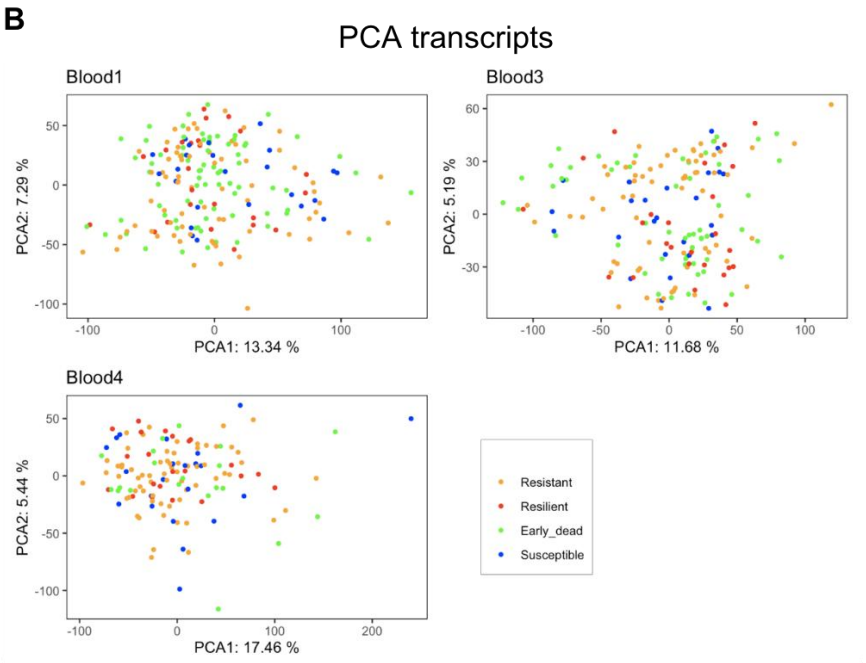
A *PIK3C2A* transcript was specifically down-regulated in the Resilient group. This gene encodes a class II phosphatidylinositol 3-phosphate kinase (PI3K), and intriguingly, was reported to play an important role in promoting bacterial dissemination (*Shigella flexneri*) (Dragoi & Agaisse, 2015). This change indicates a relatively stronger control of bacterial infection in resilient pigs. Reversely in terms of bacterial and viral infection, down-regulation of a *SRGN* transcript in the Susceptible group may reflect their weakness in control and ability to clear those pathogens.

Comparison of two groups for the DE transcripts (Figure 3.14B) did help solidify and improve the hypothetical baseline immunity model: Resistant≧Resilient>Early_dead>Susceptible. It seemed controversial that *STAT6* was more highly expressed in the Susceptible group than that in the Resilient group, as STAT6 is known to induce the transcriptional regulation downstream of IL-4/IL-13 during a type 2 response (Walford & Doherty, 2013). In addition, it was reported that STAT6 activated by STING independent of Janus kinases plays an important role in innate antiviral response (H. Chen et al., 2011). However, in the context of our poly-pathogen natural disease challenging model, this may not be the case that pigs with higher *STAT6* expression in blood cells would perform better in controlling the infections. First, the STING study only used two model viruses (a RNA virus-Sendai virus, and a

DNA virus-Herpes simplex virus 1). Second, other studies demonstrated a negative role of STAT6 in susceptibility to viruses and parasites (Dent et al., 1999; Mahalingam et al., 2001; Tekkanat et al., 2001). Thus, it is necessary to be careful to discuss the meaning of immune-related genes in terms of their role in pathogenesis especially in the context of multiple pathogens. Another two interesting hits *IFI44L* and *PTPRC* up-regulated in the Resilient group (Figure 3.14B) imply that both innate and adaptive immune potential of resilient pigs may be superior. *IFI44L* is a Type I interferon-stimulated gene and considered as a critical factor in restricting viral and bacterial infection (Busse et al., 2020; Jiang et al., 2021). While *PTPRC* encodes the surface CD45 molecule known as a marker of almost all hematopoietic cells, and its expression can not only lower the threshold for antigen receptor activation of T and B cells, but also regulates the antigen-sensing signaling in mast cells and dendritic cells (Saunders & Johnson, 2010). The higher expression of those two molecules may result in a "primed" immune defending status in resilient pigs making them faster to respond and take less effort to induce a certain level of immune response to control the invading pathogens.

By using a larger cohort size (cohort 2), the observations from cohort 1 were partially supported, but no DE genes or gene sets were validated to be used as potential markers to differentiate post-challenge pig response in our natural disease challenge model. This indicates that there was not enough difference in the whole blood transcriptome among the four groups of pigs before pathogen challenge and appearance of their defined phenotypes ("Resistant" or "Resilient" or "Early_dead" or "Susceptible"). However, our efforts did provide hints to unveil the different baseline levels for their immune defense potential or metabolic activity of the whole blood cells. Investigation of


A
PCA genes
Blood1
PCA2: 13.66 %
PCA1: 38.98 %
Blood3
PCA2: 12.39 %
PCA1: 35.49 %
Blood4
PCA2: 9.64 %
PCA1: 48.87 %
Resistant
Resilient
Early_dead
Susceptible


B
PCA transcripts
Blood1
PCA2: 7.29 %
PCA1: 13.34 %
Blood3
PCA2: 5.19 %
PCA1: 11.68 %
Blood4
PCA2: 5.44 %
PCA1: 17.46 %
Resistant
Resilient
Early_dead
Susceptible

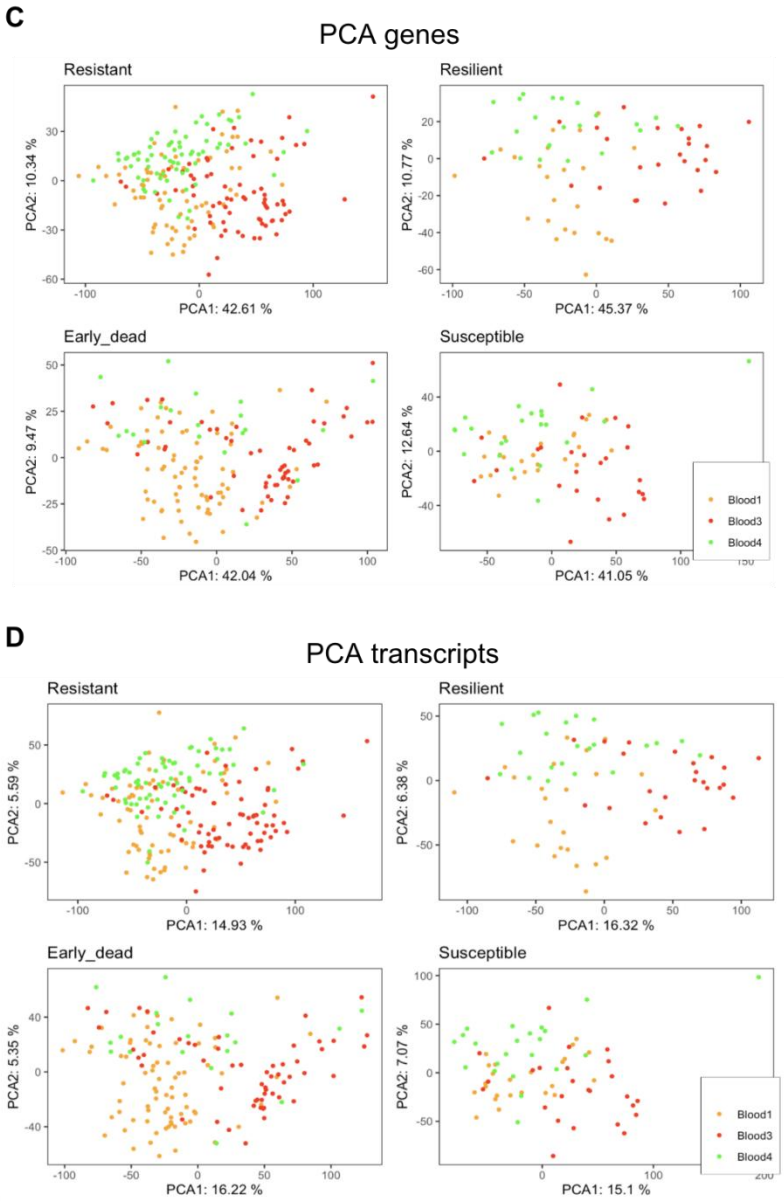



**Figure 3.12** PCA of transcriptomic profiles for pigs from each blood collection by displaying the four groups in each plot (A, B) and for pigs from each group by displaying the three blood collections in each plot (C, D).

Panel A/C and panel B/D are based on the gene or transcripts level, respectively. Each dot represents a pig. The different colors indicate pigs corresponding to the four defined pig groups (A, B) or three blood collections (C, D).

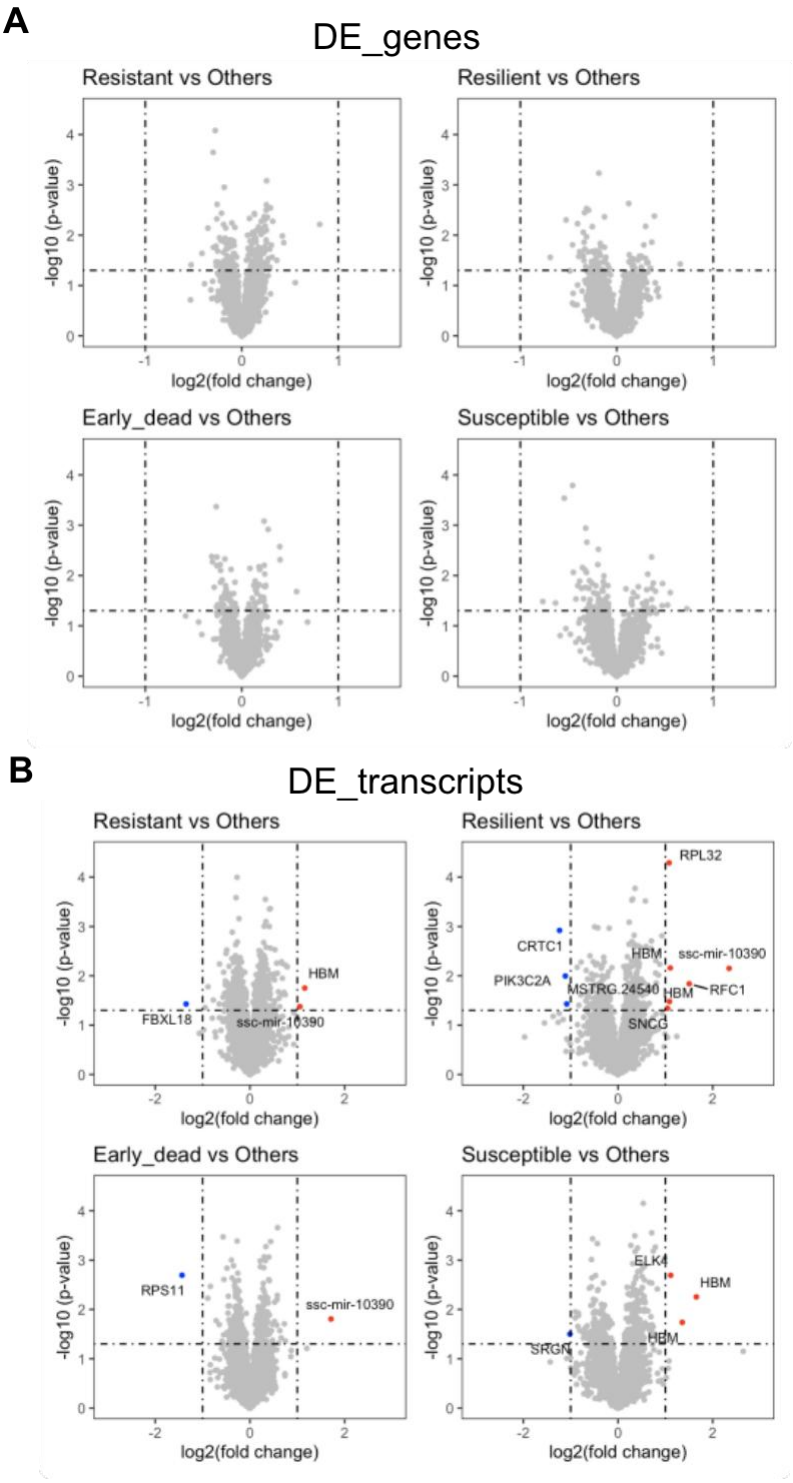



**Figure 3.13** Differential expression of genes (A) and transcripts (B) level comparing one group with the rest of the groups in cohort 2 pre-challenge samples.

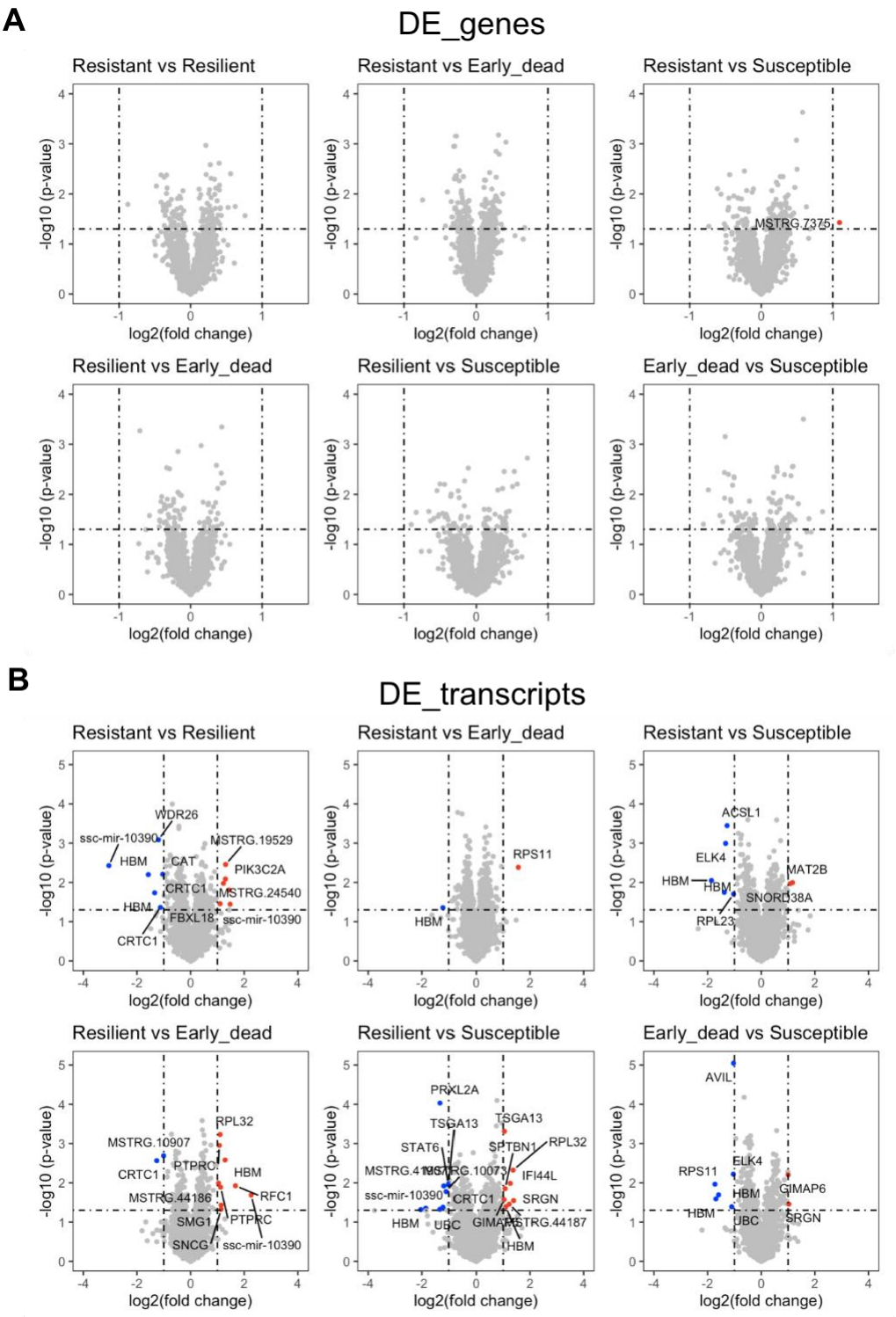



**Figure 3.14** Differential expression of genes (A) and transcripts (B) level comparing two specific groups in cohort 2 pre-challenge samples.

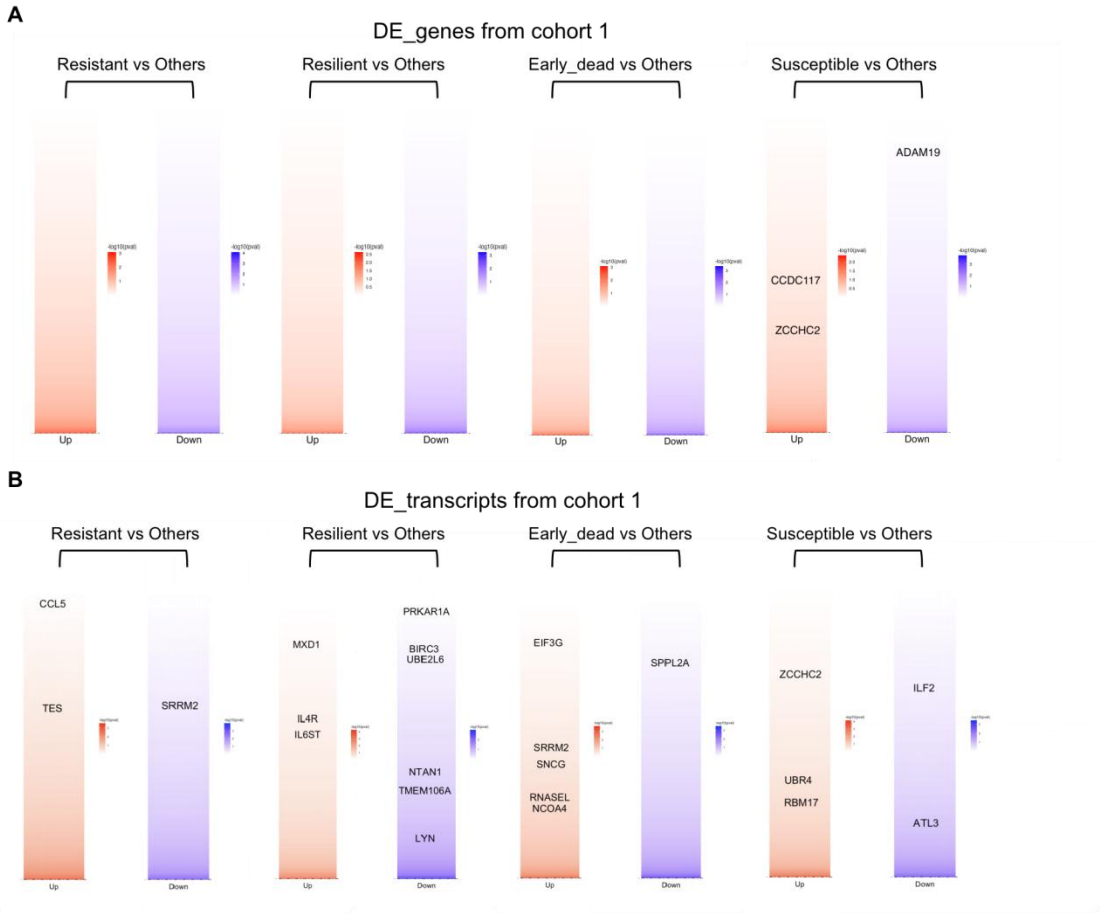
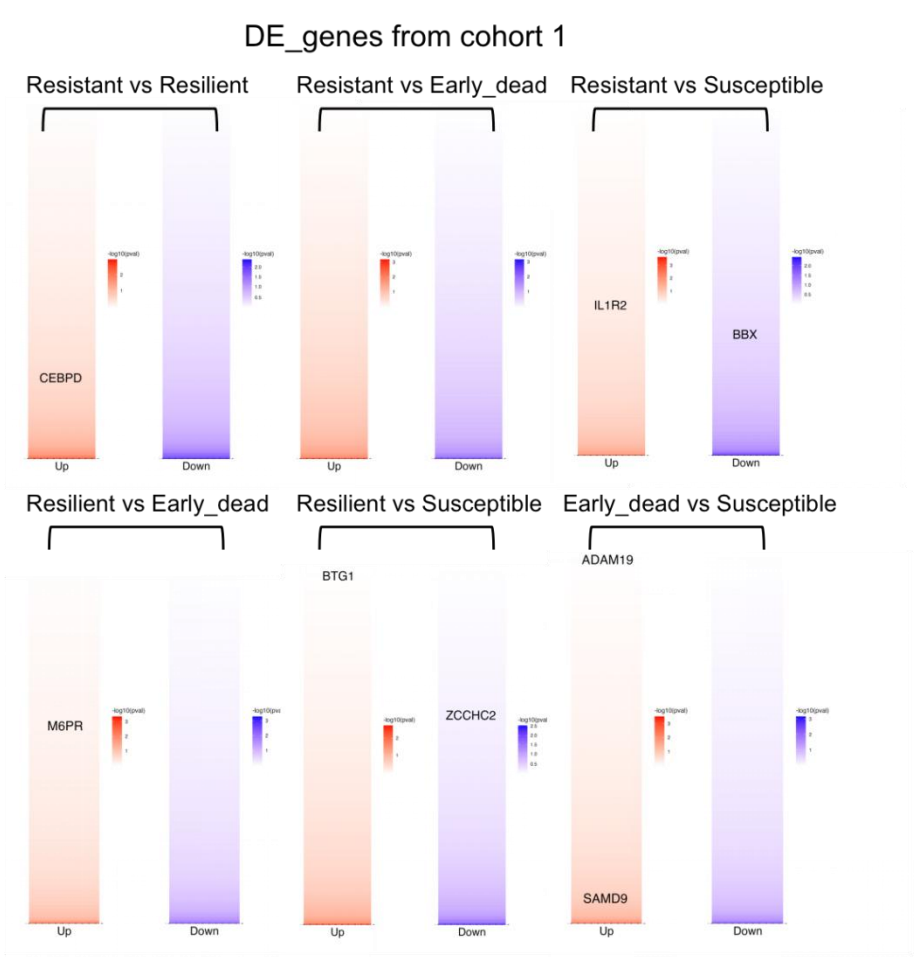



**Figure 3.15** Mapping of top DE_genes (panel A) or DE_transcripts (panel B) in cohort 1 to the list of genes or transcripts ranked by P-value of the corresponding group comparisons (one vs the rest of three groups) from cohort 2.

Red or purple color indicate the up-regulated or down-regulated genes/transcripts, respectively, in the group before "vs" than the group after.

**A**


DE_genes from cohort 1
Resistant vs Resilient
CEBPD
Up
Down
Resistant vs Early_dead
Up
Down
Resistant vs Susceptible
IL1R2
BBX
Up
Down
Resilient vs Early_dead
M6PR
Up
Down
Resilient vs Susceptible
BTG1
ZCCHC2
Up
Down
Early_dead vs Susceptible
ADAM19
SAMD9
Up
Down

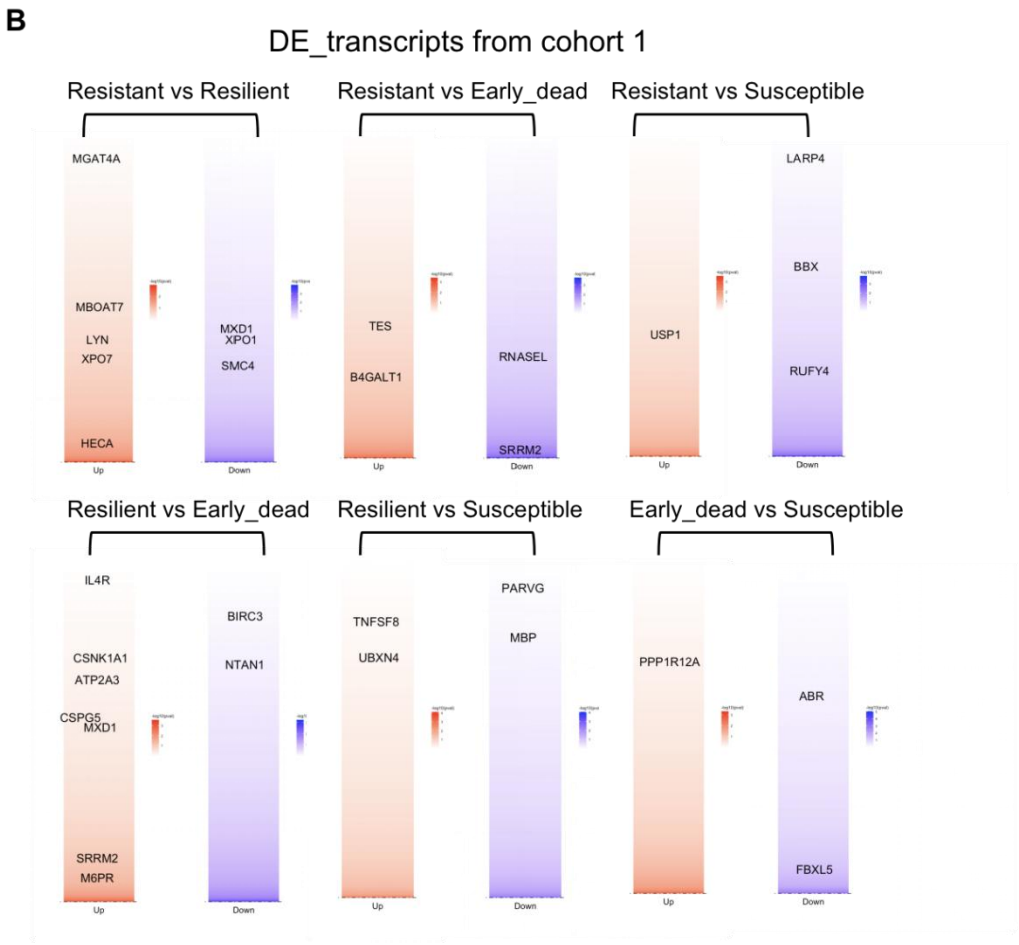



**Figure 3.16** Mapping of top DE_genes (panel A) or DE_transcripts (panel B) in cohort 1 to the list of genes or transcripts ranked by P-value of the corresponding group comparisons (one vs one group) from cohort 2.

Red or purple color indicate the up-regulated or down-regulated genes/transcripts, respectively, in the group before "vs" than the group after.

how the blood transcriptome is responding to disease challenges in different groups of pigs could further help to understand how the mobilization of the immune system and cell metabolism correlates with the pig's health and productive performance.

### 3.3.8 DE gene analysis comparing different blood sample collections in cohort 2.

The major goal of this chapter was to test whether the pre-challenge whole blood transcriptome (Blood 1) could be used to differentiate the defined four pig phenotypes that appeared after disease challenge. The subtle differences shown above did not strongly support this idea, but endogenous differences in terms of immunity and cell metabolism appeared to be exhibited in the blood transcriptome of different groups of pigs even before contacting pathogens. To further understand these biological differences and potential molecular mechanisms that may result in the differential post-challenge phenotypes, I processed the transcriptome data of the post-challenge blood samples from cohort 2 pigs. To better reflect the dynamics of gene expression change and dissect the activating and resolving of immune response, I included both Blood 3 and Blood 4 samples to represent the early and late phase of disease challenge.

Given that the dramatic change of transcriptome was observed in the PCA analysis among Blood 1, 3, 4 shown in Figure 3.12 C and D, I first performed DE analysis for Blood 3 vs Blood 1 and Blood 4 vs Blood 3 (Figure 3.17). In terms of common responses reflected by gene expression change during early infection, all the four groups up-regulated *CXCL10* which encodes a pro-inflammatory chemokine in Blood 3 compared to Blood 1. CXCL10 can be induced by viral and bacterial infections and stimulate the recruitment of activated immune cells to the infection site (Elemam et al., 2022). This

confirms that our natural disease challenge model efficiently spreads active infections to the tested pigs and activates their immune system. Meanwhile, Blood 3 has higher expression of *VWF* gene which is known to promote hemostasis (Kanaji et al., 2012) and thus indicating tissue damage happened early after infection. Down-regulation of a gene (*IL1R2*) encoding the decoy receptor for the infection-induced pro-inflammatory cytokine IL-1β (Dinarello, 2018) is also consistent with the ongoing active defensive immune response. All the groups except the Early_dead group up-regulated *CXCL10* as a common feature but each group displayed some specific top hits that may result in their phenotype differences. For example, the Resistant group up-regulated additional immune effector genes including *GZMA* and *GBP5*. This supported a strong immune response was induced in the Resistant group and excludes the possibility that they were totally tolerant to the infections. Up-regulation of *CAMK1* gene encoding a calcium-dependent kinase indicates the immune cells in resistant pigs were highly active to calcium-induced metabolic activities. Other than additional immune defensive genes including *FCGR1A* and *DDX60* were increased in Blood 3 of the Resilient pigs, what needs to be noted are the following several genes. *IOD1* plays a crucial role in preventing autoimmunity from the overstimulated immune response (van Baren & Van den Eynde, 2015). *IL27* not only participates in inducing Th1 cells but also suppresses some pro-inflammatory cytokines including IL-6 known as a major driver of cytokine storm (Jones et al., 2018). Cytokine storm is a systemic disorder of immune dysregulation characterized by elevated circulating cytokine levels, acute systemic inflammatory symptoms, and multiorgan dysfunction with the potential to lead to multiorgan failure (Fajgenbaum and June, 2020). *MANF* plays a role in alleviating endoplasmic reticulum stress and cell toxicity from

hypoxia (Bai et al., 2018). The up-regulation of the above three genes in the Resilient group provides examples indicating their better performance in immune tolerance maintenance and anti-stress response. In contrast, the Early_dead group up-regulated the *CCL5* gene which encodes a chemokine in innate immune response to infection but also a well-known factor inducing cytokine storm (Agresti et al., 2021; Patterson et al., 2020; Teijaro et al., 2014). The Susceptible group also efficiently up-regulated some immune-sensing and pro-inflammatory genes including *CSF1* and *UNC93B,* but no cytotoxicity-related genes were found that were increased in the other three groups (*GZMA, NKG7, FCGR1A*). The gene expression changes during early infection particularly highlighted the opposite performance in maintaining a well-controlled immune response in the Resilient and Early_dead group pigs. The overwhelming activation of pro-inflammatory cytokine cascade might be an important reason to cause the early death of the Early_dead group pigs.

Comparison of Blood 4 vs Blood 3 revealed the gene expression change from early infection to a later point of continuous disease challenge (Figure 3.17). All the four groups down-regulated several innate effector genes including *CXCL10*, *LTF*, and *SERPING1* as the top hits, implying a transition already happened from innate to adaptive immune response. Of note, the Resistant group up-regulated *MS4A1* and *SCIMP*, two genes involved in B cell activation, differentiation and signaling (Tedder et al., 1985), indicating an ongoing humoral immune response. The expression of these two genes were also found increased in the Resilient group, however, not in the Early_dead and Susceptible groups. Additionally, the Resilient group up-regulated two more genes (*CD72* and *CD79b*) which also play a role in B cell differentiation and signaling

transduction. Interestingly, the Resilient group had increased expression for *NPRL3* and *FBXO9*, both function to restrain mTOR1 signaling which is crucial in cell proliferation and metabolism (Fernández-Sáiz et al., 2013; Iffland et al., 2022). Moreover, *SGK1*, a gene encoding a kinase playing a role in promoting cell survival under cell stress (Ghani, 2022), was also up-regulated in the Resilient group. In comparison, the Resistant group down-regulated *NRPL3* in Blood 3 vs Blood1. These facts suggest the different strategies of organizing immune defense in pigs from the Resistant and Resilient group that the former boosted the activity of immune cells quickly after infection while the later one utilized a more economic way to maintain a reasonable but long lasting immune response.

### 3.3.9 Pathway enrichment analysis for the DE genes from comparing different blood collections

The DE genes identified between blood collections revealed activation of the innate immune system in all the groups of pigs early after disease challenge and then the contraction of innate immune response when the infection lasted longer. Although some key genes listed above could provide examples of how the four groups respond differentially in our natural disease challenge mode during the early and late phase of infection, a more integrated analysis was needed to clarify the biological meanings of those gene expression changes. To this end, I did GO pathway enrichment analysis on the DE genes shown in Figure 3.17. GO analysis was not able to enrich pathways for specific comparison, such as Blood 3 vs Blood 1 for all the groups and the Resistant group due to insufficient number of DE genes. As shown in Figure 3.18, both the Resilient and Susceptible group enriched several pathways associated with innate immune response in Blood 3 compared to Blood 1 (e.g. defense response to other organisms, innate immune

response, response to external biotic stimulus), suggesting those two groups of sick pigs were more sensitive than the Resistant group in inducing innate defensive response in our model. All the groups down-regulated those innate immune pathways in Blood 4, again suggesting an ongoing contraction phase of innate immunity as inferred above from the DE analysis. Notably, the Resilience group was the only group that down-regulated many cell proliferation and metabolic-related pathways (e.g. cell population proliferation, protein metabolic process, regulation of transport) in Blood 4 compared to Blood 3. This is also consistent with the observations from DE analysis and further supports the energy saving mode of resilient pigs during the immune response.

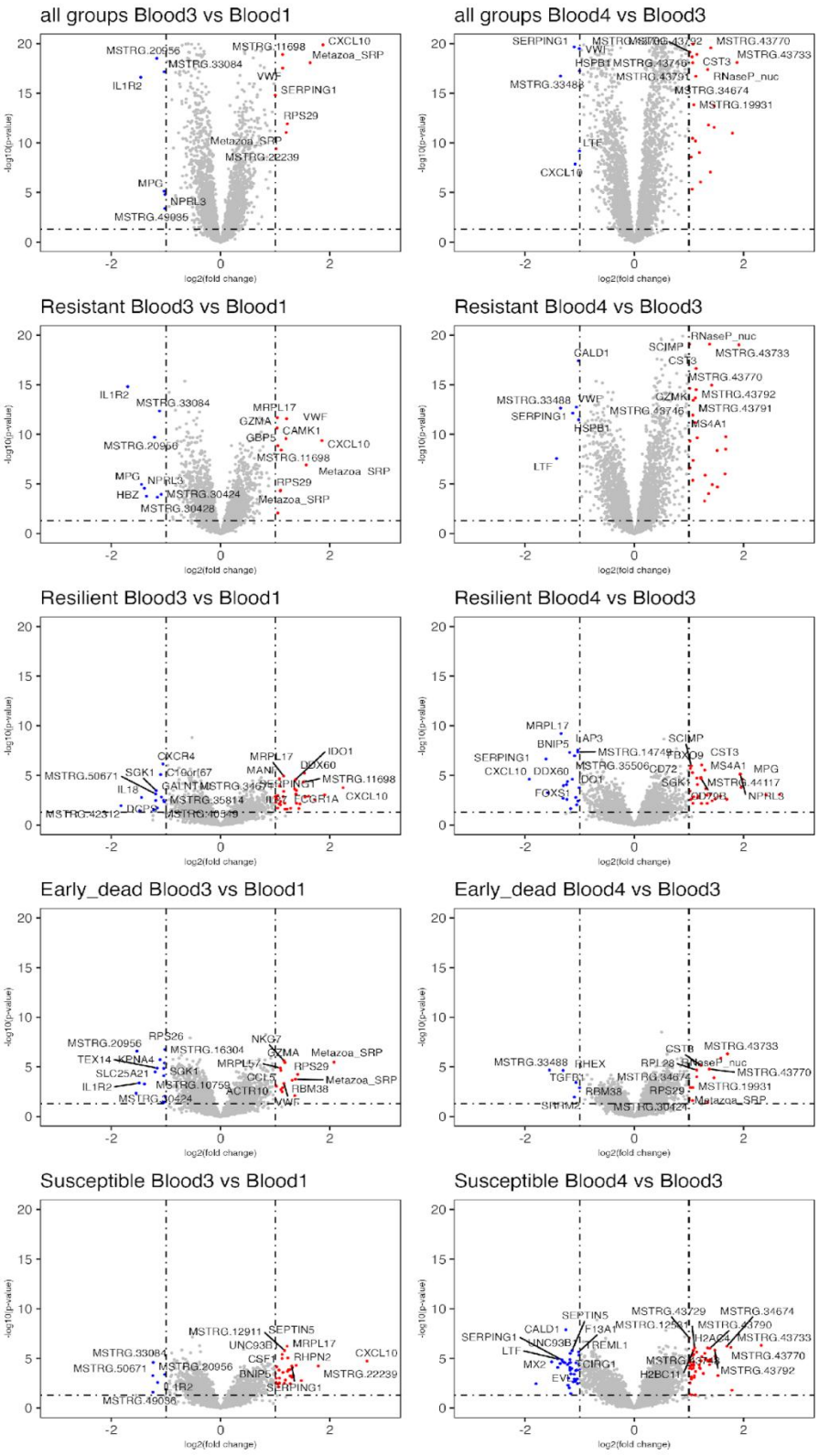



**Figure 3.17** DE genes between two blood collections in cohort 2 in all the pigs or a specific group.

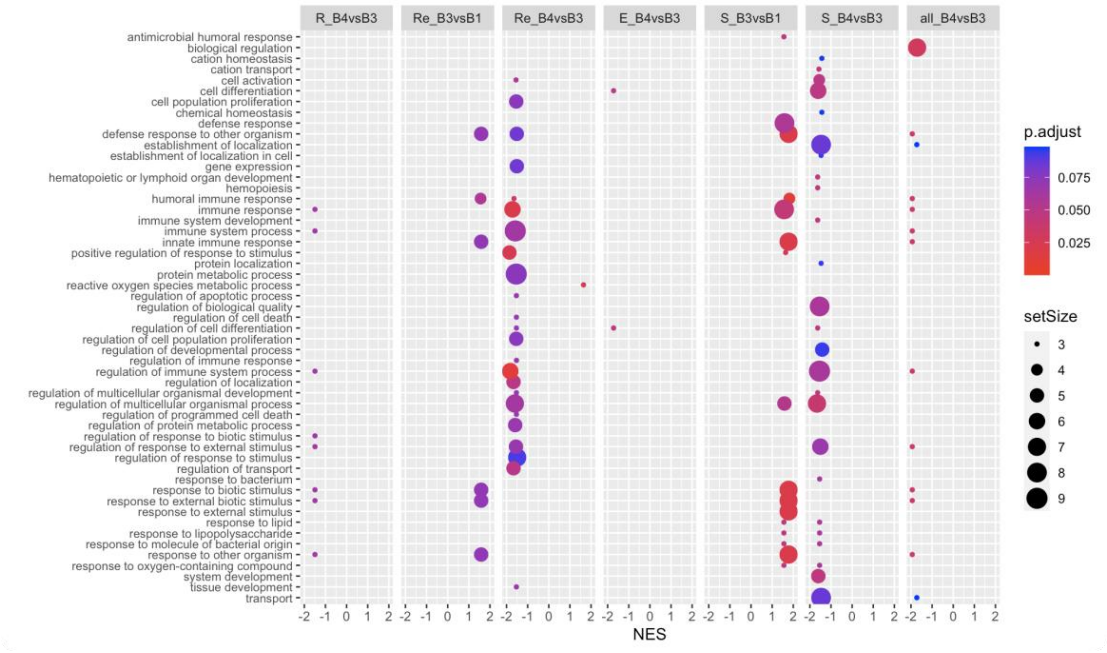



**Figure 3.18** GO term enrichment analysis for DE genes from Blood 3 vs Blood 1 (B3vsB1) and Blood 4 vs Blood 3 (B4vsB3) in each group (Resistant-R, Resilient-Re, Early_dead-E, Susceptible-S, all groups-all).

## 3.4 Discussion

### 3.4.1 Insights for differential pig disease responses from the transcriptome of blood cells before challenging

The immune cells of pigs may differ endogenously and dictate the response once pigs have become infected. I thus hypothesized that pig disease response could be predicted using the transcriptome profiles of whole blood cells prior to the pathogen challenge. Even at steady state, the innate immune cells are critical components to form the first immune defensive barrier and express multiple pattern recognition receptors (PRRs) for immune surveillance (Kumar et al., 2011). The response of innate immune cells also plays an important role in the generation of the adaptive immune response (Iwasaki & Medzhitov, 2010). To gain an insight into whether the gene expression in immune cells at the homeostatic state reflected the pigs' better control of the infectious disease, I first did RNA-Seq on the whole blood cells from four groups of pigs (58 in total) prior to the pathogen challenge. The grouping was determined based on their differential response in our natural disease challenge model determined retrospectively and used to select the samples.

Many effector genes of immune cells are largely induced upon activation after recognizing antigens. Not surprisingly, PCA analysis was not able to differentiate those four groups of pigs based on the global features of the transcriptome at their steady state (Figure 3.7). However, this does not exclude the possibility that immune cells from a different group of pigs are endogenously similar in their functional potential without the pathogen challenge. In contrast, the immune cells from different pigs might already

exhibit differences in the epigenetic level, e.g. in their genome accessibility. This idea is supported by a mouse CD8 T cell study that the researchers cannot differentiate the WT group and *IL2RA* gene knockout group in the PCA analysis of RNA-Seq data whereas the PCA on ATAC-Seq clearly separated each group (Chin et al., 2022).

Only a few group-specific genes were found in the DE genes analysis when one group was compared to the other three. This finding suggests the difficulty in identifying well-defined gene markers for each specific group but also might be due to the limited size of the experiment cohort (58 pigs in total). The DE analysis on transcripts did enrich many hits specifically for each group, including transcripts encoded by immune-related genes, e.g. higher expression of chemokine gene *CCL5* in the Resistant group, and higher expression of cytokine signaling genes *IL4R* and *IL6ST* in the Resilient groups. CCL5 is able to recruit leukocytes including T cells, NK cells, monocytes, dendritic cells, eosinophils, and basophils to the site of inflammation (Appay & Rowland-Jones, 2001). Coordinated with IL-2 and interferon (IFN) gamma, CCL5 can promote an activation and expansion of NK cells, which is crucial for innate immune cells to control viral infection (Maghazachi et al., 1996). Furthermore, through binding to its receptor CCR5, CCL5 plays a protective role in the induction of anti-apoptotic signals to macrophages during viral infection, which maintains the survival of those scavengers to clear the infected cells (Tyner et al., 2005). IL-4 plays a key role in the induction of Th2 response which is important to induce specific humoral response against extracellular pathogens e.g. helminths (Nakayama et al., 2017). Thus, T cells with higher *IL4R* gene expression are expected to be more sensitive to IL-4 stimulation. IL-6 is a pleiotropic cytokine that works on both innate and adaptive immune responses to protect the host from the

infection of a variety of pathogens (Rose-John et al., 2017). IL-6 also contributes to the cytokine storm when its production is exaggerated in response to infection (Tanaka et al., 2016). Notably, IL-6 is a mediator to induce the synthesis of APPs (Heinrich et al., 1990). Up-regulation of *IL6ST* which encodes a transmembrane protein as the subunit of the IL-6 receptor (Skiniotis et al., 2005) could probably sensitize the response to IL-6 and initiation of APR. This up-regulation may endow a more potent immune response in the resilient pigs and APR; however, I did not observe a stronger induction of the three APPs illustrated in Chapter 2 in the Resilient group. Even though some other DE transcripts might not be involved in the immune response directly, their immune-regulatory functions can also contribute dramatically to the control of infection. For example, the Resilient group specifically down-regulated the *UBE2L6* gene which encodes a ubiquitin ligase and was found to inhibit virus-induced IFN beta production (L. Huang et al., 2020), suggesting a lower threshold for type I IFN induction upon virus infection. The Resistant group up-regulated a transcription factor gene *ATOH8* which is epigenetically regulated and a target of Blimp-1, and was recently uncovered to be important in plasma cell functions (Kong et al., 2022; Minnich et al., 2016). Besides the genes directly involved in immune response or immune regulation discussed above, one transcript encoded by a gene called *ILF2* drew my attention and was down-regulated specifically in the Susceptible groups (Figure 3.8B). ILF2 has been found to interact with two PRRSV viral proteins nsp9 and nsp2 which are important for viral replication, and negatively regulate PRRSV replication (Wen et al., 2017). Given that PRRSV was included in our natural disease model, it is not surprising that the susceptible pigs had lower expression of this antiviral protein. Moreover, ILF2 was also documented to inhibit the replication of

infectious bursal disease virus and Japanese encephalitis virus (Cui et al., 2019; Stricker et al., 2010).

Among all the DE transcripts, only two, MSTRG.32803.53 and MSTRG.21242.14, were found to be up-regulated specifically in the Susceptible group with q-value<0.1. To gain more information about the potential function of these two transcripts, I found that the first exon of *MSTRG.32803.53* matched the *Ig C* gene which encodes a constant chain of immunoglobulin. Higher expression of antibodies (e.g. an auto-reactive antibody or natural antibody in the healthy host) may not indicate a better immunity. However, early studies proposed a plausibly defensive role of auto-reactive antibody in response to pathogens because of its cross-reactivity with foreign antigens; but some evidence also supported their anti-inflammatory properties (Lacroix-Desmazes et al., 1998). A recent study using this model also supported a positive correlation between the level of natural antibodies and pig resilience to polymicrobial disease (Y. Chen et al., 2020). A putative reason for the controversial role of auto-reactive antibodies could be that the traits used to define resilience in other studies were the higher survival rate and fewer treatments. Both traits were not the focus of my "Resilient" group. Six out of nine exons of *MSTRG.21242.14* matched the *SLAM9* gene which is a SLAM family receptor expressed on the surface of leukocytes. Studies in mice demonstrated that SLAM9 was expressed on phagocytes and dendritic cells, and played a role in the recognition and clearance of bacteria (Wilson et al., 2020). However, the function of this molecule is largely unknown since SLAM has only been discovered recently and its ligand has not yet been identified. In addition, SLAMF9 could suppress the migration of macrophages presumably due to its increased intercellular adhesion (Dollt et al., 2018). The function of SLAMF9 should be

verified in pigs in the context of pathogen challenge, and this could provide hints on whether a higher expression of SLAMF9 is beneficial to pigs in the immune defense against pathogen invasion. No matter what the actual function of these two transcripts in pigs, their expression was exclusively high in the Susceptible group (Figure 3.11). eQTL analysis for associated SNPs with the expression of these two transcripts led to two outstanding candidate SNPs and the rare genotypes were correlated with higher expression for either MSTRG.32803.53 or MSTRG.21242.14. The loci of the two SNPs are either in gene introns or intergenic regions distant from the gene locus, indicating an indirect association with the expression of MSTRG.32803.53 and MSTRG.21242.14.

Nevertheless, I observed more DE genes when comparing two specific groups. Several key genes including *CEBPD* and *ATRN* were reported to have positive immune regulation functions. These genes were identified for the Resistant group, which is consistent with why those pigs outperformed the others in our natural disease challenge model with regards to the control of disease development. Interestingly, *CEBPD* was found to play a role in regulating the expression of APPs in the liver including AGP, CRP, and HP (Cantwell et al., 1998). I did observe a significantly higher concentration of AGP and a trend of higher HP in resistant pigs than that in resilient pigs under an unchallenged condition (Figure 2.2). This effect however did not become more dramatic when the pigs were under pathogen challenge. Another up-regulated gene *ATRN* in the Resistant group (compared to the Early_dead and Susceptible group) was reported in an earlier study, and in which its gene polymorphism was associated with pig ADG and live weight traits (Kim et al., 2005). This is particularly interesting as, in Chapter 3, the selected pigs from the Resistant group had much higher pre-slaughter weight than the ones from the

Early_dead and Susceptible group (Figure 3.3). This is consistent with the earlier finding of the association of *ATRN* gene with pig growth. In addition, a negative immune regulator *IL1R2* was found more up-regulated in the Resistant group than in the Susceptible group (Figure 3.9A). IL1R2 was a decoy receptor for IL-1 cytokines and thus suppresses IL-1 signaling and its proinflammatory function (Schlüter et al., 2018). Although IL-1 is one of the main pro-inflammatory cytokines during viral infection and its signaling is protective to the host, dis-regulated IL-1 production may cause immunopathology and even cytokine storm syndrome, which can be lethal to the host (Conti et al., 2020; Q. Liu et al., 2016). Thus, a higher expression of a negative immune regulator such as IL1R2 might be a protective mechanism of resistant pigs to prevent excessive inflammatory response and not counteract with their superior immunity against pathogens. In contrast, this could be a great example of how pigs successfully survived from infectious diseases, and how well the pigs can balance between the pathogen clearance and inflammation intensity may be more important in actual commercial farms under the risk of complex pathogen challenge.

To gain more functional insight into the DE genes and transcripts, and better understand the biological pathways involved, GO enrichment analysis was performed. Even though the enriched pathways specific for each group (1 group vs 3) did not reveal the relative strength of immunity, multiple cytoskeleton organization pathways were down-regulated in the Resilient group compared to the others (Figure 3.8). Those down-regulated genes (*TPM3*, *ARPC2*, *DMTN*, *GMFG*, *PLEK*) in cytoskeleton organization pathways participate in cell protrusions, motility, and migration. For example, *ARPC2* encodes a component protein of the Arp2/3 complex which mediates polymerization of

actin upon stimulation that provides the force for cell motility (Welch et al., 1997). Additionally, Arp2/3 complex is required for proper function of integrin in macrophages and is critical for complement receptor 3-mediated phagocytosis (Rotty et al., 2017). The down-regulated gene *GMFG* also interacts with Arp2/3 complex and regulates its debranching (Boczkowska et al., 2013). Another down-regulated gene *DMTN* encoding a protein called "Dematin" is implicated in the formation of cell protrusions and platelet motility (Mohseni and Chishti, 2008). This finding is particularly interesting as cytoskeleton organization plays pivotal roles in immune cell activities and is also energetically demanding. especially as immune cell migration is largely dependent on cytoskeletal remodelling and requiring sufficient metabolic activity (Guak and Krawczyk, 2020). Glycolysis play a pivotal role in the motility of both immune and non-immune cells and many glycolytic enzymes were found localized at the cytoskeleton (Guak and Krawczyk, 2020). Therefore, glucose consumption from immune cell migration may have a competition with energy accumulation in terms of livestock productivity purposes. A study in mice revealed that there is an energetic trade-off between active host defense against pathogens and homeothermy resulting in hypometabolism and hypothermia (Ganeshan et al., 2019), suggesting that energy consumption from immune defense of pigs may also interfere with productive metabolism. Thus, it is suggested that less energy consumption occurs in the immune response in resilient pigs which may spare energy for use in productive metabolism. This is consistent with the higher pre-slaughter weights of resilient pigs than those of susceptible pigs (Figure 3.3). Another intriguing point is that the Susceptible group had suppressed many cell metabolism-related pathways. On one hand this may result in inadequate energy production in the immune cells when they are

activated for proliferation and effector functions, but on the other hand it might also induce immune exhaustion (Delgoffe & Powell, 2015). Combining the enriched pathways from the DE genes and transcripts, a hypothetical model depicting the immune defense potential to complex pathogen challenge was Resistant>Resilient>Susceptible>Early_dead.

Most of the findings discussed above were derived from a small cohort of 58 pigs. Thus, I used a larger cohort of 209 pigs to validate these findings. Even though the DE genes and transcripts found in cohort 1 were not found as top hits in cohort 2, most of the newly identified DE hits were reasonable and basically matched the idea that the Resistant and Resilient group have stronger immunity than the Susceptible and Early_dead group when they are healthy. The DE transcripts (Figure 3.14B) particularly for the comparisons of two specific groups did help solidify and improve the hypothetical baseline immunity model as Resistant≧Resilient>Early_dead>Susceptible. It seemed unexpected that *STAT6* was more highly expressed in the Susceptible group than that in the Resilient group, as STAT6 is known to induce the transcriptional regulation downstream of *IL-4*/*IL-13* during a type 2 response (Walford & Doherty, 2013). In addition, it was reported that STAT6 activated by STING independent of Janus kinases plays an important role in innate antiviral response (H. Chen et al., 2011). However, in the context of our poly-pathogen natural disease challenging model, this may not be the case that pigs with higher *STAT6* expression in blood cells would perform better in controlling the infections. First, the STING study only used two model viruses (a RNA virus-Sendai virus, and a DNA virus-Herpes simplex virus 1). Second, other studies demonstrated a negative role of STAT6 in susceptibility to viruses and parasites (Dent et

al., 1999; Mahalingam et al., 2001; Tekkanat et al., 2001). Thus, we could not conclude that the Susceptible group with higher *STAT6* expression had stronger immunity than the Resilience group. It should be careful to discuss the meaning of immune-related genes in terms of their role in pathogenesis, especially in the context of multiple pathogens. Another two interesting hits *IFI44L* and *PTPRC* up-regulated in the Resilient group (Figure 3.14B) imply that the Resilient group is superior for both innate and adaptive immune potential. *IFI44L* is a Type I interferon-stimulated gene and considered as a critical factor in restricting viral and bacterial infection (Busse et al., 2020; Jiang et al., 2021). While *PTPRC* encodes the surface CD45 molecule known as a marker of almost all hematopoietic cells, and its expression can not only lower the threshold for antigen receptor activation of T and B cells, but also regulates the antigen-sensing signaling in mast cells and dendritic cells (Saunders & Johnson, 2010). The higher expression of those two molecules may result in a "primed" immune defending status in resilient pigs making them faster and take less effort to induce a certain level of immune response to control the invading pathogens.

Besides the immune-related genes from cohort 2 discussed above, some genes associated with cell metabolism were also notable. *SNCG* transcript encoding γ-Synuclein was found with higher expression in the Resilient group compared to the rest of pigs. γ-Synuclein was known as microtubule-associated protein and regulating cytoskeleton rigidity (Zhang, Kouadio, et al., 2011). γ-Synuclein could support the differentiation and expansion of fat tissue through inhibiting lipid oxidation and energy consumption (Millership et al., 2013; Rodríguez-Barrueco et al., 2022). Interestingly, γ-Synuclein was also reported to promote cell survival under ER stress, which is consistent

with the role of *MANF* gene that was up-regulated early after disease challenge in the Resilient group (Figure 3.16). This matches our speculation for the energy-saving mode of the baseline immunity in resilient pigs. Consuming less energy from immune cells in resilient pigs may contribute to their higher weight gain than others. Down-regulation of a transcript from *CRTC1* gene in the Resilient group may further support this idea. *CRTC1* encodes a transcriptional co-activator activated by cyclic AMP (cAMP) and calcium signals, and mice deficient for this gene showed increased appetite but less energy expending at young age (Altarejos et al., 2008). However, *CRTC1* is known to be expressed in the brain, more research is necessary to reveal whether its expression in the haematopoietic system matters in maintaining energy balance.

The transcriptome study did come up with many interesting gene candidates that may be involved in forming the unique disease resilience phenotype upon pathogen challenge. Further studies such as loss-of-function experiments may reveal the important role of those genes in an actual immune response induced by infections.

### 3.4.2 Implications of pig immune response and pathogenesis from DE genes between pre- and post-challenge transcriptome

To understand how the whole blood transcriptome is modulated in our natural disease challenge model, I did DE analysis comparing samples collected before (Blood 1) and early after (Blood 3) disease challenge. It should be highlighted that all the pigs tested up-regulated *CXCL10* gene in Blood 3. This specific pro-inflammatory chemokine gene was also reported previously by other researchers in different infection models. These models include not only the pathogens that were verified in our experiment (i.e.

PRRSV (Xiao et al., 2010), PCV2 (Zhang, Lunney, et al., 2011), Mhyo (Ni et al., 2019)), but also some other common pig pathogens such as African swine fever virus (Fishbourne et al., 2013) and *Mycobacterium tuberculosis* (Rawat et al., 2013). Intriguingly, all these pathogens mentioned above can cause pig respiratory disease in the lung. This consistent finding indicates that CXCL10 coupled with its specific receptor CXCR3 (Groom & Luster, 2011) is a common mechanism in pigs to induce pro-inflammatory response to pulmonary infections.

After recognition of pathogens by pattern recognition receptors, innate immune cells are recruited to the infection site by chemokines as exemplified by the CXCL10 mentioned above. The recruited immune cells including neutrophils, macrophages, and NK cells are then activated to clear the pathogens by phagocytosis, cytotoxicity, production of anti-microbial and pro-inflammatory molecules. One of the important defensive mechanisms that is utilized by innate immune cells is the release of reactive oxygen species (ROS) and reactive nitrogen species (RNS) which could generate superoxide and nitric oxide (Nathan & Shiloh, 2000). This is a nonspecific response that can efficiently kill invading pathogens but also induce dramatic tissue injury (Mittal et al., 2014). Production of ROS and RNS also induce oxidative stress which could cause cell apoptosis and impaired antigen presentation (Kannan & Jain, 2000; Preynat-Seauve et al., 2003), and this may dampen the activation of adaptive immune response. Resolution of inflammation once pathogens are controlled or cleared is thus extremely important to maintain body homeostasis. Failure to resolve inflammation or induction of over-activated inflammatory response could cause irreversible tissue damage and even death. One popular example is the acute respiratory distress syndrome (ARDS) induced by the

severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection that is known as the major cause of death in the coronavirus disease 2019 (COVID-19) patients. Cytokine storm and oxidative stress are the main inducers of ARDS (Meftahi et al., 2021). As mentioned in the result shown in Figure 3.17, the Early_dead group up-regulated a major cytokine storm inducer *CCL5* but down-regulated an anti-oxidative stress gene *SGK1*, indicating a greater risk of those pigs than other survived pigs to develop cytokine storm and severe respiratory syndrome. Endoplasmic reticulum (ER) stress during viral infection can also induce production of ROS (Fung & Liu, 2014). A study from our collaborators also implies pigs from the Early_dead group may die from over-activated immune response during the natural disease challenge (Schmied et al., 2018). They applied a pre-challenge test named "High Immune Response" which included assessment of antibody and cell mediated immune response upon immunization of healthy pigs with model antigens. Surprisingly, the pigs were high with both antibody and cell mediated immune response all died after the natural disease challenge. In contrast, the Resilient pigs up-regulated an anti-ER gene (*MANF*), an anti-oxidative stress gene (*SGK1)*, and some "immune braking" genes (e.g. *IOD1*, *IL27*) after disease challenge, making them more adapted to balance the activation and resolution of inflammatory response. Intriguingly, these gene expression changes seemed unique to the Resilient group, highlighting the advantage of selecting disease resilience phenotypes.

## 3.5 Summary

To summarize this chapter, I started with using a small cohort of 58 pigs (cohort 1) to test our hypothesis that the pre-challenge transcriptome of whole blood cells is able to

identify biomarker genes to predict post-challenge phenotypes, especially disease resilience. DE and pathway enrichment analysis suggested a hierarchical model for the baseline immunity in four groups of pigs that were separated to “Resistant”, “Resilient”, “Early_dead”, and “Susceptible” based on their post-challenge health and productivity records. In this model, the Resistant and Resilience group outperformed the other two by expressing higher levels of immune defensive genes. Meanwhile, the Resilient group expressed a relatively low level of genes associated with cell activities and thus supporting a energy-saving feature of its baseline immunity. To validate the findings from cohort 1, I did the same analysis on a larger cohort 2 with 209 pigs. Many DE genes or transcripts identified in cohort 1 were more likely “noises” that were not recapitulated in cohort 2. However, many newly identified DE genes and transcripts supported the major idea that the Resistant and Resilient group had a strong baseline level of immune defensive genes. Moreover, more DE hits were also revealed that suggested the Early_dead group had even more active baseline immunity than the Susceptible group. Besides the blood samples collected before disease challenge, cohort 2 samples also included two additional collections of blood that were 4 and 8 weeks post-challenge. This made it possible to track gene expression change upon time in different groups of pigs in our natural disease challenge model. Through DE analysis between different collections of blood, mobilization and contraction of immune response could be revealed at the early and late blood collections, respectively, in all the pigs. Notably, DE and pathway enrichment analysis suggested different features of dynamic immune regulation in different groups. It should be highlighted that pigs from the Resilient group tended to implement an “economical” strategy in response to pathogen challenge by restricting the

energy consumption of immune cells and increase the expression of anti-stress genes to maintain the homeostasis after pathogens are controlled, but not inducing sharp and strong response trying to clear the infections. This unique feature for resilient pigs may confer them a more modern farm-favored phenotype by well-balancing the health status and productive performance.

# CHAPTER 4: Investigation of using plasma metabolites to predict pig disease resilience

## 4.1 Background

Metabolites are organic and inorganic low-molecular-weight molecules that are substrates, intermediates, or terminal products in biological activities or processes. Due to their essential roles in broad biological pathways, these metabolites are typically regarded as the outcome of gene expression in an organism. The metabolome is defined as a collection of all metabolites in a specific biological sample (e.g. cells, tissues, organs, or organisms). Metabolomics is the comprehensive assessment and systematic study of the metabolome (Jordan et al., 2009). Metabolites not only participate in metabolic processes but also functionally interact with and actively regulate other omics levels (Johnson et al., 2016). For example, metabolites can regulate or mediate protein activity as co-factors of enzymes or co-substrates that catalyze particular reactions (Jacob & Monod, 1961; Yang et al., 2018). They can also modulate RNA metabolism and fate by binding to ribosomes and regulating post-transcriptional modifications (Jones & Ferré-D'Amaré, 2017; Warth et al., 2018). In addition, metabolites are actively involved in epigenetic regulation by functioning as cofactors or co-substrates for chromatin-modifying proteins (Watanabe et al., 2013). Metabolites are promising biomarker candidates, representing functional phenotypes of a biological system because of the following two reasons. First, metabolites are derived from various biological processes; and second, they broadly participate or modulate gene expression. Notably, one can also reveal the mechanisms of

metabolites in phenotypic regulation through metabolic-based activity screening (Guijas et al., 2018), thereby indicating the status or predicting the performance change of an organism in certain contexts.

In previous studies, metabolomics was used to assess the progression of diseases, and some metabolites related to disease development were used in the diagnosis and prognosis of diseases, which provides guidance for subsequent treatment and to develop personalized treatment plans. One example is that some key metabolic markers (e.g. gamma glutamyl dipeptides) can be used to distinguish among nine types of liver diseases (Soga et al., 2011). Other examples include use of metabolite profiles as biomarkers of early onset of disease to help diagnose oral cancer (Tiziani et al., 2009), ovarian cancer (Chen et al., 2011), and malignant pancreatic lesions (Bathe et al., 2011), etc., which fulfills the purpose of early detection and early treatment to improve successful treatment rates. However, for livestock, there were studies linking metabolites concentration in blood with productive performance (Chapinal et al., 2012; Kim et al., 2017; Moran et al., 2019), but not health conditions. Inspired by the applications of metabolites in human disease, I thus hypothesized that metabolites may potentially predict the response of livestock to pathogen infection. Since metabolites are substrates, intermediates, or final components of livestock products, screening metabolites to predict pig disease resilience may be another promising direction in addition to using specific RNA and protein levels in the aforementioned chapters.

In this chapter, I investigated whether metabolomics is applicable to predict pig disease resilience in our natural disease challenge model. I adopted metabolomics approaches to determine the metabolome of pigs before the pathogens challenge and

analyzed the correlation of this prior-challenging metabolome profile with different health and productivity-related post-challenging phenotypes. Similar to what was described in the previous two chapters, this study aimed to test whether there are specific metabolic markers to predict the "resilience" phenotype of pigs.

## 4.2 Methods

### 4.2.1 Grouping of pigs for metabolome analysis

The rationale behind the selection of pigs was based on our natural disease challenge model elaborated in Chapters 2 and 3. A total of 460 pigs divided into seven batches (batches 26-32) from Cycle 4 and Cycle 5 were selected for the metabolome analysis experiment. Those pigs were grouped into "Resistant", "Resilient", "Early_dead", and "Susceptible" classes as described in Chapter 3 based on their pre-slaughter weights, health conditions, and survival during the experiment period. Detailed pig grouping records and methods are elaborated in section 4.3.2 below.

### 4.2.2 Blood sample collection and plasma preparation

Blood samples were collected in the tubes containing EDTA and centrifuged at 3000 rpm for 10 minutes at 4°C. Next, plasma was collected from the top transparent layer and aliquoted in 500 µL per Nunc tube and stored at -80°C before the NMR assay.

### 4.2.3 Sample preparation and NMR spectroscopy

Frozen plasma samples were thawed at room temperature and shaken well to homogenize before use. Before the NMR spectroscopy, plasma proteins were removed through a deproteinization step involving ultra-filtration as previously described

(Psychogios et al., 2011). 3 KDa cut-off centrifugal filter units (Amicon Microcon YM-3) were rinsed five times with each wash using 0.5 mL $H_2O$ and centrifuged (10,000 rpm for 10 minutes) to remove residual glycerol bound to the filter membrane. Each thawed plasma sample was then transferred into the centrifuge filter devices, and spun (10,000 rpm for 20 minutes) to remove macromolecules (primarily protein and lipoproteins) from the sample. The filtrates were checked visually to make sure that the membrane was intact after filtering. The filtrates were subsequently collected and the volumes were recorded. Subsequently, 50 mM potassium salt buffer (pH 7) was supplemented to the filtrates to ensure the total volume of washed sample was a minimum of 200 µL. Any sample supplemented with potassium salt buffer was annotated with the dilution factor and metabolite concentrations were corrected accordingly in the subsequent analysis. Next, 50 µL of the standard buffer solution (54% D2O: 46% 250 mM K2HPO4 + KH2PO4 pH 7.0 v/v containing 5 mM 2, 2-dimethyl-2-silcepentane-5-sulphonate (DSS), 5.84 mM 2-chloropyrimidine-5 carboxylate, and 0.1% NaN3 in H2O) was added to the sample.

A total of 250µL serum sample was then transferred to a 3 mm SampleJet NMR tube for NMR analysis. All $^1$H-NMR spectra were recorded in the 700 MHz Avance III (Bruker) spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) cryoprobe. $^1$H-NMR spectra were detected at 25°C using the first transient of the NOESY pre-saturation pulse sequence (noesypr1d) considering its high degree of quantitative accuracy (Saude et al., 2006). All FID's (free induction decays) were zero-filled to 250 K data points. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm). All $^1$H-NMR spectra were

processed and analyzed using the online Bayesil software package which can identify and quantify the metabolites from the biological samples using de-convolution process (Ravanbakhsh et al., 2015). The NMR assays were carried out by The Metabolomics Innovation Centre at the University of Alberta.

### 4.2.4 Metabolomics statistical analysis

The effect of batch, sire, and dam on metabolites was tested using Fitting Linear Models in R before metabolomics analysis. The cutoff for statistical significance was P-value < 0.05. All metabolomics data were processed and analyzed using an integrated web-based platform, MetaboAnalyst 4.0 (https://www.metaboanalyst.ca) (Chong et al., 2019). MetaboAnalyst offers a variety of methods commonly used in metabolomic data analysis. In this study, we primarily focused on T-test, ANOVA, PCA, Partial least squares-discriminant analysis (PLS-DA, a supervised method that uses multivariate regression techniques to extract via the linear combination of original variables the information that can predict the class membership), and Significance Analysis of Microarray (SAM, addresses the false discovery rate when running multiple tests on high-dimensional microarray data). The statistical procedures for metabolomics analysis were performed according to previously published protocols (Chong et al., 2019). In this study, we used logarithmic normalization to transfer the metabolite concentrations acquired from the NMR experiment.

### 4.2.5 Machine learning-based phenotype prediction using metabolites data

The data cleaning step was performed to remove the missing values in the NMR

dataset. Then, PCA was used to extract the principal components from the whole NMR data. The input data for machine learning to predict pig performance classification consisted of original (unadjusted) NMR data, batch-adjusted residuals of the original NMR data, and PCA-processed NMR data. First, several popular supervised machine learning prediction algorithms were used in this study including Random Forest (RF, a supervised learning algorithm suitable for high dimensional class prediction). K Nearest Neighbor (KNN, determines the category of the sample to be classified according to the category of the nearest sample or samples), Decision Tree (DT, a tree-like model to predict possible consequences), support vector machine (SVM, solving the separation hyperplane that correctly divides the training data set and has the largest geometric interval), and logistic regression (LR, to predict target values based on linear regression model) (Riyaz et al., 2022). Optimal hyper-parameters used in each algorithm were manually determined using the original concentrations of metabolites and detailed command was listed as following: RF-RandomForestClassifier (n_estimators=500, max_depth=6, random_state=0, bootstrap=True); KNN-KNeighborsClassifier (n_neighbors=60); DT-DecisionTreeClassifier(random_state=0, max_depth=6, criterion="entropy"); SVM-SVC(kernel='rbf'); LR-LogisticRegression(random_state=0, solver = 'lbfgs', multi_class='multinomial') . The prediction results for each pig grouping method and prediction algorithm were presented as predicting accuracies. To further validate the prediction accuracies from each pig grouping method, a permutation test was conducted. The function “random.permutation()” from the package “NumPy" was used to randomly shuffle the samples for a thousand rounds, and baseline predicting accuracy was determined by cutting off the top 5% permuted rounds with highest predicting

performance (If a specific machine model returns an accuracy higher than this baseline accuracy, it means there is a 95% possibility for a convincible prediction, or statistically a $P<0.05$).

Two independent publicly available datasets with clean human blood metabolome were tested in order to validate the performance of the aforementioned supervised machine learning algorithms and the analysis pipeline used. The first dataset (MTBLS90: Large-scale non-targeted serum metabolomics in the Prospective Investigation of the Vasculature in Uppsala Seniors) from the MetaboLights database (Haug et al., 2020) contains the plasma concentrations of 189 metabolites determined by liquid chromatography/mass spectrometry (LC/MS) from males (n=485) and females (n=483) at age of 70. Machine learning was conducted to predict the sex of the subjects using all the original metabolites concentrations. The second dataset (ST000369: Investigation of metabolomic blood biomarkers for detection of adenocarcinoma lung cancer (part II) within the Project PR000293) from the Metabolomics Workbench database (https://www.metabolomicsworkbench.org) includes the serum concentrations of 181 metabolites determined by gas chromatography/mass spectrometry (GC/MS) from the patients of adenocarcinoma lung cancer (n=100) and healthy controls (n=63). Machine learning was performed to predict the cancer/healthy state of the subjects using all the original metabolites concentrations. Results of the prediction accuracies are listed in Appendix 4.3 and 4.4.

Next, another unsupervised machine learning prediction algorithm called “K-means” was applied as a comparison to the above-mentioned supervised algorithms. The samples were divided into several clusters by K-means clusters based on metabolite concentration.

In metabolites analysis, the K-means cluster was performed by calculating the Euclidean distance measure on the correlation coefficients between the metabolite concentrations (Nikas & Low, 2011). Basically, K-means clustering uses standardized metabolite concentrations to generate several clusters of pigs and the cluster number equals the number of groups under a certain grouping method. The correlation between the generated clusters and groups was in turn assessed by Spearman's rank correlation coefficient since the clusters are not in accordance with normal distribution. Machine learning was programmed using sklearn and scipy on Python 3.6. Training (75%) and testing (25%) samples were split according to the command train_test_split (X, Y, random_state=0) without considering the pig phenotype grouping (Avuçlu & Elen, 2020; Woillard et al., 2021).

## 4.3 Results

### 4.3.1 Determining the effect of the batch, dam, and sire on metabolite concentrations using Fitting Linear Model

A total of 38 metabolites were resolved by NMR from the plasma samples of the 460 pigs. To determine whether the pig batch, dam, and sire had impacts on the animal metabolome, I first generated a linear model using pig batch, dam, and sire information to fit with single metabolite concentration. Then I applied ANOVA analysis (anova.lm in R) and the results were listed in Table 4.1. There were 19 metabolites significantly related to the dam with P-value < 0.001, and 14 metabolites were significantly related to the dam with $0.05 \geqq$ P-value >0.001, whereas 5 metabolites showed no relationship (P> 0.1). As expected, most metabolites were remarkably influenced by the dam since it is a genetic

factor. Surprisingly, I found that 36 out of 38 metabolites were significantly ($P< 0.001$) related to the pig batch. L-Histidine was also significantly ($P=0.007$) related to the pig batch whereas only acetic acid did not reach the threshold ($P=0.12$). Considering batch factors showed such a big impact on metabolite concentration, I generated batch-adjusted concentrations with a linear model for the metabolites detected in the NMR experiment to exclude the batch effect in the following statistical analysis.

### 4.3.2 Classification of pigs using production and health-related traits

To determine if there was an association between the specific metabolite concentration and the pig performance after challenging, we grouped pigs according to the traits as described previously (shown in Figure 3.1). These trait data for the Resistant, Resilient, Susceptible, and Early_Dead groups are also listed in Table 4.2.

In addition to the classification listed in Table 4.2, I also applied several other ways to group the pigs for statistical analysis based on the single trait or possible combination of the two traits shown in Table 4.3:

1) Health status (Healthy or Sick): number of treatments was zero or non-zero for pigs reaching slaughter;

2) Mortality (Dead or Alive): animals were dead or alive before slaughter age;

3) Productivity performance (Underweight, Normalweight, or Overweight): pre-slaughter weight was under 115kg, between 115-125kg, or over 125kg, respectively.

4) Health status + Productive performance: healthy and overweight (HO), healthy and normal weight (HN), healthy and underweight (HU), sick and overweight (SO), sick and normal weight (SN), and sick and underweight (SU).

**Table 4.1** ANOVA for Linear Model Fits (F test)

| Metabolite | P (Batch) | P (Dam) |
|---|---|---|
| 2-Hydroxybutyrate | 4.50E-05 | 4.47E-03 |
| Acetic acid | 1.24E-01 | 5.61E-01 |
| Creatine | 1.23E-22 | 5.48E-04 |
| Dimethylglycine | 2.41E-58 | 4.57E-16 |
| Citric acid | 1.48E-23 | 1.89E-04 |
| Choline | 9.98E-39 | 1.84E-01 |
| D-Glucose | 1.49E-21 | 2.04E-02 |
| Glycine | 1.44E-06 | 1.16E-03 |
| Glycerol | 1.89E-33 | 2.90E-01 |
| Formate | 1.16E-17 | 3.75E-02 |
| L-Glutamic acid | 1.98E-12 | 1.32E-03 |
| Hypoxanthine | 3.60E-50 | 5.38E-07 |
| Tyrosine | 9.91E-11 | 1.30E-04 |
| L-Phenylalanine | 2.14E-04 | 1.79E-02 |
| L-Alanine | 1.26E-10 | 2.69E-04 |
| L-Proline | 1.35E-08 | 5.55E-02 |
| L-Threonine | 3.21E-24 | 3.81E-09 |
| L-Asparagine | 1.11E-12 | 3.31E-04 |
| Mannose | 3.04E-25 | 1.28E-04 |
| Isoleucine | 3.23E-37 | 3.46E-07 |

| Metabolite | P (Batch) | P (Dam) |
|---|---|---|
| L-Histidine | 7.94E-03 | 2.27E-01 |
| L-Lysine | 3.73E-08 | 1.82E-03 |
| L-Serine | 1.99E-20 | 4.39E-05 |
| L-Lactic acid | 1.93E-12 | 3.72E-05 |
| Aspartate | 4.52E-09 | 9.83E-01 |
| Oxoglutarate | 1.01E-07 | 9.41E-03 |
| ornithine | 8.75E-07 | 2.33E-03 |
| Pyruvic acid | 1.59E-50 | 3.94E-07 |
| Succinate | 1.19E-36 | 2.14E-10 |
| Urea | 6.23E-13 | 1.48E-03 |
| L-Alpha-aminobutyric acid | 1.89E-23 | 2.40E-03 |
| 3-Methyl-2-oxovaleric acid | 3.39E-25 | 1.16E-07 |
| Creatinine | 1.02E-21 | 5.04E-13 |
| L-Glutamine | 1.78E-07 | 3.31E-02 |
| L-Leucine | 2.47E-06 | 1.30E-04 |
| Methionine | 1.59E-18 | 1.90E-03 |
| Valine | 2.40E-09 | 5.58E-07 |
| Isobutyric acid | 2.98E-14 | 1.07E-04 |

**Table 4.2** Summary of the performance traits and group classification for animals analyzed for blood metabolites.

| | **Resistant** | **Resilient** | **Susceptible** | **Early_Dead** |
|---|---|---|---|---|
| **Total numbers** | 93 | 118 | 147 | 100 |
| **Mortality** | Alive | Alive | Alive | Dead |
| **Number of treatments** | 0 | 1-5 | 1-6 | 0-7 |
| **Pre-slaughter weight[1] (kg)** | 87-139 | 120-142.5 | 63-119.5 | 5.6-95 |
| **Range for slaughter age[2] (days)** | 160-191 | 159-206 | 161-191 | 32-183 |

[1] Pre-slaughter weights: The Pre-slaughter weights under the Early_Dead group were corpse weights.
[2] Range for slaughter age: The Slaughter age under the Early_dead group was the death age.

**Table 4.3** Summary of performance traits data for additional classifications.

| | Total numbers | Morta--lity | Treatment number | Pre-slaughter weight[1] (kg) | Slaughter age[2] (days) |
|---|---|---|---|---|---|
| **Healthy** | 93 | Alive | 0 | 87-139 | 160-191 |
| **Sick** | 267 | Alive | 1-6 | 63-142.5 | 159-206 |
| **Alive** | 360 | Alive | 0-6 | 63-142.5 | 159-206 |
| **Dead** | 100 | Dead | 0-7 | 5.6-95 | 32-183 |
| **Over weight** | 99 | Alive | 0-5 | 125.5-142.5 | 162-206 |
| **Normal weight** | 120 | Alive | 0-5 | 115.5-125 | 159-191 |
| **Under weight** | 139 | Alive | 0-6 | 63-115 | 161-186 |
| **HO** | 29 | Alive | 0 | 126.5-139 | 162-186 |
| **HN** | 39 | Alive | 0 | 116-125 | 160-191 |
| **HU** | 25 | Alive | 0 | 87-115 | 161-186 |
| **SO** | 70 | Alive | 1-5 | 125.5-142.5 | 162-206 |
| **SN** | 81 | Alive | 1-5 | 115-125 | 159-191 |
| **SU** | 114 | Alive | 1-6 | 63-115 | 162-186 |

[1]Pre-slaughter weight: The Pre-slaughter weight under the Death group was corpse weight.
[2]Slaughter age: The Slaughter age under the Death group was death age.

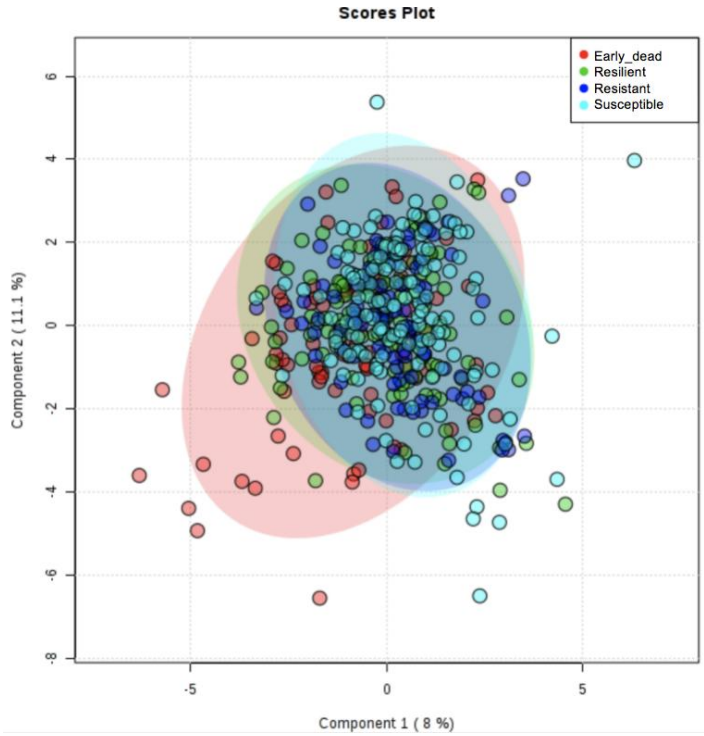
### 4.3.3 Statistical analysis for metabolic data with pig phenotypes using MetaboAnalyst

Metabolite concentration data was standardized by Quantile normalization to produce a Gaussian distribution prior to statistical analysis. In this study, I used logarithmic transformation to transfer all the metabolite data for the analysis in MetaboAnalyst. It is worth noting that the One-way ANOVA analysis for the batch effect in MetaboAnalyst revealed significant (FDR < 0.05) differences for the concentrations of 35 metabolites among different batches (Appendix 4.1), which is consistent with the ANOVA results from R shown above (Table 4.1). Therefore, without further clarification, the batch-adjusted concentrations of metabolites were used as input data in the following analysis including T-test, ANOVA, PCA, PLS-DA, and SAM. Only the results with the FDR / P-value < 0.05 were shown.

Initially, to have a global view of the relationship between classified pig groups defined in Table 4.2 and Table 4.3 regarding the metabolite concentrations, I first applied PLS-DA and observed a separation of some Early_Dead individuals from others when combining all three traits (Figure 4.1A). PLS-DA also revealed a separation of some Early_Dead individuals from other individuals when mortality was the exclusive classification parameter (Figure 4.1B). The ANOVA test for all the metabolites detected was performed to compare the concentrations among groups, and the significant results with either unadjusted or batch-adjusted data are listed below following the order of group methods introduced above. Succinate was shown to have significantly (FDR < 0.05) different concentrations between the Early_Dead and any of the remaining three groups, respectively, and Dimethylglycine was significantly (FDR < 0.05) different between Early_Dead and either Resilient or Resistant group (Table 4.4 and Figure 4.2). These

results were consistent with the result of the SAM test. Only if unadjusted concentrations were applied, was the Succinate level significantly (FDR = 0.0002) different between the Alive and Dead group in the T-test (Figure 4.3). The resultant data was also supported by the SAM test results. Of note, the T-test and SAM test based on batch-adjusted metabolite concentration revealed significant (FDR = 0.03) difference in concentrations of another metabolite, pyruvic acid, between the Alive and Dead group (Figure 4.4). This result suggests that pyruvic acid plays important roles in circulating immune cells to maintain efficient control of infection. Altogether, several metabolites in pigs from Early_Dead (or Dead) group were found to exhibit different concentrations. Dimethylglycine showed a trend to have lower plasma concentration in the Resistant and Resilient group compared to that of the other two. Nevertheless, neither of these aforementioned metabolites could be used to differentiate the favorable Resilient pigs from the Susceptible group for animal selection purpose. In addition, the ANOVA test for the grouping method (illustrated in Table 4.3) also revealed the unadjusted concentration of Dimethylglycine was significantly different (FDR = 0.01) in the comparisons SO-HN, HN-SU, HO-HU, SO-HU, SU-HU, SO-SN, SU-SN (Fisher's LSD) (Table 4.4 and Figure 4.5). Intriguingly, in healthy pigs Dimethylglycine has a higher plasma concentration in the HO than the HU group (Healthy overweight and underweight respectively). In other words, the pre-challenge concentration of Dimethylglycine is positively correlated with subsequent body weight of disease-free pigs in response to natural disease challenge. In contrast, the SO and SU groups (Sick overweight and underweight respectively) showed a comparable level of Dimethylglycine, indicating Dimethylglycine is not correlated with post-challenge body weight if the pigs were sick.

**A**


Scores Plot
Early_dead
Resilient
Resistant
Susceptible
Component 2 ( 11.1 %)
Component 1 ( 8 %)

**B**

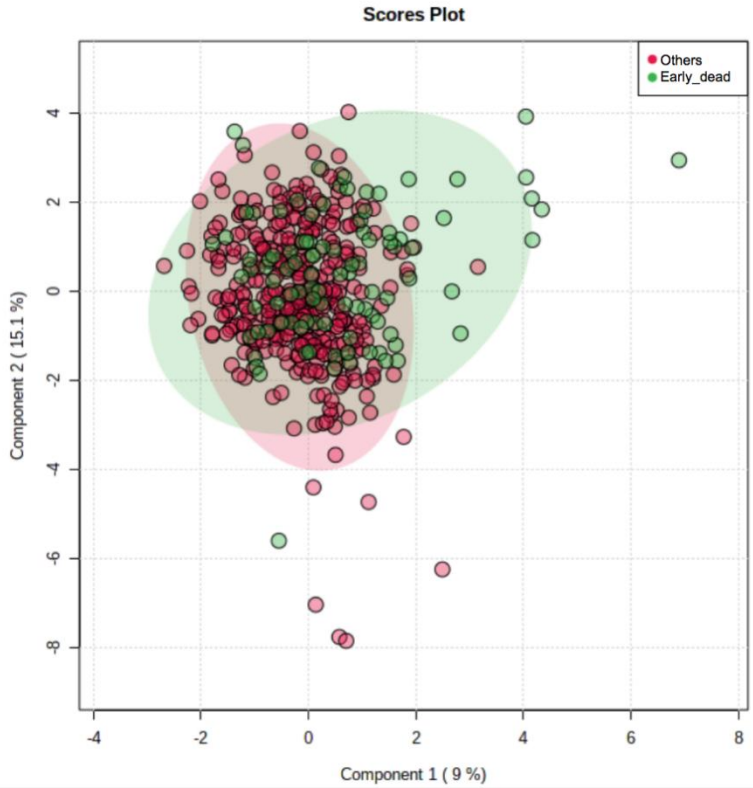

Scores Plot
Others
Early_dead
Component 2 ( 15.1 %)
Component 1 ( 9 %)

**Figure 4.1** PLS-DA analysis shows the separation of some individuals from the Early_Dead group.

(A) The result based on the grouping method combining all three performance traits. (B) The result based on the grouping method focusing on mortality.

**Table 4.4** Summary of metabolite hits with significantly different levels in the four major groups shown in Table 4.2 and six two-trait-defined groups shown in Table 4.3 from the ANOVA analysis without batch adjustment.

| Metabolites | P-value | FDR | Fisher's LSD (significant comparisons) |
|---|---|---|---|
| **Succinate** | 0.0001 | 0.0049 | Early_Dead vs Resilient;<br>Early_Dead vs Resistant;<br>Early_Dead vs Susceptible |
| **Dimethylglycine** | 0.0011 | 0.0212 | Early_Dead vs Resilient;<br>Early_Dead vs Resistant;<br>Susceptible vs Resistant |
| **Dimethylglycine** | 0.0005 | 0.01 | SO vs HN;<br>SU vs HN;<br>HO vs HU;<br>SO vs HU;<br>SU vs HU;<br>SO vs SN;<br>SU vs SN |

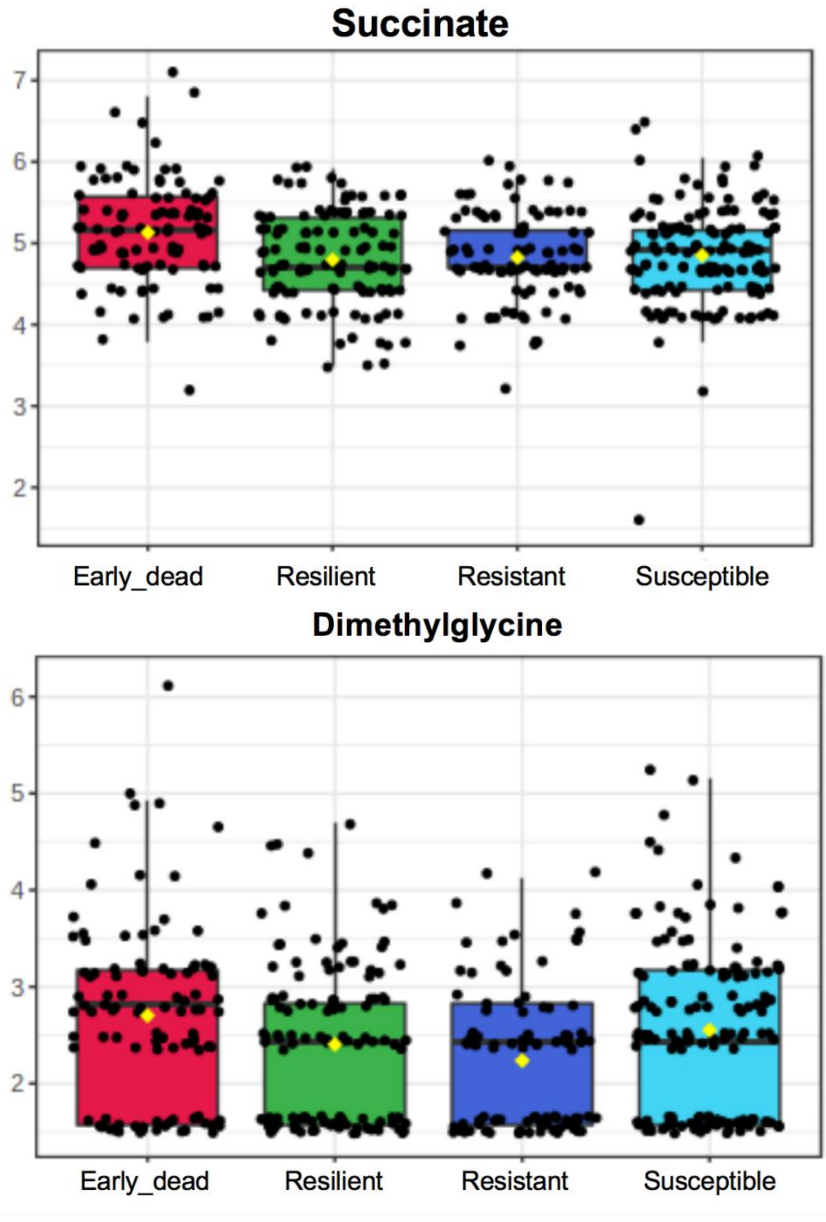



**Figure 4.2** The two metabolites that have significantly different concentrations between the Early_Dead group and the other three groups identified from the ANOVA test using the grouping method shown in Table 4.2 without batch adjustment. Result of statistical analysis is listed in Table 4.4.

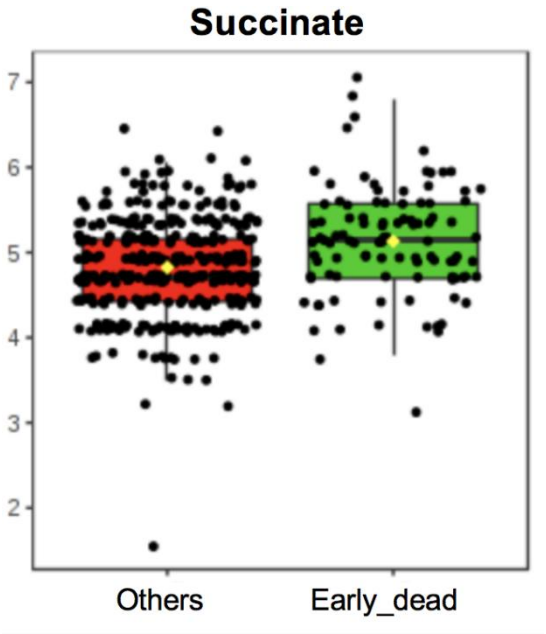



**Figure 4.3** Comparison of unadjusted concentration of Succinate between the pigs from the Alive and Dead groups (FDR = 0.0002).

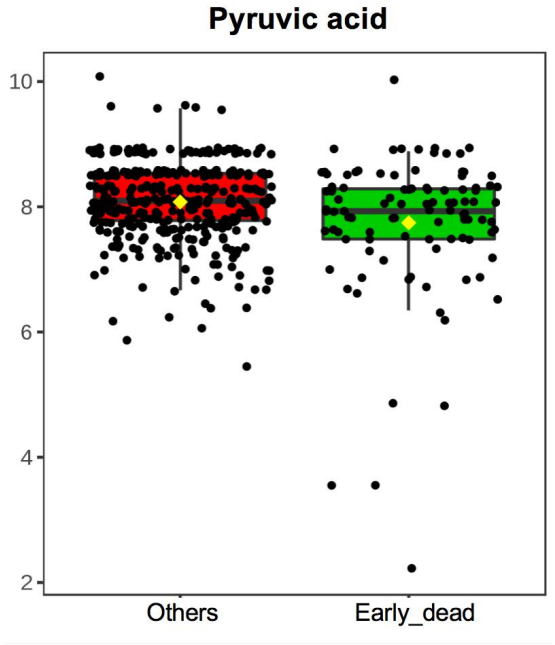



**Figure 4.4** Comparison of batch-adjusted concentration of Pyruvic acid between the pigs from the Alive and Dead groups (FDR = 0.03).

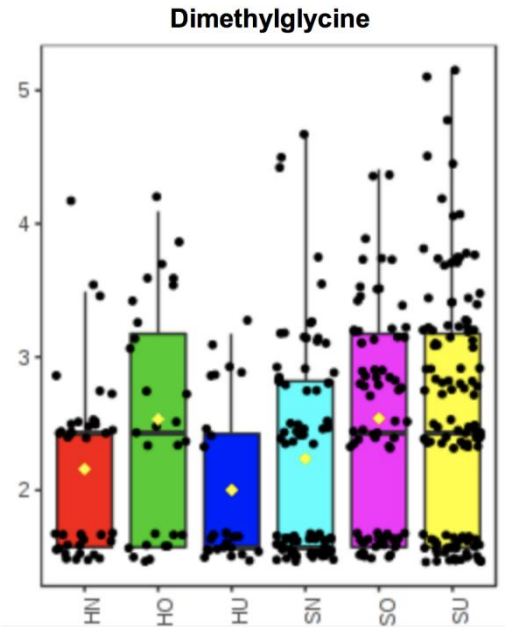



**Figure 4.5** Unadjusted concentrations of Dimethylglycine in different groups defined in Table 4.3.

Significant differences (P=0.0005, FDR=0.01) were found comparing SO vs HN, SU vs HN, HO vs HU, SO vs HU, SU vs HU, SO vs SN, and SU vs SN. Result of statistical analysis is also listed in Table 4.4.

### 4.3.4 Machine Learning

The statistical analysis mentioned above did not identify any potential metabolic biomarker to predict the pigs' potential of resilience or resistance in response to pathogen challenge. My alternative approach was then to use the integrated metabolic profile to pinpoint the desired phenotypes. Machine learning has been used on metabolites data derived from bio-fluid to advance the characterization and classification of disease developing stage and subtypes (Kouznetsova et al., 2019; Machado-Fragua et al., 2022; Reel et al., 2022). Thus, several commonly-used supervised machine learning algorithms (KNN, DT, RF, SVM, and LR) were trained using unadjusted or batch-adjusted metabolite concentrations to explore if any of those models could accurately predict the group of pigs with desired phenotypes under particular grouping methods. Since the metabolome contains dynamic systematic information of the metabolic status at a specific time point, the determined metabolome would inevitably contain some metabolites that have continuous fluctuations in concentration and may become noise or mislead the prediction of machine learning algorithms. Meanwhile, PCA was applied on the unadjusted or batch-adjusted metabolite concentrations to simplify the complexity of the training data and trained the models mentioned above based on the input of those PCA-processed data which could explain over 80% or 90% of the variance. The most representative single component only covers about 20% of the total variation (Appendix A4.2). This result from PCA analysis indicates that it is impractical to use a few PCA components to represent most variables of the whole dataset. In addition, an unsupervised machine learning algorithm called K-means clustering was performed to examine the correlations between the modeled clusters and the corresponding defined groups from

different grouping methods. The results of K-means clustering are listed below after the results of supervised machine learning algorithms.

For the grouping method combining all the three traits (number of treatments, mortality, and pre-slaughter weight), most supervised algorithms except LR returned an accuracy higher than the permuted baseline accuracy (Figure 4.6) using either unadjusted or batch-adjusted metabolite concentrations. Except for the DT model using batch-adjusted data, PCA processing did not notably improve the prediction accuracy (Table 4.5). Even though the metabolites data did help to give significantly higher predicting accuracy than random prediction, most actual accuracies were under 0.4 which is not useful for practical pig selection purposes. Similarly, neither of those supervised algorithms performed very well (most accuracies ~0.3 to 0.4) to predict the groups from the grouping method focusing on production ability only or combining health status and production ability, respectively (Table 4.8 and 4.9). Unexpectedly, all supervised algorithms returned very high prediction accuracies around 0.8 when the grouping method was based on survival statement or health status information (Table 4.6 and 4.7). It should be noted that the group size has big variations when grouping the pigs as indicated in Table 4.6 and 4.7, which may lead to biased decisions for the group that had a much larger size during the training step. However, compared to the permuted baseline accuracy (Figure 4.7 and 4.8), all the algorithms performed well with higher predicting accuracies, again indicating the metabolites data did significantly improve the prediction.

Results of the prediction accuracies of public datasets are listed in Appendix 4.3 and 4.4. Of note, trait prediction by several algorithms using metabolomic data from both datasets remarkably improved the permuted baseline accuracy to around 1.5 fold, and the

SVM (0.83) and LR (0.83) algorithm performed the best for first and second public dataset, respectively. These results validated the machine learning pipeline that was used for predicting pig response using plasma metabolome, and also revealed big variations among different algorithms in the performance to differentiate particular traits.

Furthermore, to reduce the effect of unrelated metabolites, the input data size was compacted by only using the concentration of the three metabolites (Succinate, Dimethylglycine, and Pyruvic acid) that showed significant differences among particularly defined groups (Figure 4.2 to 4.5) would increase the prediction accuracy. Even so with this modification only slight improvement was observed for particular algorithms (Table 4.10 compared to Table 4.5 and 4.6). Overall, the algorithm DT showed relatively better performance in prediction of the multi-factor-determined phenotypes (Resistant, Resilient, Susceptible, and Early_Dead) with the highest accuracy of 0.38 using resdPCA24. SVM was found more robust to predict the two-factor-determined phenotypes (HO, HN, HU, SO, SN, and SU, highest accuracy= 0.36) and the productivity trait (Underweight, Normalweight, and Overweight, highest accuracy=0.40). RF and LR were outstanding in prediction of mortality (highest accuracy=0.81) whereas KNN was more proper to predict health status (highest accuracy=0.79) (Table 4.5 to 4.10). It is promising that the tested supervised algorithms showed some potential for utilizing plasma metabolites data to predict pig health and productive phenotypes. More optimizations may be needed to improve the algorithms' performance as the predicting accuracy was still far-from-ideal to be used in practical pig selection.

**Table 4.5** Accumulated accuracy results of each different machine learning algorithm for predicting the grouping of Resistant, Resilient, Susceptible, and Early_dead.

"con", "conPCA12", and "conPCA17" indicates all the unadjusted metabolite concentrations, 12 components of the PCA explaining over 80% of the variance of the unadjusted metabolite concentrations, and 17 components of the PCA explaining over 90% of the variance of the unadjusted metabolite concentrations, respectively. "resd", "resdPCA18", and "resdPCA24" indicates all the batch-adjusted metabolite concentrations, 18 components of the PCA explaining over 80% of the variance of the batch-adjusted metabolite concentrations, and 24 components of the PCA explaining over 90% of the variance of the batch-adjusted metabolite concentrations, respectively. (this description also applies to the following tables in this chapter)

| | **con** | **conPCA12** | **conPCA17** | **resd** | **resdPCA18** | **resdPCA24** |
|---|---|---|---|---|---|---|
| **KNN** | 0.36 | 0.34 | 0.33 | 0.32 | 0.33 | 0.30 |
| **DT** | 0.34 | 0.28 | 0.24 | 0.30 | 0.37 | 0.38 |
| **RF** | 0.33 | 0.34 | 0.30 | 0.31 | 0.32 | 0.31 |
| **SVM** | 0.30 | 0.30 | 0.29 | 0.31 | 0.32 | 0.34 |
| **LR** | 0.26 | 0.32 | 0.30 | 0.25 | 0.26 | 0.28 |

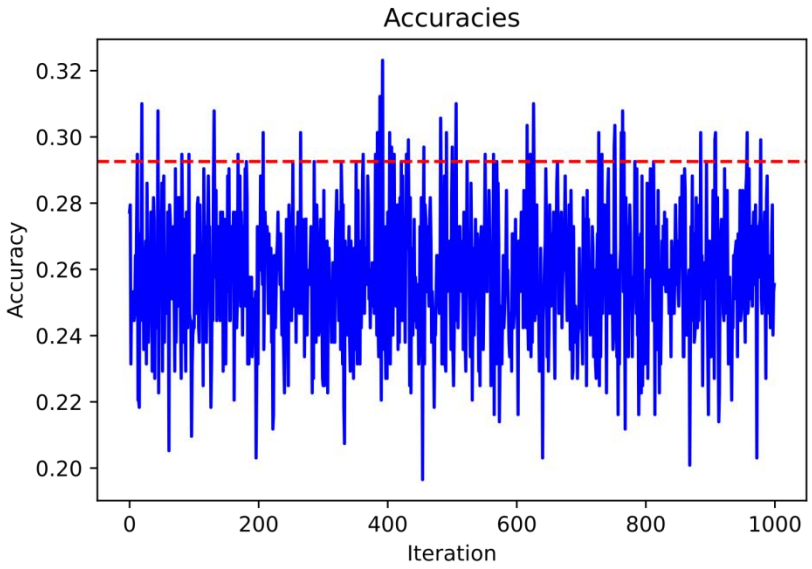



**Figure 4.6** Permutation result for the grouping method used in Table 4.5. The red dotted line represents the baseline accuracy for a convincible prediction (Also applies to Figure 4.7 to 4.10).

**Table 4.6** Accumulated accuracy results of each different machine learning algorithm for predicting the grouping of Dead and Alive.

| | con | conPCA12 | conPCA17 | resd | resdPCA18 | resdPCA24 |
|---|---|---|---|---|---|---|
| **KNN** | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 |
| **DT** | 0.80 | 0.79 | 0.74 | 0.79 | 0.76 | 0.72 |
| **RF** | 0.80 | 0.78 | 0.80 | 0.80 | 0.81 | 0.80 |
| **SVM** | 0.81 | 0.81 | 0.81 | 0.80 | 0.81 | 0.80 |
| **LR** | 0.77 | 0.78 | 0.81 | 0.79 | 0.78 | 0.79 |

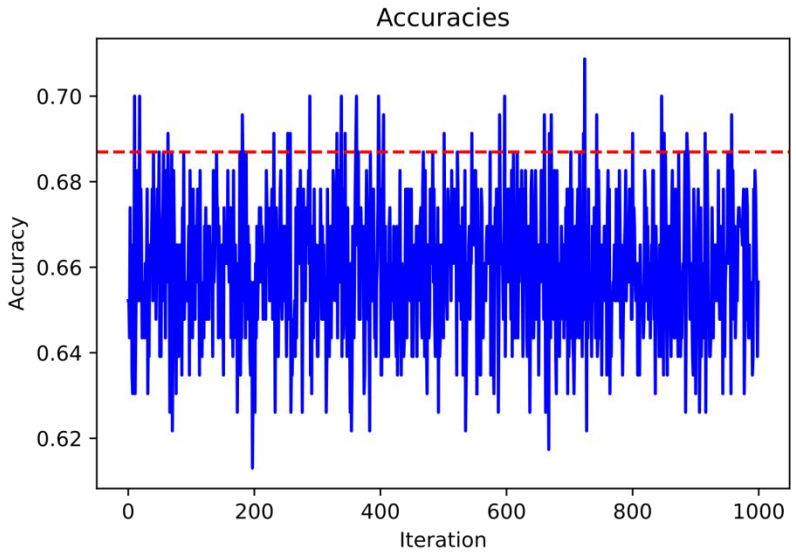



**Figure 4.7** Permutation result for the grouping method used in Table 4.6.

**Table 4.7** Accumulated accuracy of each different machine learning algorithm for predicting the grouping of Healthy or Sick.

| | **con** | **conPCA12** | **conPCA17** | **resd** | **resdPCA18** | **resdPCA24** |
|---|---|---|---|---|---|---|
| **KNN** | 0.79 | 0.77 | 0.77 | 0.79 | 0.77 | 0.77 |
| **DT** | 0.68 | 0.71 | 0.73 | 0.69 | 0.71 | 0.72 |
| **RF** | 0.79 | 0.77 | 0.77 | 0.79 | 0.76 | 0.74 |
| **SVM** | 0.79 | 0.77 | 0.77 | 0.79 | 0.76 | 0.76 |
| **LR** | 0.76 | 0.74 | 0.76 | 0.69 | 0.72 | 0.69 |

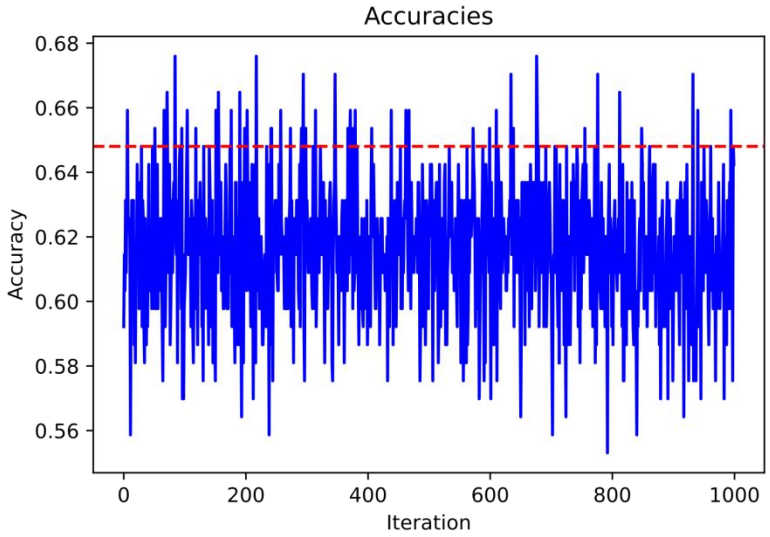



**Figure 4.8** Permutation result for the grouping method used in Table 4.7.

**Table 4.8** Accumulated accuracy of each different machine learning algorithm for predicting the grouping of Underweight, Normalweight, and Overweight.

| | **con** | **conPCA12** | **conPCA17** | **resd** | **resdPCA18** | **resdPCA24** |
|---|---|---|---|---|---|---|
| **KNN** | 0.38 | 0.40 | 0.40 | 0.38 | 0.38 | 0.36 |
| **DT** | 0.27 | 0.43 | 0.38 | 0.32 | 0.32 | 0.36 |
| **RF** | 0.40 | 0.37 | 0.40 | 0.36 | 0.30 | 0.34 |
| **SVM** | 0.40 | 0.38 | 0.40 | 0.37 | 0.33 | 0.39 |
| **LR** | 0.31 | 0.34 | 0.30 | 0.29 | 0.37 | 0.37 |

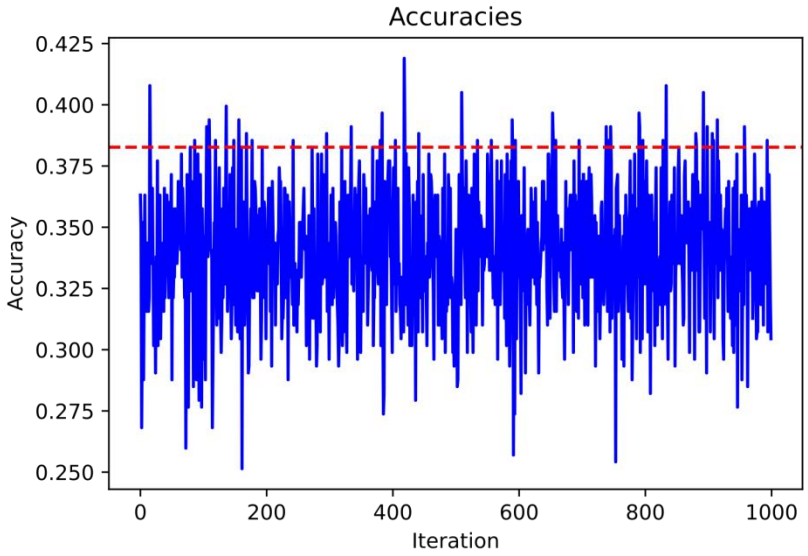



**Figure 4.9** Permutation result for the grouping method used in Table 4.8.

**Table 4.9** Accumulated accuracy of each different machine learning algorithm for predicting the grouping of HO, HN, HU, SO, SN, and SU.

| | **con** | **conPCA12** | **conPCA17** | **resd** | **resdPCA18** | **resdPCA24** |
|---|---|---|---|---|---|---|
| **KNN** | 0.33 | 0.32 | 0.34 | 0.31 | 0.32 | 0.33 |
| **DT** | 0.30 | 0.32 | 0.27 | 0.26 | 0.23 | 0.27 |
| **RF** | 0.26 | 0.28 | 0.30 | 0.30 | 0.31 | 0.28 |
| **SVM** | 0.27 | 0.31 | 0.30 | 0.36 | 0.31 | 0.34 |
| **LR** | 0.16 | 0.26 | 0.23 | 0.20 | 0.27 | 0.27 |

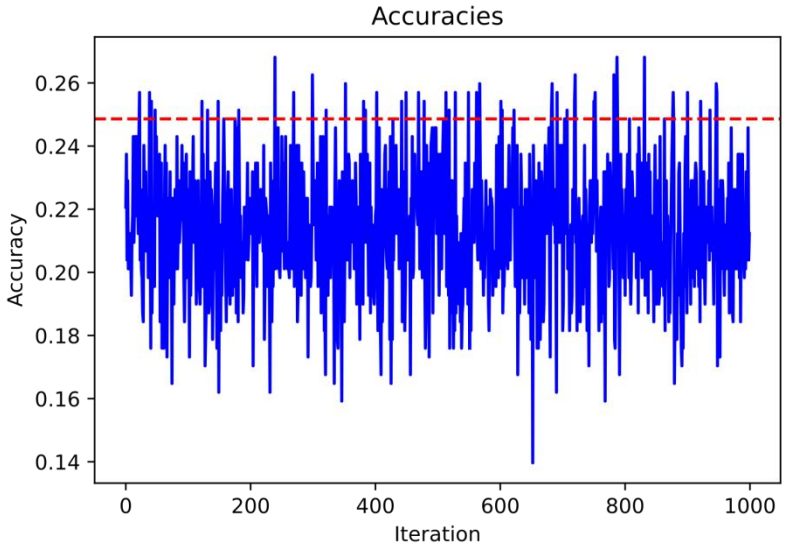



**Figure 4.10** Permutation result for the grouping method used in Table 4.9.

**Table 4.10** Accumulated accuracy of each different machine learning algorithm for predicting the groups of particular grouping method using the unadjusted or batch-adjusted concentrations of correspondingly identified significantly related metabolites.

CGSD: unadjusted concentration of Succinate and Dimethylglycine for predicting groups of Resistant, Resilient, Susceptible, and Early_Dead; CDS: unadjusted concentration of Succinate for predicting groups of Death and Alive; BCDS: unadjusted concentration of Succinate for predicting groups of Dead and Alive using balanced group size for each group; RDP: batch-adjusted residuals of Pyruvic acid for predicting groups of Dead and Alive; BRDP: batch-adjusted residuals of Pyruvic acid for predicting groups of Dead and Alive using balanced group size for each group; CHWD: unadjusted concentration of Dimethylglycine for predicting groups of HO, HN, HU, SO, SN, SU.

| | **CGSD** | **CDS** | **BCDS** | **RDP** | **BRDP** | **CHWD** |
|---|---|---|---|---|---|---|
| **KNN** | 0.36 | 0.80 | 0.64 | 0.80 | 0.48 | 0.32 |
| **DT** | 0.35 | 0.81 | 0.62 | 0.79 | 0.56 | 0.34 |
| **RF** | 0.37 | 0.80 | 0.64 | 0.79 | 0.52 | 0.29 |
| **SVM** | 0.34 | 0.80 | 0.64 | 0.80 | 0.48 | 0.32 |
| **LR** | 0.32 | 0.79 | 0.58 | 0.80 | 0.54 | 0.33 |

Next, machine learning using an unsupervised algorithm called K-means clustering was applied on the metabolome data. To verify the batch effect on metabolite concentrations as determined by the Fitting Linear model, 7 clusters were generated to match 7 batches of pigs using the K-means algorithm and found that the batch was significantly correlated with metabolite concentrations (Figure 4.11A and Table 4.13). In Figure 6B, a clear trend shows that most pigs in each batch have similar metabolomes. The coefficient was even higher when the cluster number was set to 4 (Figure 4.11C, D, and Table 4.11). However, the correlations between the clusters and the phenotype groups from all the grouping methods were very weak (most were lower than 0.1, Table 4.11). Considering the correlation analysis with K-means clusters were groups of phenotypes that had discrete variables, the Spearman correlation was assessed between the 4 clusters and three continuous variables including pre-slaughter weight, number of treatments, and growth rate. As shown in Figure 4.12, there is a negligible trend for the distribution of clusters in terms of any trait tested. The Spearman coefficients were all consistently very low (Table 4.12). Together, the machine learning results using the K-means clustering algorithm revealed that the pre-challenge metabolome is strongly dependent on and correlated with batch factor but only loosely correlated with the post-challenge phenotypes of pigs related to health, survival, and productivity performances.

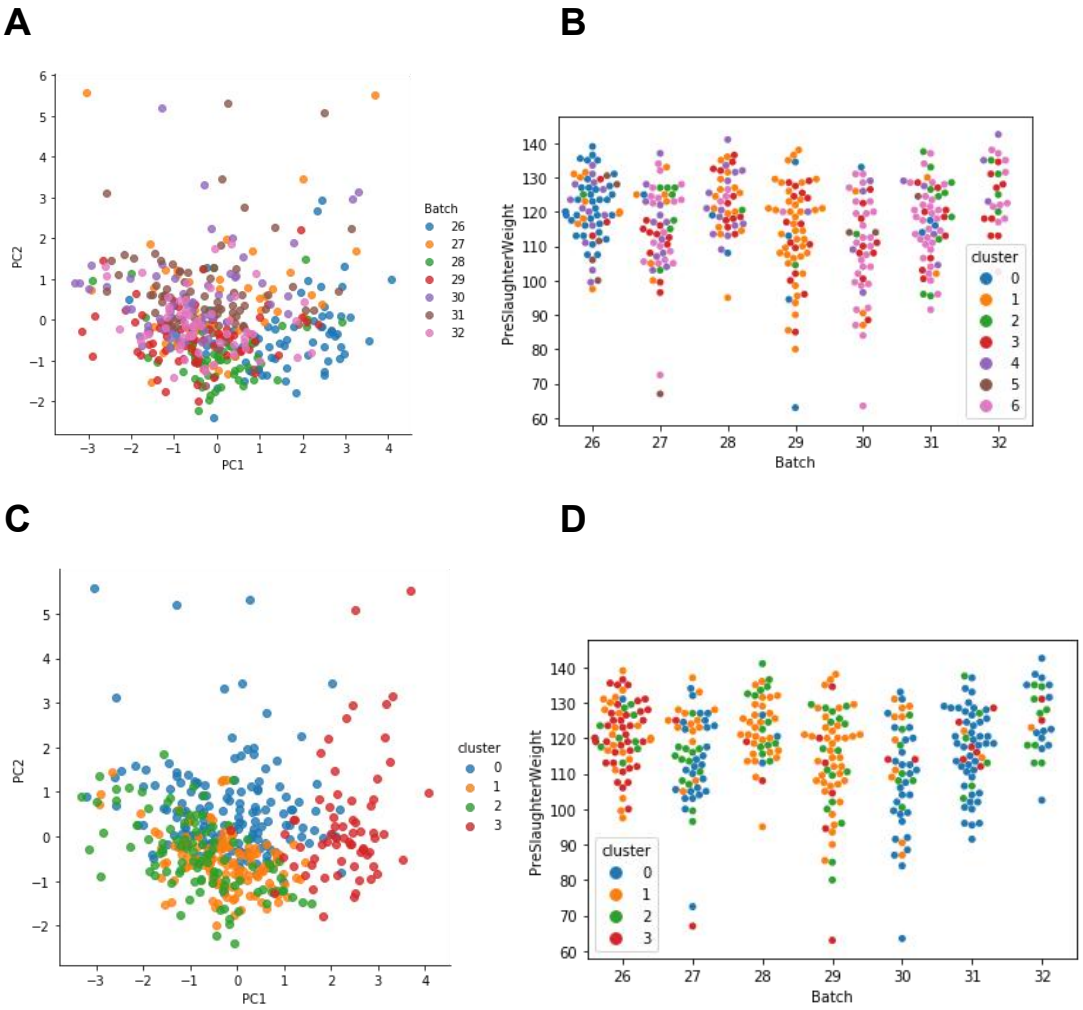



**Figure 4.11** K-means clustering of all the pigs tested based on the prior-challenge metabolomes.

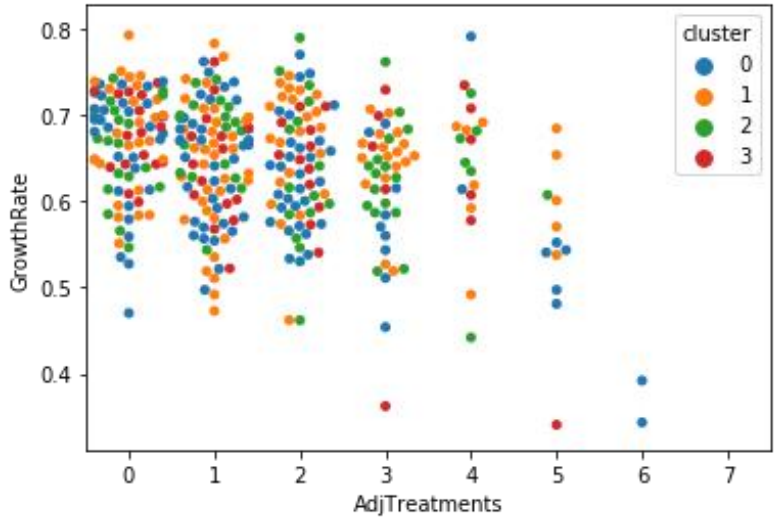
(A and C) Distribution of 7 (A) or 4(C) clusters plotted in two dimensions where the x and y axis indicate the two main components from PCA analysis that represent most variables of the dataset, respectively. (B and D) Distribution of 7 (B) or 4(D) clusters plotted in two dimensions where the x and y axis indicate the batches and the pre-slaughter weight, respectively.

**Table 4.11** Spearman correlation analysis of groups (phenotypes) under certain grouping methods with the corresponding number of K-means clusters.

"Combined" indicates the groups of Resistance, Resilience, Susceptibility, and Early_Dead.

| Grouping Method | Clusters Number | Spearman Correlation Result | |
|---|---|---|---|
| | | Correlation | P-value |
| **Batch** | 7 | 0.38 | 1.23E-17 |
| **Batch** | 4 | -0.42 | 1.03E-20 |
| **Health status** | 2 | 0.11 | 0.04 |
| **Production ability** | 3 | 0.04 | 0.46 |
| **Survival statement** | 2 | -0.07 | 0.13 |
| **Health status + Production ability** | 6 | 0.04 | 0.50 |
| **Combined** | 4 | -0.05 | 0.29 |

A




B

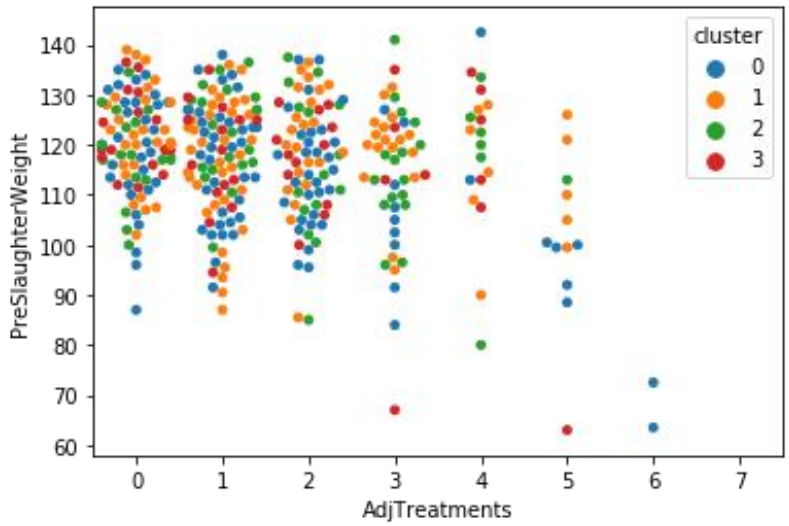



**Figure 4.12** The 4 K-means clusters of tested pigs based on the pre-challenge metabolomes.

(A) Distribution of the 4 clusters plotted in two dimensions where the x and y axis indicates numbers of treatments and growth rate, respectively (B) Distribution of the 4 clusters plotted in two dimensions where the x and y axis indicates numbers of treatments and growth rate pre-slaughter weight, respectively.

**Table 4.12** Spearman correlation analysis of pre-slaughter weight, number of treatments, and growth rate with the 4 K-means clusters, respectively.

| | Spearman Correlation Result | |
|---|---|---|
| | **Correlation** | **P value** |
| **Pre-slaughter Weight** | 0.10 | 0.06 |
| **Number of Treatments** | 0.03 | 0.54 |
| **Growth Rate** | 0.08 | 0.13 |

## 4.4 Discussion

One of our breeding aims is to select pigs with resistance or resilience to the challenge of infectious agents. Those two concepts have distinct meanings but share one common feature about the immune response to pathogens. Besides the narrow sense examples of resistance derived from complete blockage of pathogen entry due to entry receptor mutation or deficiency (Bao et al., 2012; Ren et al., 2012; Whitworth et al., 2016, 2019), successful defense of pathogens after the invasion largely depends on the immune response to clear or suppress the infection. Maintenance and activation of the immune system require energy consumption in high demand, and excessive activation of the immune response may cause appetite loss and muscle wastage (Bird, 2019; Straub, 2017), which is against the livestock production aim. The resilient pigs could maintain good productivity performance with active control of pathogen infection by their immune system, through two possible mechanisms: (1) accomplish relatively efficient immune response in an energy-effective manner; (2) establish a more active and efficient way to gain and save energy. To better understand how the resilient pigs balance the energy distribution between productivity and immune system, and how they outperform other pigs to maintain relatively high productivity performance under active infection, it is important to carefully characterize their metabolic activities.

The metabolome is closely related to biological function (phenotype) in studies of omics. Specific metabolites from biofluid have been used to diagnose or predict the progression of human diseases with examples described in the background part. Plasma metabolites were found to be potential predictive or diagnostic indicators for some chronic diseases of livestock such as ketosis during the perinatal period of dairy cows

(Cao et al., 2017; Lisuzzo et al., 2022; Shahzad et al., 2019). Although metabolites associated with specific diseases obtained in understanding biological processes are widely considered as biomarkers, they might not be applicable in clinical settings since the research purposes and research methods in the clinical settings are not equivalent to those in the laboratory settings. Biomarker identification is not aiming to understand the underlying biological mechanisms but to identify the specificity or sensitivity of an organism to a particular disease or external influence. Thus, a biomarker is an objective, measurable biological feature used to diagnose, monitor, or predict disease risk. However, mechanistic insights can not only direct better treatment, but also support the selection of biomarkers (Zhang et al., 2017). Fundamentally, the goal of biomarkers in metabolomics is to create a predictive model from a collection of metabolites that can be used to classify targets into specific groups (e.g. health group and disease group). Notably, the associated metabolome in most studies of biological processes is after the occurrence of the event, while biological markers usually require the characteristics of metabolites to be identified prior to the onset of the disease.

Thus, I hypothesized that the pigs with different responses to pathogen infection have distinct plasma metabolomes before exposure to pathogens, and metabolome differences could be used to predict the pig's health and productive performance after infection. To test this hypothesis, I determined the metabolome of pigs in our natural disease challenge model before pathogen exposure using the $H^1$ NMR approach. The metabolome is dynamically stable at the steady state, which possibly simplifies the complex effect after infection derived from multiple types of pathogens on energy consumption and production. Indeed, the NMR data revealed that acetic acid had similar

plasma concentrations in pigs from different batches or dams even though there were many other metabolites with endogenously different levels.

The multiplier pigs used in this study were separated in different batches from different breeding companies, they may have various genetic backgrounds which can cause variegated metabolome (Swain-Lenz et al., 2017). The batch factor also involves environments which can also dramatically influence metabolite concentrations (Bermingham et al., 2021; Katz et al., 2021). For these reasons, I first examined the correlation of batch factor with plasma metabolome and found that this complicated factor did significantly impact the majority of plasma metabolites (~97%) in the pigs tested. It was hard to avoid the influence of seasonal and geographic factors although we have maximized the consistency of environmental factors such as temperature. In addition, I also observed that about 87% of plasma metabolites were influenced by the dam, presumably due to the higher sensitivity of the blood metabolite concentrations to maternal inheritance including cytoplasmic inheritance (rather than paternal inheritance) (Joseph et al., 2013). Thus, the batch and dam factors should not be overlooked when analyzing the relationship between pig performance with plasma metabolome.

I used several statistical methods to compare the plasma metabolome with different pig performance phenotypes that were determined after the pathogen challenge. For the NMR-determined metabolites, Succinate (FDR=0.0049) and Dimethylglycine (FDR=0.0212) had significantly higher concentration in the Early_Dead group than that in the Resistant and Resilient groups. Succinate is an intermediate from the Kreb cycle. Interestingly, it was also considered as an inflammation mediator in macrophages and dendritic cells, and it was assumed that inhibiting its receptor on dendritic cells may

prevent exacerbated inflammation (Mill and O'Neill, 2014). Dimethylglycine is a derivative of glycine and was reported as an immune modulator many years ago with the potential to increase both humoral and cellular immunity in humans (Graber et al., 1981). Dimethylglycine was also used to boost immune response in animals including pigs (Bai et al., 2022; Reap and Lawson, 1990). The higher concentrations of these two metabolites in the Early_Dead group may imply a low threshold for immune activation, supporting our hypothesis that the pigs from the Early_Dead group might die from an overactive immune response. However, the different concentrations observed for these two metabolites may be due to difference between the batch factors. Accordingly, I tried to analyze pigs only from the same batch but failed to identify metabolites with significantly different concentrations. Notably, when batch-adjusted metabolite concentrations were applied, the aforementioned significant differences from Succinate and Dimethylglycine were not found (FDR>0.05), and only Pyruvic acid was found (FDR = 0.03) different between the Alive and Dead group. This difference, however, was very small and it is impossible to distinguish most pigs from the Alive and Dead group. Nevertheless, Pyruvic acid is the essential product of glycolysis and substrate of the Krebs cycle for energy supply (Akram, 2014). Thus, the pigs from the Dead group were expected to have a lower concentration of Pyruvic acid, critical to ensure an efficient immune response. This implies that before challenge, the Dead group may not have as robust functionality as the other groups. A recent study uncovered the important role of Pyruvic acid in inflammatory macrophages during the infection of influenza A virus (Abusalamah et al., 2020).

The subtle difference between groups and big group-internal variation of the

identified metabolites suggest that they are not ideal biomarkers to predict pig performance post-pathogen challenge. This may plausibly be due to the complexity of our model system which contains multiple pathogens and the pigs tested were from different batches. The NMR assay (being the sole method used here to determine metabolome) might also contribute to our unsuccessful outcomes. Incorporating data from MS with NMR would provide a much higher resolution to achieve more convincing hits of metabolites in a plasma sample. It should be pointed out that our group recently used a cohort with about two-fold more pigs than I used to explore plasma metabolites that had differential concentration in resilient pigs using the same NMR method and natural disease challenge model (Dervishi et al., 2023). Consistently, batch effect was found to be the most dramatic factor that influenced the concentrations of metabolites in plasma. Intriguingly, only Creatinine was found to have significantly lower concentrations in pigs classified as "dead" and "susceptible" than the "average" ones. Creatinine was also detected in my NMR data but no difference was observed between any groups based on my pig classifications. Possible reasons could be: (1) Different grouping logic (e.g. whether pigs receive medical treatment are regarded as "resilient"); (2) Different adjustment models (I only adjusted the metabolites concentrations from batch effect, the other study generated a linear model considering environment enrichment, age, pen, and litter besides batch); (3) Different cohort size; (4) Statistical methods. I used multiple methods but the other study only used the least Square Means. It is plausible that including as many factors as possible could better interpret the variations for such a complex phenotype as "disease resilience", however, each factor needs to be carefully assessed to confirm their potential effect on plasma metabolite before

artificially they are included in a linear model to prevent inclusion of unrelated errors. Intriguingly, our earlier study using the same combination of factors in the linear model failed to identify significant correlation between any plasma metabolite with health-related traits (Dervishi et al., 2021) which are important aspects in assessing resilient pigs. Thus, in the context of a highly heterogeneous pig cohort, more solid testing of mathematical models and statistical methods are needed in the future to precisely interpret the differences observed in the metabolites and other measured biological molecules.

It is worthwhile to examine the feasibility of predicting the performance and health status in our challenge model based on the information of ensemble plasma metabolome or a major population of metabolites as metabotypes of pigs. I thus applied machine learning using different commonly-applied algorithms. However, none of these algorithms could well predict defined phenotypes in response to pathogen challenge. It appears that the pre-challenged (NMR) metabolome is not enough to characterize the potential of pigs to defend against infection or maintain productivity performance in the context of multiple disease challenges. Alternatively, the model system could be simplified by reducing the number of pathogen types or even focus on a scenario of a single pathogen. Another direction might be detecting the metabolome of relevant organs that are mainly targeted by specific pathogens, e.g. lung and tonsil for PRRSV infection (Beyer et al., 2000). However, this might not be useful for screening biomarkers to predict performance since the sampling of lungs and tonsils is impracticable during the breeding process. One solution to make use of plasma metabolome for pig response prediction could be multi-omics analysis, as the study of metabolomics is usually

combined with other omics (Fan et al., 2006; Hirai et al., 2004; Verhoeckx et al., 2004; Wang et al., n.d.2018) see Chapter 5.

## 4.5 Summary

To sum up for this chapter, a total of 460 pigs were included to test the feasibility of using plasma metabolome analysis in healthy pigs to identify marker metabolites to predict their post-challenge phenotypes particularly disease resilience. The pigs were routinely grouped into Resistant, Resilient, Early_dead, and Susceptible as chapter 3, and the one collection of blood samples were acquired before the start of pathogen challenge. By utilizing NMR technique, 38 metabolites were successfully detected in the plasma samples. Notably, Succinate and Dimethylglycine were found with significantly higher concentrations in the Early_dead group than the other three groups (except Dimethylglycine in Early_dead vs Susceptible). By further dissecting the health and productivity traits, there was a positive correlation between the pre-slaughter weight and Dimethylglycine concentration in healthy pigs but not the sick ones. However, the Fitting Linear Model demonstrated that the concentrations of most detected metabolites were influenced by batch factors including pig batch, dam, and sire. After batch adjustment on the metabolites concentrations, only Pyruvic acid was found with significantly higher concentration in surviving pigs than the dead one. A series of machine learning models was also tested to explore the possibility of using global metabolome data to differentiate pig groups, but prediction accuracy was far from ideal in all the models. These findings suggested that plasma metabolites in healthy pigs are very sensitive to batch effects and largely not able to differentiate between groups of animals. It may not be used to identify

valuable markers to select pigs with potential favored phenotypes such as disease resilience, but it may provide insights to understand the correlation with common traits such as pig survival after disease challenge.

# CHAPTER 5:
# General discussion and future directions

## 5.1 Summary of findings in this study

The overall aim of this study was to test different types of biomolecules as biomarkers to predict pig disease resilience. Plasma acute phase proteins (APPs), whole blood mRNAs, and plasma metabolites from healthy pigs were each interrogated for their feasibility to predict pig health and productivity in response to multiple pathogens in our natural disease challenge model.

The study in chapter 2 focused on three common APPs in pigs and their potential applications as biomarkers in prediction of pig disease resilience. Haptoglobin (HP) and C-reactive protein (CRP), but not α1-acid glycoprotein (AGP) showed dramatically up-regulated concentration in plasma upon exposure to pathogens in the natural disease challenge model (For Blood 3/Blood 1, HP~20 fold, CRP~5 fold), which validated them as positive APPs. Even though AGP concentration was relatively stable in response to the disease challenge, its concentration prior to the challenge was the lowest in resilient pigs ($P<0.01$ compared to Group 1 and 6, although not significantly lower than group 2, 3, 4, $P>0.05$), making it a putative marker for disease resilience. The CRP in resilient pigs exhibited a different pattern of response upon disease challenge compared with the other two APPs tested. Early after disease challenge, the concentration of CRP was found with a slight up-regulation (~2 fold) but followed by a sharp increase (~10 fold) at the later stage, which could potentially be used to assess resilience with ongoing infections. None of the APPs tested showed a significant correlation with peripheral blood traits before

challenge. In contrast, the early induction of HP was positively correlated with the concentration of white blood cells (WBC) and the late induction of CRP was positively correlated with the concentration of lymphocytes (Lym), implying these two APPs respectively accompanied the mobilization of innate and adaptive immune response. Similarly, no significant correlation was observed between the APPs prior to the challenge whereas the concentrations of the three APPs at the post-challenge stage were negatively correlated with 1st treatment age but positively correlated with treatment frequency. AGP concentration before the challenge was negatively correlated with total ADG and FI while the other two APPs showed a similar correlation only in the plasma after challenge. Overall the results in Chapter 2 suggest that pre-challenge AGP has the potential to predict disease resilience phenotype. HP and CRP also showed potential value in terms of assessing pig health status and guiding disease management (See further discussion in Section 5.3 below).

Chapter 3 examined whether specific gene expression information could be exploited to predict pig disease resilience. Transcriptomic profiles from whole blood were determined by RNA-Seq in healthy pigs before they were exposed to pathogens in the natural disease challenge model. Pigs were subsequently grouped into “Resistant”, “Resilient”, “Early_dead”, and “Susceptible” based on their performance after disease challenge. Differential expression (DE) analysis was performed to identify candidate genes or transcripts with predictive values for disease resilience. In the initial cohort of 58 pigs, two transcripts derived from the *MSTRG.32803* (on unplaced chromosome AEMK02000452.1) and *MSTRG.21242* (on chromosome 4) genes mapped to *IgC* and *SLAMF9* genes, respectively, and were found to be significantly up-regulated in pigs

from the Susceptible group (q-value<0.1) compared to the rest of the pigs. eQTL analysis for those two transcripts (MSTRG.32803.53 and MSTRG.21242.14) identified expression-correlated SNPs located on chromosomes 12 and 13, respectively. DE analysis also revealed that some positive immune-regulatory genes or transcripts were found with a higher level of expression (P<0.05) in the Resistant group (i.e. *CCL5, ATRN, LYN*) and Resilient group (i.e. *CXCR5, IL4R, PPP3CA, TNFSF8*) compared to those in other specific groups or the rest of all the pigs. Gene set enrichment analysis was performed to better understand the biological contexts of the DE genes found in each group. Consistent with the DE analysis, immune response-related pathways were up-regulated in the Resistant group while those related to negative regulation of response to stimulus were more enriched in the Susceptible group. Notably, the Resilient group was found to down-regulate pathways associated with cytoskeleton organization and suggested lower energy consumption in terms of the multiple functions (e.g. migration, phagocytosis) of immune cells that the cytoskeleton is involved in. It is thus possible that resilient pigs could spare more energy from immune defense to be used in productivity purposes. This finding supports the hypothetic mechanism to explain why resilient pigs could maintain optimal productivity in response to pathogen infection. To validate the findings from the 58 pigs, a larger cohort of 209 pigs from an independent experiment cycle were tested for the same analysis. Even though the DE genes and transcripts enriched from the first cohort were not ranked as the top hits in the larger cohort, the newly identified DE transcripts (i.e. *PTPRC, IFI44L*) also supported a superior baseline immunity of the Resilient group compared to the “Early_dead” and “Susceptible” groups. Some transcripts (i.e. *CAT, WDR26*) associated with anti-stress response were found up-

regulated in the Resilient group. A proposed model arising from the results of the two cohorts supported a hierarchical baseline immunity among the four groups of pigs which were ranked as follows: Resistant≧Resilient>Early_dead>Susceptible. The RNA-Seq analysis of the larger cohort included two additional sample collections (blood 3 and 4) after disease challenge. When comparing DE genes before and after challenge, *CXCL10* was up-regulated in all of the pigs two weeks after challenge. By comparing DE genes in each group, it was suggested that the Resilient group tended to induce a more "economical" response to pathogen challenge through limiting the energy consumption of immune cells but increasing the expression of anti-stress genes to maintain homeostasis. Moreover, there is evidence (elaborated below) supporting the possibility that the Early_dead group may die from over-activation of immune response instead of uncontrolled pathogen infections.

The last part of this study was the investigation of the plasma metabolome in pigs before they were challenged, and the assessment of using these metabolites as predictive biomarkers for disease resilience. Both the resilient and resistant pigs were observed with a significantly lower plasma concentration of succinate and dimethylglycine compared to that of the Early_dead group; a significantly lower concentration of dimethylglycine was also found in the Resistant group compared with that of the Susceptible group. In addition, dimethylglycine had a significantly higher plasma concentration ($P<0.001$) in the healthy and overweight (HO) group than that in the healthy and underweight (HU) group whereas no difference was observed between the sick and overweight (SO) and sick and underweight (SU) groups. Both succinate and dimethylglycine are metabolites involved in promoting the immune system (see Chapter 4 discussion) however, the concentration

differences observed may likely be due to batch effects. Removal of the batch effect by adjusting the concentration of the metabolites left only pyruvic acid as being significantly different (FDR = 0.03). Lower concentration of pyruvic acid was observed in the Early-dead group compared with the other groups. Supervised machine learning was applied, aiming to use the integrated metabolomic profile to predict disease resilience; prediction accuracy was however sub-optimal in all the algorithms tested. The unsupervised algorithm did not perform well in clustering pigs correlated with single or combined traits. Altogether, these methods highlighted the impact of the batch factor on the variations of the metabolome in the experimental pig cohort and further research may be warranted in terms of the utility of metabolites as biomarkers (See more discussion in Section 4.4).

To summarize the finding from the whole thesis, a hypothetical model was generated to compare the differences observed in the defined pig groups in terms of the different biomolecule types investigated (Figure 5.1).

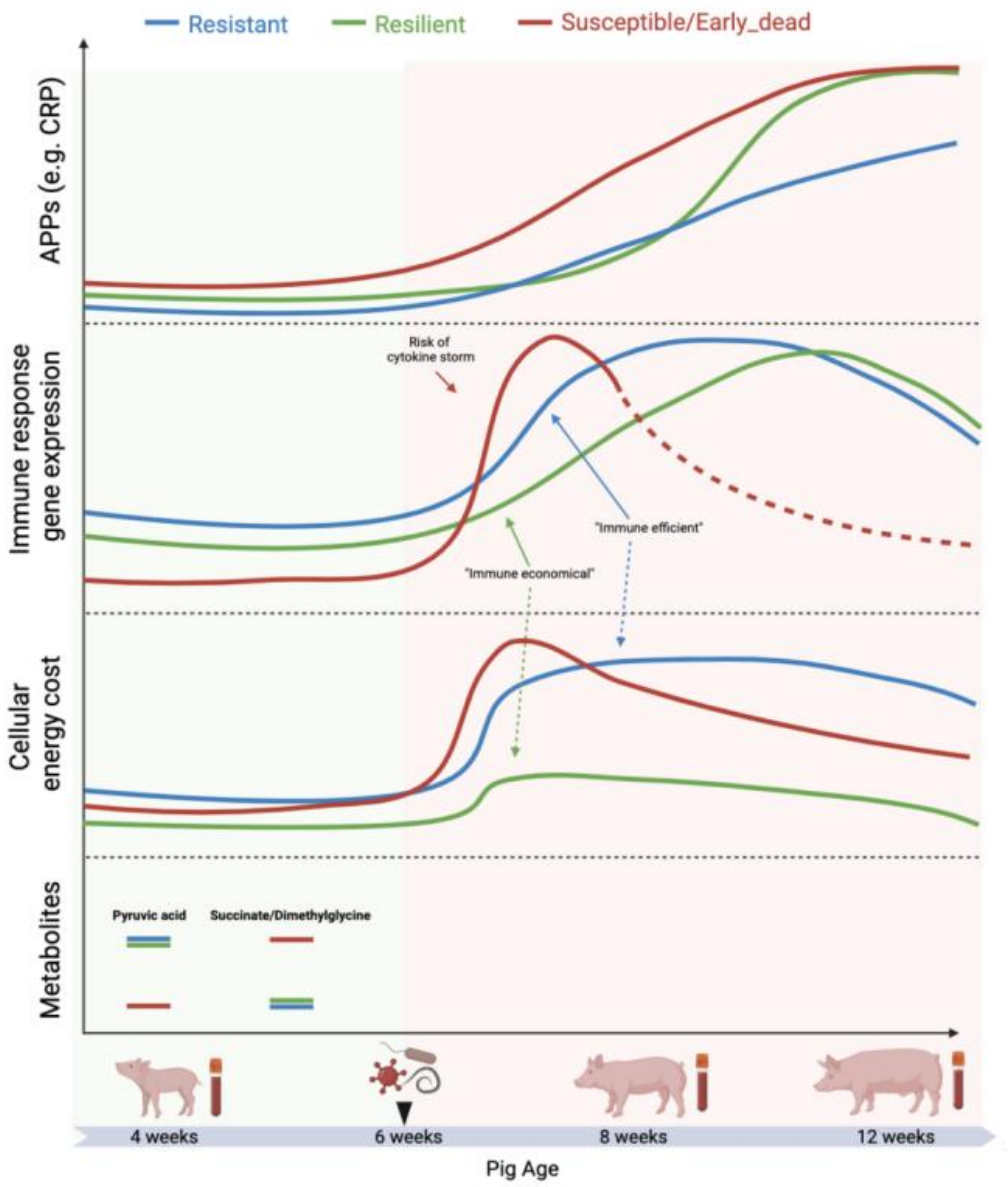



**Figure 5.1** Hypothetical model summarizing the patterns of the three molecular modalities determined in this thesis in different groups of pigs before and after the natural disease challenge.

## 5.2 Limitations of the experimental model and design

Pig resilience is a trait dependent on multiple factors, and as defined here mainly involves the ability to control infection without compromising productivity. There is still no standardized assessment or measurement of pig resilience in the pig farming field. In most studies, the level of infection was determined or estimated by the load of pathogens in the experiments, or alternatively by a health score provided by the veterinarian. However, the production performance of pigs may vary even though they are exposed to the same infection pressure, especially for their growth performance. Conversely, pigs with similar production performance may have different pathogen loads. Since the core feature of disease resilience is the maintenance of productive performance with exposure to diseases, pre-slaughter weight, treatment frequency, and mortality were mainly used to efficiently classify the pig phenotypes in response to disease challenge. More efforts have been made by others to characterize and broaden the spectrum to assess the resilient phenotype. Recent studies with the challenge model we used have started to include health score, growth rate, feed efficiency, and various carcass traits (Bai & Plastow, 2022; Lim et al., 2021). The health score tends to be limited for assessing disease response as it is subjective. For pig production traits, it is helpful to also measure feed intake (FI) besides pig weight since a recent study from our collaborators demonstrated an association between pig resilience and FI (Putz et al., 2018). Furthermore, the economy of a pig farm is usually determined by the cost of feeding, treatment of sick pigs, and the profit of slaughter weight in the market. It would be beneficial to improve the pig farm economy by taking into accounts these factors, e.g. FI to select resilient pigs. Therefore,

to reasonably define pig resilience, it is optimal to include multiple production and health-associated traits as proxy traits in future studies. These traits include but are not limited to health score, pathogen load, FI, and ADG during pathogen challenge. Meanwhile, the heritability of those known and candidate proxy traits should not be overlooked, especially when disease resilience is represented based on multiple traits. Caution on the selection of trait combinations needs to be taken care of to ensure the breeding aims based on these traits will be desirable to both health and production performance.

To simulate natural disease challenges for pigs in commercial farms, we established a system called the natural disease challenge model. This model allowed us to assess the defined phenotypes, reactive to pathogens in an environment, simulating the actual situations including infections by multiple pathogens and cross-transmission among pigs. Most importantly, we aimed to use the animal information collected from this model prior to the pathogen challenge to predict the health status and performance (phenotypes) of pigs under the pathogen challenge. Thus, the samples used in this experiment were collected within the first week of the pig's arrival from the breeding company at the healthy quarantine nursery to avoid pre-existing contamination in pigs before the following challenging step.

Commercial pig production typically takes place in areas with dense pig populations. This as expected results in unsatisfactory health status as biosecurity is difficult to maintain in areas that are densely populated with pig farms. Commercial pigs usually face various types of pathogen infections resulting in disease syndromes or complexes, e.g. PRDC. Thus, the overarching aim of this study was to identify biomarkers that can be

used to screen pigs prior to the infection that will subsequently respond better to this mixed pathogen challenge. Indeed, pigs' differential responses to pathogen challenge may involve many factors including but not limited to genetic or genomic variations, environment, supplier company, animal's physical traits (e.g. age, weight), and early-stage stress (e.g. weaning, transportation) when they are exposed to pathogens. In the actual pig industry, it is inevitable to introduce those variables. The leading aim of this study was to explore bio-molecules as biomarkers to more efficiently differentiate supplier pigs that are predicted to be disease resilient when they encounter pathogen challenge. In particular, we aimed to explore the potential of APP as biomarkers for pig health status. Research published in 2013 assessed the response of several main APPs including HP and CRP of pigs in a model of H1N1 SIV and *Pasteurella multocida* co-infection (Pomorska-Mól et al., 2013). In this study, they considered a secondary bacterial pneumonia being a frequent complication during SIV infection (Pomorska-Mól et al., 2013). Unfortunately, the authors inoculated SIV and *Pasteurella multocida* at the same time by intranasal inoculation and so were not able to differentiate the main infection and concomitant infection. In another APP response study, the group employed a different experimental infection model to investigate HP and CRP response to *Trichinella* species parasites (Gondek et al., 2020). However, they observed mild change of those APPs to the infection and only HP significantly increased two months after the infection had occurred. Conversely, our natural disease challenge model introduced infections through pig-to-pig transmission, which provides natural and continuous challenges to closely mimic the actual case of commercial farms. Specific APP may presumably respond in various patterns and exhibit different kinetics to different types of

pathogens. Our model included prevalent pathogens in pig farms including viruses, bacteria and parasites to study pig APP response, to better reflect the complicated infection case. The advantage of our model could analogously apply to the transcriptome and metabolome experiments. Most of the studies focusing on investigating the association of transcriptome profiles with pig response to pathogen infections were conducted with artificial infection models. The pathogens in these studies include PRRSV (Wilkinson et al., 2016), ASFV (Jaing et al., 2017), foot-and-mouth disease virus (Lv et al., 2018), salmonella (Kommadath et al., 2014), and mycoplasma (do Nascimento et al., 2018). The resulting data may uncover some important genes directly or indirectly involved in infection onset or pathogenesis of a particular pathogen, but these data may not represent the scenario of cross infections. A limitation of the natural disease model used in this study is that it is very difficult to track all the pathogen loads owing to the complexity of concurrent infections and pathogen/host interaction. Thus, it is improper to use pathogen load to assess pig response when employing this system. It is also impossible to ensure that all the pigs tested are exposed to the same level of pathogen stress by using seeder pigs to spread infection naturally. This could potentially cause large variations in disease initiation and development, and easily result in type I errors. It would be optimal to find a way to synchronize the exposure to pathogens for the experimental pigs in the natural disease challenge model, for example, raising the possibility of the naive pigs in contact with the seeder pigs by increasing the percentage of seeder pigs. However, the disadvantage due from this would be excessive pathogen stress compared to practical origin of a disease outbreak. Another limitation of the natural disease model is the complexity of infection transmission or pig infectivity. This may

cause variations of pathogen stress in different herds in terms of the transmission from infected pigs by the initial seeder pigs. Even though there is no direct evidence that whether resilient pigs have lower infectivity compared to suceptible pigs, a recent study from Doeschl-Wilson group revealed no supportive evidence to limit PRRSV-2 transmission in pigs with the WUR0000125 PRRS resilience allele (Chase-Topping et al., 2023). The improvement of the natural disease model would be achieved with better understanding of the pathogen pathogenesis and interaction with pigs.

## 5.3 Future directions and challenges

With the progress of omics techniques and analytic methods, the main purpose of this study was to examine the feasibility of using different types of biomolecules information from blood to predict pig disease resilience. These biomolecules with the potential to predict disease resilience phenotype can serve as "indicator" or “alternative” traits (or biomarkers) and should be easily measured or assessed in pig production units. Any biomarkers identified from the data generated from the natural disease challenge model (see also Table 5.1) should be carefully validated for their correlation with pig resilence traits using batches of pigs from different sources. eQTLs can be then determined and those new biomarkers could be potentially integrated to improve the genomic selection program. However, before implementing these newly-identified biomarkers in pig genomic selection, it is also essential to ascertain their heritability and correlation (both genetic and phenotypic) with the known, validated surrogate traits of resilience. Once heritability and correlation with other resilience traits is determined and validated, a large enough training dataset can be established using pigs with both

genotypic and phenotypic (including the known surrogate traits and new biomarkers) information to estimate the marker effect. Another training panel with SNP information only collected in healthy animals can be used to predict the genomic breeding values of each individual using the estimates (Samorè and Fontanesi, 2016), and these predicted genomic breeding value can be validated with the resilience phenotypes acquired after disease challenge. Thus, the resilient pigs can be quickly identified and selected at early age based on their genomic selection index integrating the SNP allelle effects across the whole genome.

Overall in this study, the omics data only included blood transcriptome and plasma metabolome, and the analysis for different types of biomolecules were separated, except for eQTL analysis which combined transcriptomic and genomic information. The recent work from our collaborators made a great attempt to integrate whole blood transcriptome and high-density genotyping to screen cis-eQTLs in the pig genome (Kramer et al., 2022). This provided valuable resources to link the gene expression difference correlated with economical phenotypes including the disease resilience with associated genomic variations which are broadly used in livestock selection and improvement.

Recent work from our collaborators also reported promising results of using plasma proteome of healthy pigs to predict subsequent response to disease challenge (Chen et al., 2023). This provides important additional information for this study as only three APPs were investigated in the plasma samples. With more and more multi-omics data collected for the pigs in our natural disease challenge model, how to integrate these datasets and extract meaningful information could be one of the major challenges. A very recent study on a large healthy human cohort (Xu et al., 2023) could be a great example for the future

direction. This study performed machine learning to leverage the connection of genetic variations with the multi-omics data. They trained the “genetic score” comprised of one to thousands of genetic variants from a single to several linkage disequilibrium blocks, and used these genetic scores to successfully predict more than ten thousand molecular traits from the multi-omics data. In addition, they utilized the genetic score to evaluate the genetic control and disease association of well-characterized biological pathways. With the phenome-wide correlated molecular traits, they could implicate the association of genetic variations with disease risk. Learning from this work, multi-omics data from the same pig could be integrated, including genomic, transcriptomic, proteomic, metabolomic, and microbiome data, and then use machine learning, even deep learning, to build a network of bio-information with pig molecular and farm-measured phenotypic traits to better understand the pig's health status and productive performance, which could potentially aid in predicting their disease resilience prior to disease challenge and be used in pig genomic selection. Meanwhile, several potential challenges must be addressed to fulfill the objectives of this study. One such challenge, evident from the findings in this thesis, is the limited sample size, which may impede the identification of disease resilience-associated traits. To mitigate this limitation, our laboratory is proactively obtaining new samples and generating multi-omics data, significantly expanding the datasets of the natural disease challenge model. Currently, samples from more than 3,500 animals have been collected, all of which have been carefully stored for future analysis. Another significant challenge would be the restricted genetic resources within the pig industry because of highly directed breeding strategies. This has likely led to the loss of potential genetic variations associated with disease resilience. However, advances in new

molecular techniques such as CRISPR may offer promising avenues to address this issue. By gaining a better understanding of the molecular mechanisms governing pig responses to pathogens and extracting valuable natural variations from different pig populations and even more distant wild animals, the limitations posed by the narrow genetic pool might be circumvented. Another challenge arises from the lag between experimental and analytical methods employed in pig studies as compared to human studies. It needs considerable efforts to develop compatible methods tailored specifically for pigs. In order to comprehensively elucidate our endeavors, which involve a series of omics and molecular assays in our natural disease model to characterize the genetic and phenotypic features of disease resilience, a table is compiled summarizing the relevant publications and their key findings (Table 5.1)

**Table 5.1** Summary of the publications using the natural disease challenge model in study of disease resilience. (The findings listed were based on the pigs upon exposure to the polymicrobial challenge if there is no further indication.)

| Publications/year | Biomolecules types/ Study approaches | Key findings |
|---|---|---|
| (Cheng, Fernando, et al., 2022) | DNA/ GWAS | Major histocompatibility complex and loci on some immune-related genes are associated with disease resilience-related traits including growth rates, treatment and mortality rates |
| (Cheng et al., 2020; Cheng, Lim, et al., 2022; Putz et al., 2018) | DNA/reaction norm analysis, GBLUP | Health scores and treatment rates, mortality, and growth rates re genetically highly correlated with each other and can be used to quantify disease challenge loads; Day-to-day variation in feed intake duration has low to moderate heritability but has strong genetic correlations with mortality; Day-to-day variation in feed intake or duration could be novel indicators for general resilience to various types stressors including disease |
| (Cheng et al., 2021) | DNA/random regression model | Water intake duration and number of visits have moderate genetic correlations with treatment and mortality rates |
| (Jeon, Cheng, et al., 2021) | DNA/ASE analysis | A SNP which is associated with the host response to PRRSV has high LD with *GBP5* gene; This SNP is significantly associated with average daily gain and treatment rates |
| (Bai et al., 2020, 2021) | Peripheral blood cells, DNA/GWAS, GBLUP | Resilient pigs have higher increase of Lym concentration early after challenge; Neu concentration and two red blood cell traits are highly genetically correlated with growth and treatment rates; Some immune-related SNPs are significantly associated with Eos, Mono, red blood cell and platelet traits before or after polymicrobial challenge |

| | | |
|---|---|---|
| (Bai et al., 2022; Bhatia et al., 2022) | Peripheral blood cells, immunoassays, DNA/GBLUP | Before polymicrobial challenge, some complete blood count and immunoassay traits have significant genetic correlations with resilience performance traits; Some phagocytosis traits of pigs before polymicrobial challenge have moderate/high genetic correlations with High Immune Response traits and some blood cell percentage |
| (Jeon, Gilbert, et al., 2021) | Peripheral blood mononuclear cells (PBMCs), DNA/GBLUP | Before polymicrobial challenge, PBMC proliferation in response to oncanavalin A, phytohemagglutinin, and phorbol myristate acetate show favorable trends to be genetically correlated with disease resilience-related traits but are not statistically significant |
| (Chen et al., 2020) | Natural antibodies and total IgG in plasma, DNA/GBLUP, GWAS | Before polymicrobial challenge, titers between the same natural antibody isotypes have moderate to high phenotypic and genetic correlations. Higher levels of natural antibodies are genetically correlated with lower mortality and treatment rates. Only after polymicrobial challenge, survived pigs had slightly higher levels of natural antibodies and total IgG than the dead pigs. Natural antibodies titers are associated with several genomic variations that include immune-related genes |
| (Schmied et al., 2018) | Antibody and cell mediated immune response/general linear model | Before polymicrobial challenge, pigs found with high immune responses have low mortality and longer survival post challenge than the low and average responders |
| (Kramer et al., 2022; Lim et al., 2021, 2022) | Whole blood RNA, DNA/eQTL mapping, DE analysis,Gene set enrichment analyses | Before polymicrobial challenge, chromosome 12 was enriched for cis-eQTL based on its relatively smaller size compared to other chromosomes; Gene expression involved in innate and adaptive immunity is genetically correlated with resilience-related traits before and after polymicrobial challenge |

| | | |
|---|---|---|
| (Chen et al., 2023) | Plasma protein, DNA/ GBLUP, Gene set enrichment analyses | Before polymicrobial challenge, some protein abundance has phenotypic association with disease resilience-related traits; Several proteins including IgG heavy chain and Complement component C9 are genetically correlated with disease resilience-related traits; These associated proteins are enriched within many immune-related pathways |
| (Dervishi, Bai, Cheng, et al., 2023; Dervishi et al., 2021) | Plasma metabolite, DNA, Peripheral blood cells/ Metabolite set enrichment analysis, GBLUP | Before polymicrobial challenge, pig batch and environmental enrichment significantly impact plasma concentrations of metabolites; Creatinine, L-α-aminobutyric acid, oxoglutarate, dimethylglycine, betaine, L-methionine and L-serine concentration is genetically correlated with some resilience traits before or after polymicrobial challenge. |
| (Dervishi, Bai, Dyck, et al., 2023) | Plasma metabolite, DNA, Peripheral blood cells/ BLUPF90 analysis | Before polymicrobial challenge, proline concentration was genetically positively correlated with hemoglobin concentration and L-tyrosine was negatively correlated with mean corpuscular hemoglobin; Genetic variations of a metabolites concentration is polygenic but not associated with a single region; *PLA2G3*, *DMGDH*, *BHMT*, and *BHMT2* are identified as the candidate genes for L-serine, dimethylglycine, betaine, and L-methionine (previously found to be positively correlated with pig growth rate) concentration variations. |

Many population-based studies from both human and livestock used bio-molecule information extracted from the whole blood. This makes sense in the context of the global assessment of host response, and blood sample collection is simple, fast, and cost-effective. However, some specific bio-molecules are derived from intracellular compartments and specific cell types, in such cases, the relevant biological information, such as transcriptome, from whole blood may be diluted and hard to sort out. Disease resilience is a complex phenotype that may involve coordination of multiple cells types, and the responsible cell types contributing to induce or maintain this phenotype might be changeable at different stages of infection or stress levels. Genes may have contrasting levels of transcription in different cell types. This could be a possible explanation for finding that not many DE genes were identified from the transcriptome of resilient pigs compared with others before disease challenge. Single-cell RNA-Seq (scRNA-Seq) has developed quickly in recent years and is able to efficiently dissect the transcriptomic information from specific cell types. scRNA-Seq has also been quickly applied to pig studies to power the characterization of cell components in different tissues (Cai et al., 2023; Han et al., 2022; Herrera-Uribe et al., 2021; Zhang et al., 2021). For example, scRNA-Seq was utilized recently to profile the macrophages in pigs infected by African swine fever virus, which provided novel insights to viral pathogenesis and anti-viral strategies (Zheng et al., 2022). Therefore, scRNA-Seq is a potentially useful tool that could be performed with the samples from the natural disease challenge model to power the characterization of cell types and genes associated with the disease resilience phenotype in specific cell types. In Chapter 2, AGP, which showed relatively low concentration in resilient pigs before natural disease challenge, was found to be inversely

correlated with monocyte concentration in the blood early after challenge (Figure 2.4). The RNA-Seq result in Chapter 3 (Figure 3.6) showed that *IL4R* was up-regulated in the resilient pigs before challenge. Coupled with the fact that IL-4 can polarize monocytes to differentiate to anti-inflammatory M2 macrophages in tissues (Hao et al., 2017), this may imply that monocyte migration from blood to tissues could be indicated by the level of AGP concentration and *IL4R* expression in monocytes or macrophages could be used as an indicator for tissue repair and immune resolution phase. However, *IL4R* is also expressed in B cells and T cells and plays a crucial role in cellular immunity (Nelms et al., 1999). In this case, single-cell RNA-Seq would potentially be a powerful tool to dissect the gene expression in different types of immune cells to resolve whether a gene expression change is cell type-specific.

The recent use of genomic tools has definitely made great contributions to livestock improvement especially for difficult and expensive to measure traits. However, genomic variation can only partially explain the observed phenotypic differences. One additional source of variation is the epigenome. The epigenome contains heritable information in the form of chemical modifications on genomic DNA that can modulate the expression of specific genes without altering DNA sequence. The epigenome is sensitive to both internal and external perturbation such as stress, nutrition level, pathogen infection, and climate (Pértille et al., 2019; Corbett et al., 2021; Baker et al., 2020). In the past few years, epigenomics is gaining more attention in livestock studies for exploring novel strategies to enhance livestock improvement and welfare (Ibeagha-Awemu & Zhao, 2015; Nery da Silva et al., 2021). Thus, besides genomic variations, epigenomic regulation may be also involved in establishment of disease resilience and could play important roles in

contributing to the endogenous variations of gene expression that lead to the phenotypic differences in transcriptome, proteome, and metabolome of pigs before or after exposure to pathogens. The epigenome information consists of various aspects such as DNA methylation, histone modification, chromatin structure, and gene regulation by non-coding RNA. Using the assay for transposase-accessible chromatin with sequencing (ATAC-Seq) can help determine the landscape of the epigenome to help determine the regions of the genome that are available for gene expression by determining the genome-wide DNA accessibility information (Buenrostro et al., 2015). ATAC-seq does not need specific antibodies for pigs and could be simply applied in future studies to gain insights into the epigenomic variations in different groups of pigs, and also provide evidence for the regulation of the DE genes observed from RNA-Seq.

## 5.4 Concluding remarks

Based on the efforts to test the potential of using acute phase proteins, whole blood transcriptome, or plasma metabolome to identify biomarkers to predict pig disease resilience, it seems impossible to use one technique or one molecular type to accurately distinguish the pigs with favorable phenotypes. To circumvent the difficulty in understanding and predicting the complicated phenotypes such as disease resilience and resistance, multi-omics could potentially be the most powerful tool. Broader aspects of omics data and more solid integrative analyzing pipelines are needed in the future to better characterize and interpret the association between bio-molecules and disease resilience phenotype. With identifying more biological traits associated with disease resilience and the development of simple and fast detecting technologies, there are still

opportunities to find easy-to-measure and sensitive biomarkers to predict this valuable phenotype (Bai & Plastow, 2022). Various biomarkers assessed in each pig could be integrated as the above mentioned “genetic score” to be used as routine evaluation of resilience to guide genomic selection.

# References:


Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdisciplinary Reviews. Computational Statistics*, *2*(1), 97–106.

Abella, G., Novell, E., Tarancon, V., Varona, L., Pena, R. N., Estany, J., & Fraile, L. (2019). Identification of resilient sows in porcine reproductive and respiratory syndrome virus–infected farms. *Journal of animal science*, 97(8), 3228-3236.

Abul Abbas, M., Lichtman, A., & Shiv Pillai, M. (2019). *Basic Immunology: Functions and Disorders of the Immune System, 6e: Sae-E-Book*. Elsevier Health Sciences.

Abusalamah, H., Reel, J. M., & Lupfer, C. R. (2020). Pyruvate affects inflammatory responses of macrophages during influenza A virus infection. *Virus Research*, *286*, 198088.

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., ... & Venter, J. C. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science,* 252(5013), 1651-1656.

Agresti, N., Lalezari, J. P., Amodeo, P. P., Mody, K., Mosher, S. F., Seethamraju, H., Kelly, S. A., Pourhassan, N. Z., Sudduth, C. D., Bovinet, C., ElSharkawi, A. E., Patterson, B. K., Stephen, R., Sacha, J. B., Wu, H. L., Gross, S. A., & Dhody, K. (2021). Disruption of CCR5 signaling to treat COVID-19-associated cytokine storm: Case series of four critically ill patients treated with leronlimab. *Journal of Translational Autoimmunity*, *4*, 100083.

Ahnach, M., Zbiri, S., Nejjari, S., Ousti, F., & Elkettani, C. (2020). C-reactive protein as an early predictor of COVID-19 severity. *Journal of Medicine and Biochemistry*, *39*(4), 500–507.

Akram, M. (2014). Citric acid cycle and role of its intermediates in metabolism. *Cell Biochemistry and Biophysics*, *68*(3), 475–478.

Albers, G. A., Gray, G. D., Piper, L. R., Barker, J. S., Le Jambre, L. F., & Barger, I. A. (1987). The genetics of resistance and resilience to Haemonchus contortus infection in young merino sheep. *International Journal for Parasitology*, *17*(7), 1355–1363.

Alexopoulos, C., Kritas, S. K., Kyriakis, C. S., Tzika, E., & Kyriakis, S. C. (2005). Sow performance in an endemically porcine reproductive and respiratory syndrome (PRRS)-infected farm after sow vaccination with an attenuated PRRS vaccine. *Veterinary Microbiology*, *111*(3-4), 151–157.

Allan, G. M., McNeilly, F., Kennedy, S., Daft, B., Clarke, E. G., Ellis, J. A., Haines, D. M., Meehan, B. M., & Adair, B. M. (1998). Isolation of porcine circovirus-like viruses from pigs with a wasting disease in the USA and Europe. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *10*(1), 3–10.

Alonso, C., Goede, D. P., Morrison, R. B., Davies, P. R., Rovira, A., Marthaler, D. G., & Torremorell, M. (2014). Evidence of infectivity of airborne porcine epidemic diarrhea virus and detection of airborne viral RNA at long distances from infected herds. *Veterinary Research*, *45*(1), 73.

Altarejos, J. Y., Goebel, N., Conkright, M. D., Inoue, H., Xie, J., Arias, C. M., Sawchenko, P. E., & Montminy, M. (2008). The Creb1 coactivator Crtc1 is required for energy balance and fertility. *Nature Medicine*, *14*(10), 1112–1117.

Alwine, J. C., Kemp, D. J., & Stark, G. R. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5350–5354.

Anderson, M. W., & Schrijver, I. (2010). Next generation DNA sequencing and the future of genomic medicine. *Genes*, *1*(1), 38–69.

Appay, V., & Rowland-Jones, S. L. (2001). RANTES: a versatile and controversial chemokine. *Trends in Immunology*, *22*(2), 83–87.

Arceo, M. E., Ernst, C. W., Lunney, J. K., Choi, I., Raney, N. E., Huang, T., Tuggle, C. K., Rowland, R. R. R., & Steibel, J. P. (2012). Characterizing differential individual response to porcine reproductive and respiratory syndrome virus infection through statistical and

functional analysis of gene expression. *Frontiers in Genetics*, *3*, 321.

Armitage, E. G., & Barbas, C. (2014). Metabolomics in cancer biomarker discovery: current trends and future perspectives. *Journal of Pharmaceutical and Biomedical Analysis*, *87*, 1–11.

Asai, T., Mori, M., Okada, M., Uruno, K., Yazawa, S., & Shibata, I. (1999). Elevated serum haptoglobin in pigs infected with porcine reproductive and respiratory syndrome virus. *Veterinary Immunology and Immunopathology*, *70*(1-2), 143–148.

Asai, T., Mori, M., Okada, M., Uruno, K., Yazawa, S., & Shibata, I. (1999). Elevated serum haptoglobin in pigs infected with porcine reproductive and respiratory syndrome virus. *Veterinary Immunology and Immunopathology*, *70*(1-2), 143–148.

Avuçlu, E., & Elen, A. (2020). Evaluation of train and test performance of machine learning algorithms and Parkinson diagnosis with statistical measurements. *Medical & Biological Engineering & Computing*, *58*(11), 2775–2788.

Bürger, W., Ewald, C., & Fennert, E.-M. (1998). Increase in C-reactive protein in the serum of piglets (pCRP) following ACTH or corticosteroid administration. *Journal of Veterinary Medicine. B, Infectious Diseases and Veterinary Public Health*, *45*(1-10), 1–6.

Bai, K., Jiang, L., Wang, T., & Wang, W. (2022). Treatment of immune dysfunction in intrauterine growth restriction piglets via supplementation with dimethylglycine sodium salt during the suckling period. *Animal Nutrition*, 11, 215-227.

Bai, M., Vozdek, R., Hnízda, A., Jiang, C., Wang, B., Kuchar, L., Li, T., Zhang, Y., Wood, C., Feng, L., Dang, Y., & Ma, D. K. (2018). Conserved roles of C. elegans and human MANFs in sulfatide binding and cytoprotection. *Nature Communications*, *9*(1), 897.

Bai, X., & Plastow, G. S. (2022). Breeding for disease resilience: opportunities to manage polymicrobial challenge and improve commercial performance in the pig industry. *CABI Agriculture and Bioscience*, *3*(1), 6.

Bai, X., Cheng, J., Fortin, F., Harding, J. C. S., Dyck, M. K., Dekkers, J. C. M., Field, C. J., Rogel-Gaillard, C., Blanc, F., & Plastow, G. S. (2022, December 31). Indicators of disease

resilience from complete blood count and *in vitro* immunoassays data from young-healthy pigs. *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*. World Congress on Genetics Applied to Livestock Production, Rotterdam, the Netherlands. https://doi.org/10.3920/978-90-8686-940-4_161

Bai, X., Putz, A. M., Wang, Z., Fortin, F., Harding, J. C. S., Dyck, M. K., Dekkers, J. C. M., Field, C. J., Plastow, G. S., & Canada, P. (2020). Exploring Phenotypes for Disease Resilience in Pigs Using Complete Blood Count Data From a Natural Disease Challenge Model. *Frontiers in Genetics*, *11*, 216.

Bai, X., Putz, A. M., Wang, Z., Fortin, F., Harding, J. C. S., Dyck, M. K., Dekkers, J. C. M., Field, C. J., Plastow, G. S., & Canada, P. (2020). Exploring Phenotypes for Disease Resilience in Pigs Using Complete Blood Count Data From a Natural Disease Challenge Model. *Frontiers in Genetics*, *11*, 216.

Bai, X., Yang, T., Putz, A. M., Wang, Z., Li, C., Fortin, F., Harding, J. C. S., Dyck, M. K., PigGen Canada, Dekkers, J. C. M., Field, C. J., & Plastow, G. S. (2021). Investigating the genetic architecture of disease resilience in pigs by genome-wide association studies of complete blood count traits collected from a natural disease challenge model. *BMC Genomics*, *22*(1), 535.

Baker, E. C., Cilkiz, K. Z., Riggs, P. K., Littlejohn, B. P., Long, C. R., Welsh, T. H., ... & Riley, D. G. (2020). Effect of prenatal transportation stress on DNA methylation in Brahman heifers. *Livestock Science*, 240, 104116.

Bao, W.-B., Ye, L., Zi, C., Su, X.-M., Pan, Z.-Y., Zhu, J., Zhu, G.-Q., Huang, X.-G., & Wu, S.-L. (2012). Study on the age-dependent tissue expression of FUT1 gene in porcine and its relationship to E. coli F18 receptor. *Gene*, *497*(2), 336–339.

Bathe, O. F., Shaykhutdinov, R., Kopciuk, K., Weljie, A. M., McKay, A., Sutherland, F. R., Dixon, E., Dunse, N., Sotiropoulos, D., & Vogel, H. J. (2011). Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored*

*by the American Society of Preventive Oncology*, *20*(1), 140–147.

Baumann, H., & Gauldie, J. (1990). Regulation of hepatic acute phase plasma protein genes by hepatocyte stimulating factors and other mediators of inflammation. *Molecular Biology & Medicine*, *7*(2), 147–159.

Baumann, H., & Gauldie, J. (1994). The acute phase response. *Immunology Today*, *15*(2), 74–80.

Becker-André, M., & Hahlbrock, K. (1989). Absolute mRNA quantification using the polymerase chain reaction (PCR). A novel approach by a PCR aided transcript titration assay (PATTY). *Nucleic Acids Research*, *17*(22), 9437–9446.

Benfield, D. A., Nelson, E., Collins, J. E., Harris, L., Goyal, S. M., Robison, D., Christianson, W. T., Morrison, R. B., Gorcyca, D., & Chladek, D. (1992). Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332). *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *4*(2), 127–133.

Bergland, F. D. M., Collins, J., & Frank, R. (2005). A Guide to Porcine Sample Submission and Diagnostic Tests.

Bermingham, K. M., Brennan, L., Segurado, R., Barron, R. E., Gibney, E. R., Ryan, M. F., Gibney, M. J., & O'Sullivan, A. M. (2021). Genetic and Environmental Contributions to Variation in the Stable Urinary NMR Metabolome over Time: A Classic Twin Study. *Journal of Proteome Research*, *20*(8), 3992–4000.

Berry, M. P. R., Graham, C. M., McNab, F. W., Xu, Z., Bloch, S. A. A., Oni, T., Wilkinson, K. A., Banchereau, R., Skinner, J., Wilkinson, R. J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J. J., Mejias, A., Ramilo, O., Kon, O. M., Pascual, V., Banchereau, J., … O'Garra, A. (2010). An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*, *466*(7309), 973–977.

Best, A., White, A., & Boots, M. (2008). Maintenance of host variation in tolerance to pathogens and parasites. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(52), 20786–20791.

Beyer, J., Fichtner, D., Schirrmeier, H., Polster, U., Weiland, E., & Wege, H. (2000). Porcine reproductive and respiratory syndrome virus (PRRSV): kinetics of infection in lymphatic organs and lung. *Journal of Veterinary Medicine. B, Infectious Diseases and Veterinary Public Health*, *47*(1), 9–25.

Bhatia, V., Schmied, J., Cheng, J., Bai, X., Mallard, B., Fortin, F., ... & Dekkers, J. C. M. (2022, December). Genetic relationships among immune response traits of young healthy pigs evaluated by immunoassays. *In Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges* (pp. 3282-3285). Wageningen Academic Publishers.

Bird, L. (2019). Getting enough energy for immunity [Review of *Getting enough energy for immunity*]. *Nature Reviews. Immunology*, *19*(5), 269.

Bishop, S. C., & Woolliams, J. A. (2014). Genomics and disease resistance studies in livestock. *Livestock Science*, *166*, 190–198.

Bisset, S. A., & Morris, C. A. (1996). Feasibility and implications of breeding sheep for resilience to nematode challenge. *International Journal for Parasitology*, *26*(8-9), 857–868.

Blanco, I., Canals, A., Evans, G., Mellencamp, M. A., Cia, C., Deeb, N., Wang, L., & Galina-Pantoja, L. (2008). Differences in susceptibility to Haemophilus parasuis infection in pigs. *Canadian Journal of Veterinary Research*, *72*(3), 228–235.

Bloom, C. I., Graham, C. M., Berry, M. P. R., Wilkinson, K. A., Oni, T., Rozakeas, F., Xu, Z., Rossello-Urgell, J., Chaussabel, D., Banchereau, J., Pascual, V., Lipman, M., Wilkinson, R. J., & O'Garra, A. (2012). Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PloS One*, *7*(10), e46191.

Boczkowska, M., Rebowski, G., & Dominguez, R. (2013). Glia maturation factor (GMF) interacts with Arp2/3 complex in a nucleotide state-dependent manner. *Journal of Biological Chemistry*, 288(36), 25683-25688.

Boddicker, N., Waide, E. H., Rowland, R. R. R., Lunney, J. K., Garrick, D. J., Reecy, J. M., &

Dekkers, J. C. M. (2012). Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. *Journal of Animal Science*, *90*(6), 1733–1746.

Boddicker, N., Waide, E. H., Rowland, R. R. R., Lunney, J. K., Garrick, D. J., Reecy, J. M., & Dekkers, J. C. M. (2012). Evidence for a major QTL associated with host response to porcine reproductive and respiratory syndrome virus challenge. *Journal of Animal Science*, *90*(6), 1733–1746.

Bossé, J. T., Li, Y., Sárközi, R., Gottschalk, M., Angen, Ø., Nedbalcova, K., Rycroft, A. N., Fodor, L., & Langford, P. R. (2017). A Unique Capsule Locus in the Newly Designated *Actinobacillus pleuropneumoniae* Serovar 16 and Development of a Diagnostic PCR Assay. *Journal of Clinical Microbiology*, *55*(3), 902–907.

Bourneuf, E., Hérault, F., Chicault, C., Carré, W., Assaf, S., Monnier, A., Mottier, S., Lagarrigue, S., Douaire, M., Mosser, J., & Diot, C. (2006). Microarray analysis of differential gene expression in the liver of lean and fat chickens. *Gene*, *372*, 162–170.

Brian, B. F., 4th, & Freedman, T. S. (2021). The Src-family Kinase Lyn in Immunoreceptor Signaling. *Endocrinology*, *162*(10). https://doi.org/10.1210/endocr/bqab152

Brisson, Y. (2014). É volution de l'industrie porcine canadienne / : Changing face of the Canadian hog industry.

Brown, L. R. (2012). *Full Planet, Empty Plates: The New Geopolitics of Food Scarcity*. W. W. Norton & Company.

Brown, W. E., Garcia, M., Mamedova, L. K., Christman, K. R., Zenobi, M. G., Staples, C. R., Leno, B. M., Overton, T. R., Whitlock, B. K., Daniel, J. A., & Bradford, B. J. (2021). Acute-phase protein α-1-acid glycoprotein is negatively associated with feed intake in postpartum dairy cows. *Journal of Dairy Science*, *104*(1), 806–817.

Bteich, M. (2019). An overview of albumin and alpha-1-acid glycoprotein main characteristics: highlighting the roles of amino acids in binding kinetics and molecular interactions. *Heliyon*, *5*(11), e02879.

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Current protocols in molecular biology*, 109(1), 21-29.

Buitenhuis, A. J., Sundekilde, U. K., Poulsen, N. A., Bertram, H. C., Larsen, L. B., & Sørensen, P. (2013). Estimation of genetic parameters and detection of quantitative trait loci for metabolites in Danish Holstein milk. *Journal of Dairy Science*, *96*(5), 3285–3295.

Busse, D. C., Habgood-Coote, D., Clare, S., Brandt, C., Bassano, I., Kaforou, M., Herberg, J., Levin, M., Eléouët, J.-F., Kellam, P., & Tregoning, J. S. (2020). Interferon-Induced Protein 44 and Interferon-Induced Protein 44-Like Restrict Replication of Respiratory Syncytial Virus. *Journal of Virology*, *94*(18). https://doi.org/10.1128/JVI.00297-20

Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D., & Craig, D. W. (2016). Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nature Reviews. Genetics*, *17*(5), 257–271.

Cai, S., Hu, B., Wang, X., Liu, T., Lin, Z., Tong, X., Xu, R., Chen, M., Duo, T., Zhu, Q., Liang, Z., Li, E., Chen, Y., Li, J., Liu, X., & Mo, D. (2023). Integrative single-cell RNA-seq and ATAC-seq analysis of myogenic differentiation in pig. *BMC Biology*, *21*(1), 19.

Calvert, J. G., Slade, D. E., Shields, S. L., Jolie, R., Mannan, R. M., Ankenbauer, R. G., & Welch, S.-K. W. (2007). CD163 expression confers susceptibility to porcine reproductive and respiratory syndrome viruses. *Journal of Virology*, *81*(14), 7371–7379.

Campbell, F. M., Waterston, M., Andresen, L. O., Sorensen, N. S., Heegaard, P. M. H., & Eckersall, P. D. (2005). The negative acute phase response of serum transthyretin following Streptococcus suis infection in the pig. *Veterinary Research*, *36*(4), 657–664.

Cantwell, C. A., Sterneck, E., & Johnson, P. F. (1998). Interleukin-6-Specific Activation of the C/EBPδ Gene in Hepatocytes Is Mediated by Stat3 and Sp1. *Molecular and Cellular Biology*, *18*(4), 2108–2117.

Cao, Y., Zhang, J., Yang, W., Xia, C., Zhang, H.-Y., Wang, Y.-H., & Xu, C. (2017). Predictive Value of Plasma Parameters in the Risk of Postpartum Ketosis in Dairy Cows. *Journal of*

*Veterinary Research*, *61*(1), 91–95.

Cardoso, F. S., Ricardo, L. B., Oliveira, A. M., Horta, D. V., Papoila, A. L., Deus, J. R., & Canena, J. (2015). C-Reactive Protein at 24 Hours after Hospital Admission may have Relevant Prognostic Accuracy in Acute Pancreatitis: A Retrospective Cohort Study. *GE Portuguese Journal of Gastroenterology*, *22*(5), 198–203.

Carpintero, R., Piñeiro, M., Andrés, M., Iturralde, M., Alava, M. A., Heegaard, P. M. H., Jobert, J. L., Madec, F., & Lampreave, F. (2005). The concentration of apolipoprotein A-I decreases during experimentally induced acute-phase processes in pigs. *Infection and Immunity*, *73*(5), 3184–3187.

Caruso, J. P., & Ross, R. F. (1990). Effects of Mycoplasma hyopneumoniae and Actinobacillus (Haemophilus) pleuropneumoniae infections on alveolar macrophage functions in swine. *American Journal of Veterinary Research*, *51*(2), 227–231.

Carvajal, A., Lanza, I., Diego, R., Rubio, P., & Cármenes, P. (1995). Evaluation of a blocking ELISA using monoclonal antibodies for the detection of porcine epidemic diarrhea virus and its antibodies. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *7*(1), 60–64.

Ceron, J. J., Eckersall, P. D., & Martýnez-Subiela, S. (2005). Acute phase proteins in dogs and cats: current knowledge and future perspectives. *Veterinary Clinical Pathology / American Society for Veterinary Clinical Pathology*, *34*(2), 85–99.

Chapinal, N., Leblanc, S. J., Carson, M. E., Leslie, K. E., Godden, S., Capel, M., Santos, J. E. P., Overton, M. W., & Duffield, T. F. (2012). Herd-level association of serum metabolites in the transition period with disease, milk production, and early lactation reproductive performance. *Journal of Dairy Science*, *95*(10), 5676–5682.

Chaplin, D. D. (2010). Overview of the immune response. *The Journal of Allergy and Clinical Immunology*, *125*(2 Suppl 2), S3–S23.

Chase-Topping, M., Plastow, G., Dekkers, J., Li, Y., Fang, Y., Gerdts, V., ... & Doeschl-Wilson, A. (2023). The WUR0000125 PRRS resilience SNP had no apparent effect on pigs'

infectivity and susceptibility in a novel transmission trial. *Genetics Selection Evolution*, 55(1), 51.

Chaudhari, M. (2016). Significant and Roles of Microarray Technology in Genetics. *AGROBIOS NEWSLETTER*, *XV*(07), 83–84.

Chefetz, I., Amitai, D. B., Browning, S., Skorecki, K., Adir, N., Thomas, M. G., Kogleck, L., Topaz, O., Indelman, M., Uitto, J., Richard, G., Bradman, N., & Sprecher, E. (2008). Normophosphatemic Familial Tumoral Calcinosis Is Caused by Deleterious Mutations in SAMD9, Encoding a TNF-α Responsive Protein. *The Journal of Investigative Dermatology*, *128*(6), 1423–1429.

Chen, H., Sun, H., You, F., Sun, W., Zhou, X., Chen, L., Yang, J., Wang, Y., Tang, H., Guan, Y., Xia, W., Gu, J., Ishikawa, H., Gutman, D., Barber, G., Qin, Z., & Jiang, Z. (2011). Activation of STAT6 by STING is critical for antiviral innate immunity. *Cell*, *147*(2), 436–446.

Chen, H.-H., Lin, J.-H., Fung, H.-P., Ho, L.-L., Yang, P.-C., Lee, W.-C., Lee, Y.-P., & Chu, R.-M. (2003). Serum acute phase proteins and swine health status. *Canadian Journal of Veterinary Research*, *67*(4), 283–290.

Chen, J., Zhang, X., Cao, R., Lu, X., Zhao, S., Fekete, A., Huang, Q., Schmitt-Kopplin, P., Wang, Y., Xu, Z., Wan, X., Wu, X., Zhao, N., Xu, C., & Xu, G. (2011). Serum 27-nor-5β-cholestane-3,7,12,24,25 pentol glucuronide discovered by metabolomics as potential diagnostic biomarker for epithelium ovarian cancer. *Journal of Proteome Research*, *10*(5), 2625–2632.

Chen, Y., Lonergan, S., Lim, K.-S., Cheng, J., Putz, A. M., Dyck, M. K., Canada, P., Fortin, F., Harding, J. C. S., Plastow, G. S., & Dekkers, J. C. M. (2023). Plasma protein levels of young healthy pigs as indicators of disease resilience. *Journal of Animal Science*, *101*. https://doi.org/10.1093/jas/skad014

Chen, Y., Tibbs-Cortes, L. E., Ashley, C., Putz, A. M., Lim, K.-S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., Harding, J. C. S., & PigGen Canada. (2020). The genetic

basis of natural antibody titers of young healthy pigs and relationships with disease resilience. *BMC Genomics*, *21*(1), 648.

Chen, Y., Tibbs-Cortes, L. E., Ashley, C., Putz, A. M., Lim, K.-S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., Harding, J. C. S., & PigGen Canada. (2020). The genetic basis of natural antibody titers of young healthy pigs and relationships with disease resilience. *BMC Genomics*, *21*(1), 648.

Cheng, J., Fernando, R., Cheng, H., Kachman, S. D., Lim, K., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Canada, P., & Dekkers, J. C. M. (2022). Genome-wide association study of disease resilience traits from a natural polymicrobial disease challenge model in pigs identifies the importance of the major histocompatibility complex region. *G3* , *12*(3). https://doi.org/10.1093/g3journal/jkab441

Cheng, J., Lim, K., Putz, A. M., Wolc, A., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., PigGen Canada, & Dekkers, J. C. M. (2022). Genetic analysis of disease resilience of wean-to-finish pigs under a natural disease challenge model using reaction norms. *Genetics, Selection, Evolution*, *54*(1), 11.

Cheng, J., Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Canada, P., & Dekkers, J. C. M. (2020). Genetic analysis of disease resilience in wean-to-finish pigs from a natural disease challenge model. *Journal of Animal Science*, *98*(8).

Cheng, J., Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Canada, P., & Dekkers, J. C. M. (2020). Genetic analysis of disease resilience in wean-to-finish pigs from a natural disease challenge model. *Journal of Animal Science*, *98*(8). https://doi.org/10.1093/jas/skaa244

Cheng, J., Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Canada, P. G., & Dekkers, J. C. M. (2021). Genetic parameters of drinking and feeding traits of wean-to-finish pigs under a polymicrobial natural disease challenge. *Journal of Animal Science and Biotechnology*, *12*(1), 105.

Cheresh, D. A., Haynes, D. H., & Distasio, J. A. (1984). Interaction of an acute phase reactant,

alpha 1-acid glycoprotein (orosomucoid), with the lymphoid cell surface: a model for non-specific immune suppression. *Immunology*, *51*(3), 541–548.

Chin, S. S., Guillen, E., Chorro, L., Achar, S., Ng, K., Oberle, S., Alfei, F., Zehn, D., Altan-Bonnet, G., Delahaye, F., & Lauvau, G. (2022). T cell receptor and IL-2 signaling strength control memory CD8+ T cell functional fitness via chromatin remodeling. *Nature Communications*, *13*(1), 2240.

Chiou, M. T., Jeng, C. R., Chueh, L. L., Cheng, C. H., & Pang, V. F. (2000). Effects of porcine reproductive and respiratory syndrome virus (isolate tw91) on porcine alveolar macrophages in vitro. *Veterinary Microbiology*, *71*(1-2), 9–25.

Chiu, K. M., Mortensen, R. F., Osmand, A. P., & Gewurz, H. (1977). Interactions of alpha1-acid glycoprotein with the immune system. I. Purification and effects upon lymphocyte responsiveness. *Immunology*, *32*(6), 997–1005.

Choi, I., Bao, H., Kommadath, A., Hosseini, A., Sun, X., Meng, Y., Stothard, P., Plastow, G. S., Tuggle, C. K., Reecy, J. M., Fritz-Waters, E., Abrams, S. M., Lunney, J. K., & Guan, L. L. (2014). Increasing gene discovery and coverage using RNA-seq of globin RNA reduced porcine blood samples. *BMC Genomics*, *15*(1), 954.

Chong, J., Wishart, D. S., & Xia, J. (2019). Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]*, *68*(1), e86.

Clapperton, M., Bishop, S. C., Cameron, N. D., & Glass, E. J. (2005). Associations of acute phase protein levels with growth performance and with selection for growth performance in Large White pigs. *Animal Science* , *81*(2), 213–220.

Clapperton, M., Bishop, S. C., Cameron, N. D., & Glass, E. J. (2005). Associations of acute phase protein levels with growth performance and with selection for growth performance in Large White pigs. *Animal Science* , *81*(2), 213–220.

Clapperton, M., Bishop, S. C., Piñeiro, M., Campbell, F. M., & Glass, E. J. (2007). The association between plasma levels of acute phase proteins, haptoglobin, alpha-1 acid

glycoprotein (AGP), Pig-MAP, transthyretin and serum amyloid A (SAA) in Large White and Meishan pigs. *Veterinary Immunology and Immunopathology*, *119*(3), 303–309.

Conti, P., Caraffa, A., Gallenga, C. E., Ross, R., Kritas, S. K., Frydas, I., Younes, A., & Ronconi, G. (2020). Coronavirus-19 (SARS-CoV-2) induces acute severe lung inflammation via IL-1 causing cytokine storm in COVID-19: a promising inhibitory strategy. *Journal of Biological Regulators and Homeostatic Agents*, *34*(6), 1971–1975.

Corbett, R. J., Luttman, A. M., Wurtz, K. E., Siegford, J. M., Raney, N. E., Ford, L. M., & Ernst, C. W. (2021). Weaning induces stress-dependent DNA methylation and transcriptional changes in piglet PBMCs. *Frontiers in Genetics*, 12, 633564.

Cravatt, B. F., Prospero-Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Boger, D. L., & Lerner, R. A. (1995). Chemical characterization of a family of brain lipids that induce sleep. *Science*, *268*(5216), 1506–1509.

Cray, C., Zaias, J., & Altman, N. H. (2009). Acute phase response in animals: a review. *Comparative Medicine*, *59*(6), 517–526.

Cray, C., Zaias, J., & Altman, N. H. (2009). Acute phase response in animals: a review. *Comparative Medicine*, *59*(6), 517–526.

Cruijsen, A. L. M. (1995). *Actinobacillus pleuropneumoniae infections in swine: pathogenetic and epidemiological studies*. Universiteit Utrecht, Faculteit Diergeneeskunde.

Cui, X., Qian, P., Rao, T., Wei, Y., Zhao, F., Zhang, H., Chen, H., & Li, X. (2019). Cellular Interleukin Enhancer-Binding Factor 2, ILF2, Inhibits Japanese Encephalitis Virus Replication In Vitro. *Viruses*, *11*(6). https://doi.org/10.3390/v11060559

Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., Bolduc, V., Waddell, L. B., Sandaradura, S. A., O'Grady, G. L., Estrella, E., Reddy, H. M., Zhao, F., Weisburd, B., Karczewski, K. J., O'Donnell-Luria, A. H., Birnbaum, D., Sarkozy, A., Hu, Y., … MacArthur, D. G. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Science Translational Medicine*, 9(386).

Das, P. B., Dinh, P. X., Ansari, I. H., de Lima, M., Osorio, F. A., & Pattnaik, A. K. (2010). The

minor envelope glycoproteins GP2a and GP4 of porcine reproductive and respiratory syndrome virus interact with the receptor CD163. *Journal of Virology*, *84*(4), 1731–1740.

Davies, P. R. (2012). One world, one health: the threat of emerging swine diseases. A North American perspective. *Transboundary and Emerging Diseases*, *59 Suppl 1*, 18–26.

Dawson, H. D., Loveland, J. E., Pascal, G., Gilbert, J. G. R., Uenishi, H., Mann, K. M., Sang, Y., Zhang, J., Carvalho-Silva, D., Hunt, T., Hardy, M., Hu, Z., Zhao, S.-H., Anselmo, A., Shinkai, H., Chen, C., Badaoui, B., Berman, D., Amid, C., … Tuggle, C. K. (2013). Structural and functional annotation of the porcine immunome. *BMC Genomics*, *14*, 332.

de Lange, C. F. M., Marty, B. J., Birkett, S., Morel, P., & Szkotnicki, B. (2001). Application of pig growth models in commercial pork production. *Canadian Journal of Animal Science*, *81*(1), 1–8.

Dekkers, J. C. M., Putz, A. M., Harding, J. C. S., Lim, K.-S., Dyck, M. K., Fortin, F., Chen, Y., Tibbs, L. E., Plastow, G. S., & Ashley, C. (2019). The Genetic Basis of Natural Antibody Titers and Relationships with Disease Resilience in Pigs. *Iowa State University Animal Industry Report,* 16(1).

Delgoffe, G. M., & Powell, J. D. (2015). Feeding an army: The metabolism of T cells in activation, anergy, and exhaustion. *Molecular Immunology*, *68*(2 Pt C), 492–496.

Dent, A. L., Doherty, T. M., Paul, W. E., Sher, A., & Staudt, L. M. (1999). BCL-6-deficient mice reveal an IL-4-independent, STAT6-dependent pathway that controls susceptibility to infection by Leishmania major. *Journal of Immunology* , *163*(4), 2098–2103.

Dervishi, E., Bai, X., Cheng, J., Fortin, F., Dyck, M. K., Harding, J. C. S., Seddon, Y. M., Dekkers, J. C. M., Canada, P., & Plastow, G. (2023). Exploration of plasma metabolite levels in healthy nursery pigs in response to environmental enrichment and disease resilience. *Journal of Animal Science*, *101*. https://doi.org/10.1093/jas/skad033

Dervishi, E., Bai, X., Cheng, J., Fortin, F., Dyck, M. K., Harding, J. C. S., Seddon, Y. M., Dekkers, J. C. M., Canada, P., & Plastow, G. (2023). Exploration of plasma metabolite levels in healthy nursery pigs in response to environmental enrichment and disease

resilience. *Journal of Animal Science*, *101*. https://doi.org/10.1093/jas/skad033

Dervishi, E., Bai, X., Dyck, M. K., Harding, J. C. S., Fortin, F., Dekkers, J. C. M., & Plastow, G. (2023). GWAS and genetic and phenotypic correlations of plasma metabolites with complete blood count traits in healthy young pigs reveal implications for pig immune response. *Frontiers in Molecular Biosciences*, *10*, 1140375.

Dervishi, E., Yang, T., Dyck, M. K., Harding, J. C. S., Fortin, F., PigGen Canada, Cheng, J., Dekkers, J. C. M., & Plastow, G. (2021). Heritability and genetic correlations of plasma metabolites of pigs with production, resilience and carcass traits under natural polymicrobial disease challenge. *Scientific Reports*, *11*(1), 20628.

Dervishi, E., Yang, T., Dyck, M. K., Harding, J. C. S., Fortin, F., PigGen Canada, Cheng, J., Dekkers, J. C. M., & Plastow, G. (2021). Heritability and genetic correlations of plasma metabolites of pigs with production, resilience and carcass traits under natural polymicrobial disease challenge. *Scientific Reports*, *11*(1), 20628.

Desrosiers, R. (2001). A review of some aspects of the epidemiology, diagnosis, and control of Mycoplasma hyopneumoniae infections. *Journal of Swine Health and Production*, *9*(5), 233–237.

Dinarello, C. A. (2018). Overview of the IL-1 family in innate inflammation and acquired immunity. *Immunological Reviews*, *281*(1), 8–27.

do Nascimento, N. C., Guimaraes, A. M. S., Dos Santos, A. P., Chu, Y., Marques, L. M., & Messick, J. B. (2018). RNA-Seq based transcriptome of whole blood from immunocompetent pigs (Sus scrofa) experimentally infected with Mycoplasma suis strain Illinois. *Veterinary Research*, *49*(1), 49.

Doeschl-Wilson, A. B., Bishop, S. C., Kyriazakis, I., & Villanueva, B. (2012). Novel methods for quantifying individual host response to infectious pathogens for genetic analyses. *Frontiers in Genetics*, *3*, 266.

Dollt, C., Michel, J., Kloss, L., Melchers, S., Schledzewski, K., Becker, K., Sauer, A., Krewer, A., Koll, F., & Schmieder, A. (2018). The novel immunoglobulin super family receptor

SLAMF9 identified in TAM of murine and human melanoma influences pro-inflammatory cytokine secretion and migration. *Cell Death & Disease*, *9*(10), 939.

Dong, F., Zhang, L., Hao, F., Tang, H., & Wang, Y. (2013). Systemic responses of mice to dextran sulfate sodium-induced acute ulcerative colitis using 1H NMR spectroscopy. *Journal of Proteome Research*, *12*(6), 2958–2966.

Donovan, T. (2005). The role of influenza on growing pig performance.

Dragoi, A.-M., & Agaisse, H. (2015). The class II phosphatidylinositol 3-phosphate kinase PIK3C2A promotes Shigella flexneri dissemination through formation of vacuole-like protrusions. *Infection and Immunity*, *83*(4), 1695–1704.

Duke-Cohan, J. S., Tang, W., & Schlossman, S. F. (2000). Attractin: a cub-family protease involved in T cell-monocyte/macrophage interactions. *Advances in Experimental Medicine and Biology*, *477*, 173–185.

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews*, *40*(1), 387–426.

Duthie, S., Eckersall, P. D., Addie, D. D., Lawrence, C. E., & Jarrett, O. (1997). Value of alpha 1-acid glycoprotein in the diagnosis of feline infectious peritonitis. *The Veterinary Record*, *141*(12), 299–303.

Eckersall, P. D. (n.d.). Meat inspection. The potential of acute phase protein assay. *Meat Focus International*.

Eckersall, P. D., & Bell, R. (2010). Acute phase proteins: Biomarkers of infection and inflammation in veterinary medicine. *Veterinary Journal* , *185*(1), 23–27.

Eckersall, P. D., Saini, P. K., & McComb, C. (1996). The acute phase response of acid soluble glycoprotein, α1-acid glycoprotein, ceruloplasmin, haptoglobin and C-reactive protein, in the pig. *Veterinary Immunology and Immunopathology*, *51*(3), 377–385.

Eckersall, P. D., Saini, P. K., & McComb, C. (1996). The acute phase response of acid soluble glycoprotein, α1-acid glycoprotein, ceruloplasmin, haptoglobin and C-reactive protein, in the

pig. *Veterinary Immunology and Immunopathology*, *51*(3), 377–385.

Elemam, N. M., Talaat, I. M., & Maghazachi, A. A. (2022). CXCL10 Chemokine: A Critical Player in RNA and DNA Viral Infections. *Viruses*, *14*(11). https://doi.org/10.3390/v14112445

Elmore, S. (2007). Apoptosis: A Review of Programmed Cell Death. *Toxicologic Pathology*, *35*(4), 495–516.

Engering, A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen–host–environment interplay and disease emergence. *Emerging Microbes & Infections*, *2*(1), 1–7.

Fablet, C., Rose, N., Grasland, B., Robert, N., Lewandowski, E., & Gosselin, M. (2018). Factors associated with the growing-finishing performances of swine herds: an exploratory study on serological and herd level indicators. *Porcine Health Management*, *4*, 6.

Fachinger, V., Bischoff, R., Jedidia, S. B., Saalmüller, A., & Elbers, K. (2008). The effect of vaccination against porcine circovirus type 2 in pigs suffering from porcine respiratory disease complex. *Vaccine*, *26*(11), 1488–1499.

Fajgenbaum, D. C., & June, C. H. (2020). Cytokine storm. New England Journal of Medicine, 383(23), 2255-2273.

Fan, T. W. M., Bandura, L. L., Higashi, R. M., & Lane, A. N. (2006). Metabolomics-edited transcriptomics analysis of Se anticancer action in human lung cancer cells. *Metabolomics: Official Journal of the Metabolomic Society*, *1*(4), 325–339.

Fan, T. W. M., Lane, A. N., Higashi, R. M., Farag, M. A., Gao, H., Bousamra, M., & Miller, D. M. (2009). Altered regulation of metabolic pathways in human lung cancer discerned by 13C stable isotope-resolved metabolomics (SIRM). *Molecular Cancer*, *8*(1), 1–19.

Fano, E., Pijoan, C., Dee, S., & Deen, J. (2007). Effect of Mycoplasma hyopneumoniae colonization at weaning on disease severity in growing pigs. *Canadian Journal of Veterinary Research*, *71*(3), 195–200.

Fenaux, M., Halbur, P. G., Haqshenas, G., Royer, R., Thomas, P., Nawagitgul, P., Gill, M., Toth, T. E., & Meng, X. J. (2002). Cloned genomic DNA of type 2 porcine circovirus is infectious

when injected directly into the liver and lymph nodes of pigs: characterization of clinical disease, virus distribution, and pathologic lesions. *Journal of Virology*, *76*(2), 541–551.

Fernández-Sáiz, V., Targosz, B.-S., Lemeer, S., Eichner, R., Langer, C., Bullinger, L., Reiter, C., Slotta-Huspenina, J., Schroeder, S., Knorn, A.-M., Kurutz, J., Peschel, C., Pagano, M., Kuster, B., & Bassermann, F. (2013). SCFFbxo9 and CK2 direct the cellular response to growth factor withdrawal via Tel2/Tti1 degradation and promote survival in multiple myeloma. *Nature Cell Biology*, *15*(1), 72–81.

Fiehn, O. (2002). Metabolomics–the link between genotypes and phenotypes. *Plant Molecular Biology*, *48*(1/2), 155–171.

Fishbourne, E., Hutet, E., Abrams, C., Cariolet, R., Le Potier, M.-F., Takamatsu, H.-H., & Dixon, L. K. (2013). Increase in chemokines CXCL10 and CCL2 in blood from pigs infected with high compared to low virulence African swine fever virus isolates. *Veterinary Research*, *44*(1), 87.

Flori, L., Gao, Y., Laloë, D., Lemonnier, G., Leplat, J.-J., Teillaud, A., Cossalter, A.-M., Laffitte, J., Pinton, P., de Vaureix, C., Bouffaud, M., Mercat, M.-J., Lefèvre, F., Oswald, I. P., Bidanel, J.-P., & Rogel-Gaillard, C. (2011). Immunity traits in pigs: substantial genetic variation and limited covariation. *PloS One*, *6*(7), e22717.

Fontanesi, L. (2016). Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. *Animal Frontiers*, *6*(1), 73–79.

Fornoni, J., Núñez-Farfán, J., Valverde, P. L., & Rausher, M. D. (2004). Evolution of mixed strategies of plant defense allocation against natural enemies. *Evolution; International Journal of Organic Evolution*, *58*(8), 1685–1695.

Friggens, N. C., Blanc, F., Berry, D. P., & Puillet, L. (2017). Review: Deciphering animal robustness. A synthesis to facilitate its use in livestock breeding and management. *Animal: An International Journal of Animal Bioscience*, *11*(12), 2237–2251.

Fung, T. S., & Liu, D. X. (2014). Coronavirus infection, ER stress, apoptosis and innate immunity. *Frontiers in Microbiology*, *5*, 296.

Ganeshan, K., Nikkanen, J., Man, K., Leong, Y. A., Sogawa, Y., Maschek, J. A., ... & Chawla, A. (2019). Energetic trade-offs and hypometabolic states promote disease tolerance. *Cell*, 177(2), 399-413.

Gates, S. C., & Sweeley, C. C. (1978). Quantitative metabolic profiling based on gas chromatography. *Clinical Chemistry*, *24*(10), 1663–1673.

German, J. B., Hammock, B. D., & Watkins, S. M. (2005). Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics: Official Journal of the Metabolomic Society*, *1*(1), 3–9.

Gerner, W., Käser, T., & Saalmüller, A. (2009). Porcine T lymphocytes and NK cells--an update. *Developmental and Comparative Immunology*, *33*(3), 310–320.

Ghani, M. J. (2022). SGK1, autophagy and cancer: an overview. *Molecular Biology Reports*, *49*(1), 675–685.

Ghosh, M., Sodhi, S. S., Song, K.-D., Kim, J. H., Mongre, R. K., Sharma, N., Singh, N. K., Kim, S. W., Lee, H. K., & Jeong, D. K. (2015). Evaluation of body growth and immunity-related differentially expressed genes through deep RNA sequencing in the piglets of Jeju native pig and Berkshire. *Animal Genetics*, *46*(3), 255–264.

Glass, E. J. (2012). The molecular pathways underlying host resistance and tolerance to pathogens. *Frontiers in Genetics*, *3*, 263.

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PloS One*, *12*(5), e0177675.

Goldberg, T. L., Weigel, R. M., Hahn, E. C., & Scherba, G. (2000). Associations between genetics, farm characteristics and clinical disease in field outbreaks of porcine reproductive and respiratory syndrome virus. *Preventive Veterinary Medicine*, *43*(4), 293–302.

Gondek, M., Knysz, P., Pomorska-Mól, M., Ziomek, M., & Bień-Kalinowska, J. (2020). Acute phase protein pattern and antibody response in pigs experimentally infected with a moderate dose of Trichinella spiralis, T. britovi, and T. pseudospiralis. *Veterinary Parasitology*, *288*,

109277.

González-Ramón, N., Hoebe, K., Alava, M. A., Van Leengoed, L., Piñeiro, M., Carmona, S., Iturralde, M., Lampreave, F., & Piñeiro, A. (2000). Pig MAP/ITIH4 and haptoglobin are interleukin-6-dependent acute-phase plasma proteins in porcine primary cultured hepatocytes. *European Journal of Biochemistry / FEBS*, *267*(6), 1878–1885.

Goodwin, R., Haque, S., Neto, F., & Myers, L. B. (2009). Initial psychological responses to Influenza A, H1N1 ("Swine flu"). *BMC Infectious Diseases*, *9*(1), 166.

Graber, C. D., Goust, J. M., Glassman, A. D., Kendall, R., & Loadholt, C. B. (1981). Immunomodulating properties of dimethylglycine in humans. *Journal of Infectious Diseases*, 143(1), 101-105.

Grau-Roma, L., Fraile, L., & Segalés, J. (2011). Recent advances in the epidemiology, diagnosis and control of diseases caused by porcine circovirus type 2. *Veterinary Journal* , *187*(1), 23–32.

Gray, G. D. (1997). The use of genetically resistant sheep to control nematode parasitism. *Veterinary Parasitology*, *72*(3-4), 345–357; discussion 357–366.

Gregg, B. A., Parker, P. A., Waller, K. M., Schneider, L. G., Garcia, M., Bradford, B., Daniel, J. A., & Whitlock, B. K. (2019). Effects of central and peripheral administration of an acute-phase protein, α-1-acid-glycoprotein, on feed intake and rectal temperature in sheep. *Journal of Animal Science*, *97*(12), 4783–4791.

Groenen, M. A. M., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J., Li, S., Larkin, D. M., Kim, H., Frantz, L. A. F., Caccamo, M., Ahn, H., Aken, B. L., Anselmo, A., Anthon, C., … Schook, L. B. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, *491*(7424), 393–398.

Groom, J. R., & Luster, A. D. (2011). CXCR3 in T cell function. *Experimental Cell Research*, *317*(5), 620–631.

Guak, H., & Krawczyk, C. M. (2020). Implications of cellular metabolism for immune

cell migration. *Immunology*, 161(3), 200-208.

Guijas, C., Montenegro-Burke, J. R., Warth, B., Spilker, M. E., & Siuzdak, G. (2018). Metabolomics activity screening for identifying metabolites that modulate phenotype. *Nature Biotechnology*, *36*(4), 316–320.

Guijas, C., Montenegro-Burke, J. R., Warth, B., Spilker, M. E., & Siuzdak, G. (2018). Metabolomics activity screening for identifying metabolites that modulate phenotype. *Nature Biotechnology*, *36*(4), 316–320.

Gultiken, N., Serhat, A., S., Gul, F., Yarim, Elvan, A., & Findik, M. (2012). Evaluation of plasma and milk haptoglobin concentrations in the diagnosis and treatment follow-up of subclinical mastitis in dairy cows. *Acta Veterinaria*, *62*(2-3), 271–279.

Guy, S. Z. Y., Li, L., Thomson, P. C., & Hermesch, S. (2019). Quantifying the health challenges in an Australian piggery using medication records for the definition of disease resilience1. *Journal of Animal Science*, *97*(3), 1076–1089.

Guy, S. Z. Y., Thomson, P. C., & Hermesch, S. (2012). Selection of pigs for improved coping with health and environmental challenges: breeding for resistance or tolerance? *Frontiers in Genetics*, *3*, 281.

Halbur, P. G., Paul, P. S., Meng, X.-J., Lum, M. A., Andrews, J. J., & Rathje, J. A. (1996). Comparative Pathogenicity of Nine US Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) Isolates in a Five-Week-Old Cesarean-Derived, Colostrum-Deprived Pig Model. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *8*(1), 11–20.

Hammond-Kosack, K. E., & Jones, J. D. G. (1997). PLANT DISEASE RESISTANCE GENES. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*, 575–607.

Han, L., Jara, C. P., Wang, O., Shi, Y., Wu, X., Thibivilliers, S., Wóycicki, R. K., Carlson, M. A., Velander, W. H., Araújo, E. P., Libault, M., Zhang, C., & Lei, Y. (2022). Isolating and cryopreserving pig skin cells for single-cell RNA sequencing study. *PloS One*, *17*(2), e0263869.

Hao, J., Hu, Y., Li, Y., Zhou, Q., & Lv, X. (2017). Involvement of JNK signaling in IL4-induced M2 macrophage polarization. *Experimental Cell Research*, 357(2), 155-162.

Harding, J. C., Baarsch, M. J., & Murtaugh, M. P. (1997). Association of tumour necrosis factor and acute phase reactant changes with post arrival disease in swine. *Zentralblatt Fur Veterinarmedizin. Reihe B. Journal of Veterinary Medicine. Series B*, *44*(7), 405–413.

Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., & O'Donovan, C. (2020). MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic Acids Research*, *48*(D1), D440–D444.

Heegaard, P. M. H., Klausen, J., Nielsen, J. P., González-Ramón, N., Piñeiro, M., Lampreave, F., & Alava, M. A. (1998). The porcine acute phase response to infection with *Actinobacillus pleuropneumoniae*. Haptoglobin, C-reactive protein, major acute phase protein and serum amyloid A protein are sensitive indicators of infection. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, *119*(2), 365–373.

Heegaard, P. M. H., Klausen, J., Nielsen, J. P., González-Ramón, N., Piñeiro, M., Lampreave, F., & Alava, M. A. (1998). The porcine acute phase response to infection with Actinobacillus pleuropneumoniae. Haptoglobin, C-reactive protein, major acute phase protein and serum amyloid A protein are sensitive indicators of infection. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, *119*(2), 365–373.

Heegaard, P. M. H., Miller, I., Sorensen, N. S., Soerensen, K. E., & Skovgaard, K. (2013). Pig α1-acid glycoprotein: characterization and first description in any species as a negative acute phase protein. *PloS One*, *8*(7), e68110.

Heinrich, P. C., Castell, J. V., & Andus, T. (1990). Interleukin-6 and the acute phase response. *Biochemical Journal*, *265*(3), 621–636.

Henryon, M., Berg, P., Jensen, J., & Andersen, S. (2001). Genetic variation for resistance to clinical and subclinical diseases exists in growing pigs. *Animal Science* , *73*(3), 375–387.

Hermesch, S., Li, L., Doeschl-Wilson, A. B., & Gilbert, H. (2015). Selection for productivity and robustness traits in pigs. *Animal Production Science*, *55*(12), 1437–1447.

Herrera-Uribe, J., Wiarda, J. E., Sivasankaran, S. K., Daharsh, L., Liu, H., Byrne, K. A., Smith, T. P. L., Lunney, J. K., Loving, C. L., & Tuggle, C. K. (2021). Reference Transcriptomes of Porcine Peripheral Immune Cells Created Through Bulk and Single-Cell RNA Sequencing. *Frontiers in Genetics*, *12*, 689406.

Hirai, M. Y., Yano, M., Goodenowe, D. B., Kanaya, S., Kimura, T., Awazuhara, M., Arita, M., Fujiwara, T., & Saito, K. (2004). Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(27), 10205–10210.

Hiss, S., & Sauerwein, H. (2003). Influence of dietary ss-glucan on growth performance, lymphocyte proliferation, specific immune response and haptoglobin plasma concentrations in pigs. *Journal of Animal Physiology and Animal Nutrition*, *87*(1-2), 2–11.

Hochepied, T., Berger, F. G., Baumann, H., & Libert, C. (2003). α1-Acid glycoprotein: an acute phase protein with inflammatory and immunomodulating properties. *Cytokine & Growth Factor Reviews*, *14*(1), 25–34.

Holko, I., Urbanova, J., THolkova, & Kmet, V. (2004). Diagnostics of main bacterial agents of porcine respiratory diseases complex (PRDC) using PCR detection of Mycoplasma hyopneumoniae. *Veterinární Medicína; Prague*, *49*(2), 35–41.

Horning, E. C., Brooks, C. J., & Vanden Heuvel, W. J. (1968). Gas phase analytical methods for the study of steroids. *Advances in Lipid Research*, *6*, 273–392.

Hoult, D. I., Busby, S. J., Gadian, D. G., Radda, G. K., Richards, R. E., & Seeley, P. J. (1974). Observation of tissue metabolites using 31P nuclear magnetic resonance. *Nature*, *252*(5481), 285–287.

Hu, L., Shi, Q., Shi, M., Liu, R., & Wang, C. (2017). Diagnostic Value of PCT and CRP for Detecting Serious Bacterial Infections in Patients With Fever of Unknown Origin: A Systematic Review and Meta-analysis. *Applied Immunohistochemistry & Molecular Morphology: AIMM / Official Publication of the Society for Applied Immunohistochemistry*,

*25*(8), e61–e69.

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*(11), 801–811.

Huang, J. X., Azad, M. A. K., Yuriev, E., Baker, M. A., Nation, R. L., Li, J., Cooper, M. A., & Velkov, T. (2012). Molecular Characterization of Lipopolysaccharide Binding to Human α-1-Acid Glycoprotein. *Journal of Lipids*, *2012*, 475153.

Huang, L., Liu, H., Zhang, K., Meng, Q., Hu, L., Zhang, Y., Xiang, Z., Li, J., Yang, Y., Chen, Y., Cui, S., Tang, H., Pei, H., Bu, Z., & Weng, C. (2020). Ubiquitin-Conjugating Enzyme 2S Enhances Viral Replication by Inhibiting Type I IFN Production through Recruiting USP15 to Deubiquitinate TBK1. *Cell Reports*, *32*(7), 108044.

Huang, T.-H., Uthe, J. J., Bearson, S. M. D., Demirkale, C. Y., Nettleton, D., Knetter, S., Christian, C., Ramer-Tait, A. E., Wannemuehler, M. J., & Tuggle, C. K. (2011). Distinct peripheral blood RNA responses to Salmonella in pigs differing in Salmonella shedding levels: intersection of IFNG, TLR and miRNA pathways. *PloS One*, *6*(12), e28768.

Huang, T.-H., Zhu, M.-J., Li, X.-Y., & Zhao, S.-H. (2008). Discovery of porcine microRNAs and profiling from skeletal muscle tissues during development. *PloS One*, *3*(9), e3225.

Huntoon, K. M., Wang, Y., Eppolito, C. A., Barbour, K. W., Berger, F. G., Shrikant, P. A., & Baumann, H. (2008). The acute phase protein haptoglobin regulates host immunity. *Journal of Leukocyte Biology*, *84*(1), 170–181.

Ibeagha-Awemu, E. M., & Zhao, X. (2015). Epigenetic marks: regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. *Frontiers in Genetics*, *6*, 302.

Iffland, P. H., Everett, M. E., Cobb-Pitstick, K. M., Bowser, L. E., Barnes, A. E., Babus, J. K., Romanowski, A. J., Baybis, M., Elziny, S., Puffenberger, E. G., Gonzaga-Jauregui, C., Poulopoulos, A., Carson, V. J., & Crino, P. B. (2022). NPRL3 loss alters neuronal morphology, mTOR localization, cortical lamination and seizure threshold. *Brain: A Journal of Neurology*, *145*(11), 3872–3885.

Itoh, H., Tamura, K., Izumi, M., Motoi, Y., Kidoguchi, K., & Funayama, Y. (1993). The influence of age and health status on the serum alpha 1-acid glycoprotein level of conventional and specific pathogen-free pigs. *Canadian Journal of Veterinary Research*, *57*(2), 74–78.

Iwasaki, A., & Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. *Science*, *327*(5963), 291–295.

Jégou, M., Gondret, F., Vincent, A., Tréfeu, C., Gilbert, H., & Louveau, I. (2016). Whole Blood Transcriptomics Is Relevant to Identify Molecular Changes in Response to Genetic Selection for Feed Efficiency and Nutritional Status in the Pig. *PloS One*, *11*(1), e0146550.

Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*, 318–356.

Jain, S., Gautam, V., & Naseem, S. (2011). Acute-phase proteins: As diagnostic tool. *Journal of Pharmacy & Bioallied Sciences*, *3*(1), 118–127.

Jørgensen, C. B., Cirera, S., Anderson, S. I., Archibald, A. L., Raudsepp, T., Chowdhary, B., Edfors-Lilja, I., Andersson, L., & Fredholm, M. (2003). Linkage and comparative mapping of the locus controlling susceptibility towards E. COLI F4ab/ac diarrhoea in pigs. *Cytogenetic and Genome Research*, *102*(1-4), 157–162.

Jaing, C., Rowland, R. R. R., Allen, J. E., Certoma, A., Thissen, J. B., Bingham, J., Rowe, B., White, J. R., Wynne, J. W., Johnson, D., Gaudreault, N. N., & Williams, D. T. (2017). Gene expression analysis of whole blood RNA from pigs infected with low and high pathogenic African swine fever viruses. *Scientific Reports*, *7*(1), 10115.

Jakab, G. J., Warr, G. A., & Sannes, P. L. (1980). Alveolar macrophage ingestion and phagosome-lysosome fusion defect associated with virus pneumonia. *Infection and Immunity*, *27*(3), 960–968.

Jeon, R. L., Cheng, J., Putz, A. M., Dong, Q., Harding, J. C. S., Dyck, M. K., Plastow, G. S., Fortin, F., Lunney, J., Rowland, R., Canada, P., & Dekkers, J. C. M. (2021). Effect of a genetic marker for the GBP5 gene on resilience to a polymicrobial natural disease challenge

in pigs. *Livestock Science*, *244*, 104399.

Jeon, R. L., Gilbert, C., Cheng, J., Putz, A. M., Dyck, M. K., Plastow, G. S., Fortin, F., PigGen Canada, Dekkers, J. C. M., & Harding, J. C. S. (2021). Proliferation of peripheral blood mononuclear cells from healthy piglets after mitogen stimulation as indicators of disease resilience. *Journal of Animal Science*, *99*(8). https://doi.org/10.1093/jas/skab084

Jeong, J., Kim, S., Park, K. H., Kang, I., Park, S.-J., Park, C., & Chae, C. (2017). Evaluation of the effect of a porcine reproductive and respiratory syndrome (PRRS) modified-live virus vaccine on sow reproductive performance in endemic PRRS farms. *Veterinary Microbiology*, *208*, 47–52.

Jiang, H., Tsang, L., Wang, H., & Liu, C. (2021). IFI44L as a Forward Regulator Enhancing Host Antituberculosis Responses. *Journal of Immunology Research*, *2021*, 5599408.

Jiang, Y., Shang, H., Xu, H., Ding, X., Zhao, L., Fang, L., & Chen, W. (2010). Detection and genotyping of porcine circovirus in naturally infected pigs by oligo-microarray. *Research in Veterinary Science*, *89*(1), 133–139.

Jiménez-Chillarón, J. C., Díaz, R., & Ramón-Krauel, M. (2014). Chapter 4 - Omics Tools for the Genome-Wide Analysis of Methylation and Histone Modifications. In C. Simó, A. Cifuentes, & V. García-Cañas (Eds.), *Comprehensive Analytical Chemistry* (Vol. 63, pp. 81–110). Elsevier.

Jo, P., Nietert, M., Gusky, L., Kitz, J., Conradi, L. C., Müller-Dornieden, A., Schüler, P., Wolff, H. A., Rüschoff, J., Ströbel, P., Grade, M., Liersch, T., Beißbarth, T., Ghadimi, M. B., Sax, U., & Gaedcke, J. (2016). Neoadjuvant Therapy in Rectal Cancer - Biobanking of Preoperative Tumor Biopsies. *Scientific Reports*, *6*, 35589.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews. Molecular Cell Biology*, *17*(7), 451–459.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews. Molecular Cell Biology*, *17*(7), 451–459.

Johnson, R. (2009). Potential Farm Sector Effects of 2009 H1N1 Swine Flu: Questions and

Answers. *Congressional Research Service, the Library of Congress.*

Jones, C. P., & Ferré-D'Amaré, A. R. (2017). Long-Range Interactions in Riboswitch Control of Gene Expression. *Annual Review of Biophysics*, *46*, 455–481.

Jones, G. W., Hill, D. G., Cardus, A., & Jones, S. A. (2018). IL-27: a double agent in the IL-6 family. *Clinical and Experimental Immunology*, *193*(1), 37–46.

Jordan, K. W., Nordenstam, J., Lauwers, G. Y., Rothenberger, D. A., Alavi, K., Garwood, M., & Cheng, L. L. (2009). Metabolomic characterization of human rectal adenocarcinoma with intact tissue magnetic resonance spectroscopy. *Diseases of the Colon and Rectum*, *52*(3), 520–525.

Joseph, B., Corwin, J. A., Li, B., Atwell, S., & Kliebenstein, D. J. (2013). Cytoplasmic genetic variation and extensive cytonuclear interactions influence natural variation in the metabolome. *eLife*, *2*, e00776.

Kadarmideen, H. N., Watson-Haigh, N. S., & Andronicos, N. M. (2011). Systems biology of ovine intestinal parasite resistance: disease gene modules and biomarkers. *Molecular bioSystems*, *7*(1), 235–246.

König, S., & May, K. (2019). Invited review: Phenotyping strategies and quantitative-genetic background of resistance, tolerance and resilience associated traits in dairy cattle. *Animal: An International Journal of Animal Bioscience*, *13*(5), 897–908.

Kanaji, S., Fahs, S. A., Shi, Q., Haberichter, S. L., & Montgomery, R. R. (2012). Contribution of platelet vs. endothelial VWF to platelet adhesion and hemostasis. *Journal of Thrombosis and Haemostasis: JTH*, *10*(8), 1646–1652.

Kankainen, M., Gopalacharyulu, P., Holm, L., & Orešič, M. (2011). MPEA—metabolite pathway enrichment analysis. *Bioinformatics* , *27*(13), 1878–1879.

Kannan, K., & Jain, S. K. (2000). Oxidative stress and apoptosis. *Pathophysiology: The Official Journal of the International Society for Pathophysiology / ISP*, *7*(3), 153–163.

Kanter, I., & Kalisky, T. (2015). Single cell transcriptomics: methods and applications. *Frontiers in Oncology*, *5*, 53.

Kapetanovic, R., Fairbairn, L., Downing, A., Beraldi, D., Sester, D. P., Freeman, T. C., Tuggle, C. K., Archibald, A. L., & Hume, D. A. (2013). The impact of breed and tissue compartment on the response of pig macrophages to lipopolysaccharide. *BMC Genomics*, *14*, 581.

Kasture, V. S., Musmade, D. S., Vakte, M. B., Sonawane, S. B., & Patil, P. P. (2012). Metabolomics: current technologies and future trends. *Int J Res Dev Pharm Life Sci*, 2, 206-217.

Katz, E., Li, J.-J., Jaegle, B., Ashkenazy, H., Abrahams, S. R., Bagaza, C., Holden, S., Pires, C. J., Angelovici, R., & Kliebenstein, D. J. (2021). Genetic variation, environment and demography intersect to shape Arabidopsis defense metabolite variation across Europe. *eLife*, *10*. https://doi.org/10.7554/eLife.67784

Kause, A., & Ødegård, J. (2012). The genetic analysis of tolerance to infections: a review. *Frontiers in Genetics*, 3.

Kause, A., van Dalen, S., & Bovenhuis, H. (2012). Genetics of ascites resistance and tolerance in chicken: a random regression approach. *G3* , *2*(5), 527–535.

Kell, D. B., & Oliver, S. G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *26*(1), 99–105.

Kilka, S., Erdmann, F., Migdoll, A., Fischer, G., & Weiwad, M. (2009). The proline-rich N-terminal sequence of calcineurin Abeta determines substrate binding. *Biochemistry*, *48*(9), 1900–1910.

Kim, J. B., Zeng, Z. X., Nam, Y. J., Kim, Y., Yang, S. L., Wu, X., & Lee, C. (2005). Association of mahogany/attractin gene (ATRN) with porcine growth and fat. *Asian-Australasian Journal of Animal Sciences*, *18*(10), 1383–1386.

Kim, J. S., Ingale, S. L., Hosseindoust, A. R., Lee, S. H., Lee, J. H., & Chae, B. J. (2017). Effects of mannan level and β-mannanase supplementation on growth performance, apparent total tract digestibility and blood metabolites of growing pigs. *Animal: An International Journal*

*of Animal Bioscience*, *11*(2), 202–208.

Kimman, T. G., Cornelissen, L. A., Moormann, R. J., Rebel, J. M. J., & Stockhofe-Zurwieden, N. (2009). Challenges for porcine reproductive and respiratory syndrome virus (PRRSV) vaccinology. *Vaccine*, *27*(28), 3704–3718.

King, R. H., & Pluske, J. R. (2003). Nutritional management of the pig in preparation for weaning. *Weaning the pig-concepts and consequences*, 37-51.

Klein, M. S., Buttchereit, N., Miemczyk, S. P., Immervoll, A.-K., Louis, C., Wiedemann, S., Junge, W., Thaller, G., Oefner, P. J., & Gronwald, W. (2012). NMR metabolomic analysis of dairy cows reveals milk glycerophosphocholine to phosphocholine ratio as prognostic biomarker for risk of ketosis. *Journal of Proteome Research*, 11(2), 1373–1381.

Knap, P. W. (2005). Breeding robust pigs. *Australian Journal of Experimental Agriculture*, *45*(8), 763–773.

Knap, P. W., & Bishop, S. C. (2000). Relationships between genetic change and infectious disease in domestic livestock. *BSAP Occasional Publication*, *27*, 65–80.

Knap, P. W., & Bishop, S. C. (2000). Relationships between genetic change and infectious disease in domestic livestock. *BSAP Occasional Publication*, *27*, 65–80.

Ko, C.-Y., Chang, W.-C., & Wang, J.-M. (2015). Biological roles of CCAAT/Enhancer-binding protein delta during inflammation. *Journal of Biomedical Science*, *22*(1), 6.

Koch, B. J., Hungate, B. A., & Price, L. B. (2017). Food-animal production and the spread of antibiotic resistance: the role of ecology. *Frontiers in Ecology and the Environment*, *15*(6), 309–318.

Koenderman, L., Buurman, W., & Daha, M. R. (2014). The innate immune response. *Immunology Letters*, *162*(2 Pt B), 95–102.

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., & Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Molecular Cell*, *58*(4), 610–620.

Kommadath, A., Bao, H., Arantes, A. S., Plastow, G. S., Tuggle, C. K., Bearson, S. M. D., Guan, L. L., & Stothard, P. (2014). Gene co-expression network analysis identifies porcine genes

associated with variation in Salmonella shedding. *BMC Genomics*, *15*(1), 452.

Kommadath, A., Bao, H., Arantes, A. S., Plastow, G. S., Tuggle, C. K., Bearson, S. M. D., Guan, L. L., & Stothard, P. (2014). Gene co-expression network analysis identifies porcine genes associated with variation in Salmonella shedding. *BMC Genomics*, *15*(1), 452.

Kommadath, A., Bao, H., Choi, I., Reecy, J. M., Koltes, J. E., Fritz-Waters, E., Eisley, C. J., Grant, J. R., Rowland, R. R. R., Tuggle, C. K., Dekkers, J. C. M., Lunney, J. K., Guan, L. L., Stothard, P., & Plastow, G. S. (2017). Genetic architecture of gene expression underlying variation in host response to porcine reproductive and respiratory syndrome virus infection. *Scientific Reports*, *7*, 46203.

Kong, I. Y., Trezise, S., Light, A., Todorovski, I., Arnau, G. M., Gadipally, S., Yoannidis, D., Simpson, K. J., Dong, X., Whitehead, L., Tempany, J. C., Farchione, A. J., Sheikh, A. A., Groom, J. R., Rogers, K. L., Herold, M. J., Bryant, V. L., Ritchie, M. E., Willis, S. N., … Hawkins, E. D. (2022). Epigenetic modulators of B cell fate identified through coupled phenotype-transcriptome analysis. *Cell Death and Differentiation*, *29*(12), 2519–2530.

Kothalawala, H., Toussaint, M. J. M., & Gruys, E. (2006). An overview of swine influenza. *The Veterinary Quarterly*, *28*(2), 46–53.

Kouznetsova, V. L., Kim, E., Romm, E. L., Zhu, A., & Tsigelny, I. F. (2019). Recognition of early and late stages of bladder cancer using metabolites and machine learning. *Metabolomics: Official Journal of the Metabolomic Society*, *15*(7), 94.

Kramer, L. M., Teng, J., Lim, K. S., Gao, Y., Yin, H., Bai, L., Liu, G. E., Zhang, Z., Fang, L., Plastow, G. S., Tuggle, C. K., & Dekkers, J. C. M. (2022). Large-scale cis-eQTL analysis of gene expression in blood of young healthy pigs using PigGTEx. In *Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP)* (pp. 2321–2324). Wageningen Academic Publishers.

Kramer, T. T., Griffith, R. W., & Saucke, L. (1985). Iron and transferrin in acute experimental Salmonella cholerae-suis infection in pigs. *American Journal of Veterinary Research*, *46*(2), 451–455.

Krueger, W. S., & Gray, G. C. (2012). Swine influenza virus infections in man. *Swine Influenza*, 201-225.

Kumar, H., Kawai, T., & Akira, S. (2011). Pathogen recognition by the innate immune system. *International Reviews of Immunology*, *30*(1), 16–34.

Lacroix-Desmazes, S., Kaveri, S. V., Mouthon, L., Ayouba, A., Malanchère, E., Coutinho, A., & Kazatchkine, M. D. (1998). Self-reactive antibodies (natural autoantibodies) in healthy individuals. *Journal of Immunological Methods*, *216*(1-2), 117–137.

Laghouaouta, H., Fraile, L., Suárez-Mesa, R., Ros-Freixedes, R., Estany, J., & Pena, R. N. (2022). A genome-wide screen for resilient responses in growing pigs. *Genetics, Selection, Evolution: GSE*, *54*(1), 50.

Laghouaouta, H., Pena, R. N., Ros-Freixedes, R., Reixach, J., Díaz, M., Estany, J., Armengol, R., Bassols, A., & Fraile, L. (2021). A Methodology to Quantify Resilience in Growing Pigs. *Animals : An Open Access Journal from MDPI*, *11*(10). https://doi.org/10.3390/ani11102970

Lalla, E., Lamster, I. B., Hofmann, M. A., Bucciarelli, L., Jerud, A. P., Tucker, S., Lu, Y., Papapanou, P. N., & Schmidt, A. M. (2003). Oral Infection With a Periodontal Pathogen Accelerates Early Atherosclerosis in Apolipoprotein E–Null Mice. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *23*(8), 1405–1411.

Lampreave, F., González-Ramón, N., Martínez-Ayensa, S., Hernández, M. A., Lorenzo, H. K., García-Gil, A., & Piñeiro, A. (1994). Characterization of the acute phase serum protein response in pigs. *Electrophoresis*, *15*(5), 672–676.

Laplana, M., Estany, J., Fraile, L. J., & Pena, R. N. (2020). Resilience Effects of SGK1 and TAP1 DNA Markers during PRRSV Outbreaks in Reproductive Sows. *Animals : An Open Access Journal from MDPI*, 10(5).

Lenz, E. M., & Wilson, I. D. (2007). Analytical strategies in metabonomics. *Journal of Proteome Research*, *6*(2), 443–458.

Li, X., Galliher-Beckley, A., Nietfeld, J. C., Faaberg, K. S., & Shi, J. (2013). MontanideTM Gel01 ST adjuvant enhances PRRS modified live vaccine efficacy by regulating porcine

humoral and cellular immune responses. *World Journal of Vaccines*.

Li, Y., Liu, H., Wang, P., Wang, L., Sun, Y., Liu, G., Zhang, P., Kang, L., Jiang, S., & Jiang, Y. (2016). RNA-Seq Analysis Reveals Genes Underlying Different Disease Responses to Porcine Circovirus Type 2 in Pigs. *PloS One*, *11*(5), e0155502.

Lillico, S. G., Proudfoot, C., King, T. J., Tan, W., Zhang, L., Mardjuki, R., Paschon, D. E., Rebar, E. J., Urnov, F. D., Mileham, A. J., McLaren, D. G., & Whitelaw, C. B. A. (2016). Mammalian interspecies substitution of immune modulatory alleles by genome editing. *Scientific Reports*, *6*, 21645.

Lim, K.-S., Cheng, J., Putz, A., Dong, Q., Bai, X., Beiki, H., Tuggle, C. K., Dyck, M. K., Canada, P. G., Fortin, F., Harding, J. C. S., Plastow, G. S., & Dekkers, J. C. M. (2021). Quantitative analysis of the blood transcriptome of young healthy pigs and its relationship with subsequent disease resilience. *BMC Genomics*, *22*(1), 614.

Lim, K.-S., Cheng, J., Tuggle, C. K., Dyck, M., Canada, P., Fortin, F., Harding, J., Plastow, G., & Dekkers, J. C. (2022). PSXIII-B-14 Investigation of the Blood Transcriptome of Young Healthy Pigs to Identify Genetic Indicators for Disease Resilience. *Journal of Animal Science*, *100*(Supplement_3), 325–325.

Lincoln, J. A., Lefkowitz, D. L., Cain, T., Castro, A., Mills, K. C., Lefkowitz, S. S., ... & Bollen, A. (1995). Exogenous myeloperoxidase enhances bacterial phagocytosis and intracellular killing by macrophages. I*nfection and immunity*, 63(8), 3042-3047.

Link, H., Fuhrer, T., Gerosa, L., Zamboni, N., & Sauer, U. (2015). Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nature Methods*, *12*(11), 1091–1097.

Lisuzzo, A., Laghi, L., Faillace, V., Zhu, C., Contiero, B., Morgante, M., Mazzotta, E., Gianesella, M., & Fiore, E. (2022). Differences in the serum metabolome profile of dairy cows according to the BHB concentration revealed by proton nuclear magnetic resonance spectroscopy (1H-NMR). *Scientific Reports*, *12*(1), 2525.

Liu, J., & McFadden, G. (2015). SAMD9 is an innate antiviral host factor with stress response

properties that can be antagonized by poxviruses. *Journal of Virology*, *89*(3), 1925–1931.

Liu, Q., Zhou, Y.-H., & Yang, Z.-Q. (2016). The cytokine storm of severe influenza and development of immunomodulatory therapy. *Cellular & Molecular Immunology*, *13*(1), 3–10.

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. *PLoS Computational Biology*, *13*(5), e1005457.

Lowen, A. C., Mubareka, S., Steel, J., & Palese, P. (2007). Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathogens*, *3*(10), 1470–1476.

Lowrie, M., Penderis, J., Eckersall, P. D., McLaughlin, M., Mellor, D., & Anderson, T. J. (2009). The role of acute phase proteins in diagnosis and management of steroid-responsive meningitis arteritis in dogs. *Veterinary Journal* , *182*(1), 125–130.

Luan, Y.-Y., & Yao, Y.-M. (2018). The Clinical Significance and Potential Role of C-Reactive Protein in Chronic Inflammatory and Neurodegenerative Diseases. *Frontiers in Immunology*, *9*, 1302.

Lunney, J. K., & Chen, H. (2010). Genetic control of host resistance to porcine reproductive and respiratory syndrome virus (PRRSV) infection. *Virus Research*, *154*(1-2), 161–169.

Lunney, J. K., Fang, Y., Ladinig, A., Chen, N., Li, Y., Rowland, B., & Renukaradhya, G. J. (2016). Porcine Reproductive and Respiratory Syndrome Virus (PRRSV): Pathogenesis and Interaction with the Immune System. *Annual Review of Animal Biosciences*, *4*, 129–154.

Lunney, J. K., Steibel, J. P., Reecy, J. M., Fritz, E., Rothschild, M. F., Kerrigan, M., Trible, B., & Rowland, R. R. (2011). Probing genetic control of swine responses to PRRSV infection: current progress of the PRRS host genetics consortium. *BMC Proceedings*, *5 Suppl 4*(Suppl 4), S30.

Luster, A. D. (2002). The role of chemokines in linking innate and adaptive immunity. *Current Opinion in Immunology*, *14*(1), 129–135.

Lv, J., Ding, Y., Liu, X., Pan, L., Zhang, Z., Zhou, P., Zhang, Y., & Hu, Y. (2018). Gene expression analysis of porcine whole blood cells infected with foot-and-mouth disease virus

using high-throughput sequencing technology. *PloS One*, *13*(7), e0200081.

Machado-Fragua, M. D., Landré, B., Chen, M., Fayosse, A., Dugravot, A., Kivimaki, M., Sabia, S., & Singh-Manoux, A. (2022). Circulating serum metabolites as predictors of dementia: a machine learning approach in a 21-year follow-up of the Whitehall II cohort study. *BMC Medicine*, *20*(1), 334.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*, 281–297.

Maes, D., Segales, J., Meyns, T., Sibila, M., Pieters, M., & Haesebrouck, F. (2008). Control of Mycoplasma hyopneumoniae infections in pigs. *Veterinary Microbiology*, *126*(4), 297–309.

Maes, D., Sibila, M., & Pieters, M. (2021). *Mycoplasmas in Swine*. CABI.

Magar, R., Larochelle, R., Thibault, S., & Lamontagne, L. (2000). Experimental transmission of porcine circovirus type 2 (PCV2) in weaned pigs: a sequential study. *Journal of Comparative Pathology*, *123*(4), 258–269.

Maghazachi, A. A., Al-Aoukaty, A., & Schall, T. J. (1996). CC chemokines induce the generation of killer cells from CD56+ cells. *European Journal of Immunology*, *26*(2), 315–319.

Magnusson, U., Wilkie, B., Artursson, K., & Mallard, B. (1999). Interferon-alpha and haptoglobin in pigs selectively bred for high and low immune response and infected with Mycoplasma hyorhinis. *Veterinary Immunology and Immunopathology*, *68*(2-4), 131–137.

Mahalingam, S., Karupiah, G., Takeda, K., Akira, S., Matthaei, K. I., & Foster, P. S. (2001). Enhanced resistance in STAT6-deficient mice to infection with ectromelia virus. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(12), 6812–6817.

Maher, C. A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., Sam, L., Barrette, T., Palanisamy, N., & Chinnaiyan, A. M. (2009). Transcriptome sequencing to detect gene fusions in cancer. *Nature*, *458*(7234), 97–101.

Mair, K. H., Sedlak, C., Käser, T., Pasternak, A., Levast, B., Gerner, W., Saalmüller, A.,

Summerfield, A., Gerdts, V., Wilson, H. L., & Meurens, F. (2014). The porcine innate immune system: an update. *Developmental and Comparative Immunology*, *45*(2), 321–343.

Marra, M. A., Hillier, L., & Waterston, R. H. (n.d.). Expressed sequence tags—ESTablishing bridges between genomes. *Trends in Genetics: TIG*.

Martelli, P., & Segalés, J. (2021). *Swine Respiratory Disease*. Grupo Asís Biomedia S.L.

Mavrommatis, B., Offord, V., Patterson, R., Watson, M., Kanellos, T., Steinbach, F., Grierson, S., & Werling, D. (2014). Global gene expression profiling of myeloid immune cell subsets in response to in vitro challenge with porcine circovirus 2b. *PloS One*, *9*(3), e91081.

Mazé-Guilmo, E., Loot, G., Páez, D. J., Lefèvre, T., & Blanchet, S. (2014). Heritable variation in host tolerance and resistance inferred from a wild host–parasite system. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1779), 20132567.

McCabe, M., Waters, S., Morris, D., Kenny, D., Lynn, D., & Creevey, C. (2012). RNA-seq analysis of differential gene expression in liver from lactating dairy cows divergent in negative energy balance. *BMC Genomics*, *13*(1), 1–11.

McCleary, S., Strong, R., McCarthy, R. R., Edwards, J. C., Howes, E. L., Stevens, L. M., Sánchez-Cordón, P. J., Núñez, A., Watson, S., Mileham, A. J., Lillico, S. G., Tait-Burkard, C., Proudfoot, C., Ballantyne, M., Whitelaw, C. B. A., Steinbach, F., & Crooke, H. R. (2020). Substitution of warthog NF-κB motifs into RELA of domestic pigs is not sufficient to confer resilience to African swine fever virus. *Scientific Reports*, *10*(1), 8951.

McGuckin, M. M., Giesy, S. L., Davis, A. N., Abyeta, M. A., Horst, E. A., Saed Samii, S., Zang, Y., Butler, W. R., Baumgard, L. H., McFadden, J. W., & Boisclair, Y. R. (2020). The acute phase protein orosomucoid 1 is upregulated in early lactation but does not trigger appetite-suppressing STAT3 signaling via the leptin receptor. *Journal of Dairy Science*, *103*(5), 4765–4776.

Medzhitov, R., Schneider, D. S., & Soares, M. P. (2012). Disease tolerance as a defense strategy. *Science*, *335*(6071), 936–941.

Meftahi, G. H., Neuroscience Research Center, Baqiyatallah University of Medical Sciences,

Tehran, Iran, Bahari, Z., Jangravi, Z., Iman, M., Neuroscience Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran;, Department of Physiology and Medical Physics, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran;, Nanobiotechnology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran;, Department of Biochemistry, Faculty of Medicine, Baqiyatallah University of Medical Sciences, Tehran, Iran;, Nanobiotechnology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran;, & Department of Pharmaceutics, Faculty of Pharmacy, Baqiyatallah University of Medical Sciences, Tehran, Iran; (2021). A vicious circle between oxidative stress and cytokine storm in acute respiratory distress syndrome pathogenesis at COVID-19 infection. *Ukrainian Biochemical Journal*, *93*(1), 18–29.

Meijerink, E., Fries, R., Vögeli, P., Masabanda, J., Wigger, G., Stricker, C., Neuenschwander, S., Bertschinger, H. U., & Stranzinger, G. (1997). Two alpha(1,2) fucosyltransferase genes on porcine chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, *8*(10), 736–741.

Meijerink, E., Neuenschwander, S., Fries, R., Dinter, A., Bertschinger, H. U., Stranzinger, G., & Vögeli, P. (2000). A DNA polymorphism influencing alpha(1,2)fucosyltransferase activity of the pig FUT1 enzyme determines susceptibility of small intestinal epithelium to Escherichia coli F18 adhesion. *Immunogenetics*, *52*(1-2), 129–136.

Melé, M., Ferreira, P. G., Reverter, F., DeLuca, D. S., Monlong, J., Sammeth, M., Young, T. R., Goldmann, J. M., Pervouchine, D. D., Sullivan, T. J., Johnson, R., Segrè, A. V., Djebali, S., Niarchou, A., GTEx Consortium, Wright, F. A., Lappalainen, T., Calvo, M., Getz, G., … Guigó, R. (2015). Human genomics. The human transcriptome across tissues and individuals. *Science*, *348*(6235), 660–665.

Melnikow, E., Dornan, S., Sargent, C., Duszenko, M., Evans, G., Gunkel, N., Selzer, P. M., & Ullrich, H. J. (2005). Microarray analysis of Haemophilus parasuis gene expression under in vitro growth conditions mimicking the in vivo environment. *Veterinary Microbiology*,

*110*(3-4), 255–263.

Melzer, N., Wittenburg, D., Hartwig, S., Jakubowski, S., Kesting, U., Willmitzer, L., Lisec, J., Reinsch, N., & Repsilber, D. (2013). Investigating associations between milk metabolite profiles and milk traits of Holstein cows. *Journal of Dairy Science*, *96*(3), 1521–1534.

Memar, M. Y., Alizadeh, N., Varshochi, M., & Kafil, H. S. (2019). Immunologic biomarkers for diagnostic of early-onset neonatal sepsis. *The Journal of Maternal-Fetal & Neonatal Medicine: The Official Journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians*, *32*(1), 143–153.

Meng, X.-J. (2013). Porcine circovirus type 2 (PCV2): pathogenesis and interaction with the immune system. *Annual Review of Animal Biosciences*, *1*, 43–64.

Meroni, G., Zamarian, V., Prussiani, C., Bronzo, V., Lecchi, C., Martino, P. A., & Ceciliani, F. (2019). The bovine acute phase protein α1-acid glycoprotein (AGP) can disrupt Staphylococcus aureus biofilm. *Veterinary Microbiology*, *235*, 93–100.

Micheel, C. M., Nass, S. J., Omenn, G. S., Board on Health Care Services, Board on Health Sciences Policy, & Institute of Medicine. (2012). *Omics-Based Clinical Discovery: Science, Technology, and Applications*. National Academies Press (US).

Miller, M. R., White, A., & Boots, M. (2006). The evolution of parasites in response to tolerance in their hosts: the good, the bad, and apparent commensalism. *Evolution; International Journal of Organic Evolution*, *60*(5), 945–956.

Millership, S., Ninkina, N., Rochford, J. J., & Buchman, V. L. (2013). γ-synuclein is a novel player in the control of body lipid metabolism. *Adipocyte*, *2*(4), 276–280.

Mills, E., & O'Neill, L. A. (2014). Succinate: a metabolic signal in inflammation. *Trends in cell biology*, 24(5), 313-320.

Minnich, M., Tagoh, H., Bönelt, P., Axelsson, E., Fischer, M., Cebolla, B., Tarakhovsky, A., Nutt, S. L., Jaritz, M., & Busslinger, M. (2016). Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. *Nature Immunology*, *17*(3), 331–343.

Mittal, M., Siddiqui, M. R., Tran, K., Reddy, S. P., & Malik, A. B. (2014). Reactive Oxygen Species in Inflammation and Tissue Injury. *Antioxidants & Redox Signaling*, *20*(7), 1126–1167.

Mohseni, M., & Chishti, A. H. (2008). The headpiece domain of dematin regulates cell shape, motility, and wound healing by modulating RhoA activation. *Molecular and cellular biology*, 28(15), 4712-4718.

Monger, X. C., Gilbert, A.-A., Saucier, L., & Vincent, A. T. (2021). Antibiotic Resistance: From Pig to Meat. *Antibiotics (Basel, Switzerland)*, 10(10).

Moran, K., Wilcock, P., Elsbernd, A., Zier-Rush, C., Boyd, R. D., & van Heugten, E. (2019). Effects of super-dosing phytase and inositol on growth performance and blood metabolites of weaned pigs housed under commercial conditions1. *Journal of Animal Science*, *97*(7), 3007–3015.

Morris, C. A., Bisset, S. A., Vlassoff, A., Wheeler, M., West, C. J., Devantier, B. P., & Mackay, A. D. (2010). Selecting for resilience in Romney sheep under nematode parasite challenge, 1994–2007. *New Zealand Journal of Agricultural Research*, *53*(3), 245–261.

Moser, B. (2015). CXCR5, the Defining Marker for Follicular B Helper T (TFH) Cells. *Frontiers in Immunology*, *6*, 296.

Mudaliar, M., Thomas, F. C., & Eckersall, P. D. (2017). Omic Approaches to a Better Understanding of Mastitis in Dairy Cows. In B. N. Ametaj (Ed.), *Periparturient Diseases of Dairy Cows: A Systems Biology Approach* (pp. 139–183). Springer International Publishing.

Mulder, H. A., & Rashidi, H. (2017). Selection on resilience improves disease resistance and tolerance to infections. *Journal of Animal Science*, *95*(8), 3346–3358.

Muñoz-Pinedo, C., El Mjiyad, N., & Ricci, J.-E. (2012). Cancer metabolism: current perspectives and future directions. *Cell Death & Disease*, *3*(1), e248.

Murata, H., Shimada, N., & Yoshioka, M. (2004). Current research on acute phase proteins in veterinary diagnosis: an overview. *Veterinary Journal* , *168*(1), 28–40.

Murata, H., Shimada, N., & Yoshioka, M. (2004). Current research on acute phase proteins in

veterinary diagnosis: an overview. *Veterinary Journal* , *168*(1), 28–40.

Murtaugh, M. P., Baarsch, M. J., Zhou, Y., Scamurra, R. W., & Lin, G. (1996). Inflammatory cytokines in animal health and disease. *Veterinary Immunology and Immunopathology*, *54*(1-4), 45–55.

Mussá, T., Rodríguez-Cariño, C., Sánchez-Chardi, A., Baratelli, M., Costa-Hurtado, M., Fraile, L., Domínguez, J., Aragon, V., & Montoya, M. (2012). Differential interactions of virulent and non-virulent H. parasuis strains with naïve or swine influenza virus pre-infected dendritic cells. *Veterinary Research*, *43*(1), 80.

Nakayama, T., Hirahara, K., Onodera, A., Endo, Y., Hosokawa, H., Shinoda, K., Tumes, D. J., & Okamoto, Y. (2017). Th2 Cells in Health and Disease. *Annual Review of Immunology*, *35*, 53–84.

Nathan, C., & Shiloh, M. U. (2000). Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(16), 8841–8848.

Nazki, S., Khatun, A., Jeong, C.-G., Mattoo, S. U. S., Gu, S., Lee, S.-I., Kim, S.-C., Park, J.-H., Yang, M.-S., Kim, B., Park, C.-K., Lee, S.-M., & Kim, W.-I. (2020). Evaluation of local and systemic immune responses in pigs experimentally challenged with porcine reproductive and respiratory syndrome virus. *Veterinary Research*, *51*(1), 66.

Nelms, K., Keegan, A. D., Zamorano, J., Ryan, J. J., & Paul, W. E. (1999). The IL-4 receptor: signaling mechanisms and biologic functions. *Annual review of immunology*, 17(1), 701-738.

Nelson, E. A., Christopher-Hennings, J., & Benfield, D. A. (1994). Serum immune responses to the proteins of porcine reproductive and respiratory syndrome (PRRS) virus. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *6*(4), 410–415.

Nery da Silva, A., Silva Araujo, M., Pértille, F., & Zanella, A. J. (2021). How epigenetics can enhance pig welfare?. *Animals*, 12(1), 32.

Neumann, E. J., Kliebenstein, J. B., Johnson, C. D., Mabry, J. W., Bush, E. J., Seitzinger, A. H., Green, A. L., & Zimmerman, J. J. (2005). Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. *Journal of the American Veterinary Medical Association*, *227*(3), 385–392.

Ng, P. M. L., Jin, Z., Tan, S. S. H., Ho, B., & Ding, J. L. (2004). C-reactive protein: a predominant LPS-binding acute phase protein responsive to Pseudomonas infection. *Journal of Endotoxin Research*, *10*(3), 163–174.

Ni, L., Song, C., Wu, X., Zhao, X., Wang, X., Li, B., & Gan, Y. (2019). RNA-seq transcriptome profiling of porcine lung from two pig breeds in response to Mycoplasma hyopneumoniae infection. *PeerJ*, *7*, e7900.

Nicholson, J. K., O'Flynn, M. P., Sadler, P. J., Macleod, A. F., Juul, S. M., & Sönksen, P. H. (1984). Proton-nuclear-magnetic-resonance studies of serum, plasma and urine from fasting normal and diabetic subjects. *Biochemical Journal*, *217*(2), 365–375.

Nicholson, T. L., Kukielka, D., Vincent, A. L., Brockmeier, S. L., Miller, L. C., & Faaberg, K. S. (2011). Utility of a panviral microarray for detection of swine respiratory viruses in clinical samples. *Journal of Clinical Microbiology*, *49*(4), 1542–1548.

Nielsen, M. J., & Moestrup, S. K. (2009). Receptor targeting of hemoglobin mediated by the haptoglobins: roles beyond heme scavenging. *Blood*, *114*(4), 764–771.

Niemann, C. U., Abrink, M., Pejler, G., Fischer, R. L., Christensen, E. I., Knight, S. D., & Borregaard, N. (2007). Neutrophil elastase depends on serglycin proteoglycan for localization in granules. *Blood*, *109*(10), 4478–4486.

Nikas, J. B., & Low, W. C. (2011). Application of clustering analyses to the diagnosis of Huntington disease in mice and other diseases with well-defined group boundaries. *Computer Methods and Programs in Biomedicine*, *104*(3), e133–e147.

Nikas, J. B., & Low, W. C. (2011). Application of clustering analyses to the diagnosis of Huntington disease in mice and other diseases with well-defined group boundaries. *Computer Methods and Programs in Biomedicine*, *104*(3), e133–e147.

Nisha, A. (2008). Antibiotic residues - A global health hazard. *Veterinary World*, *2*(2), 375.

Nodelijk, G. (2002). Porcine reproductive and respiratory syndrome (PRRS) with special reference to clinical aspects and diagnosis. A review. *The Veterinary Quarterly*, *24*(2), 95–100.

Nounamo, B., Li, Y., O'Byrne, P., Kearney, A. M., Khan, A., & Liu, J. (2017). An interaction domain in human SAMD9 is essential for myxoma virus host-range determinant M062 antagonism of host anti-viral function. *Virology*, *503*, 94–102.

Nyamundanda, G., Brennan, L., & Gormley, I. C. (2010). Probabilistic principal component analysis for metabolomic data. *BMC Bioinformatics*, *11*, 571.

Opriessnig, T., Fenaux, M., Thomas, P., Hoogland, M. J., Rothschild, M. F., Meng, X. J., & Halbur, P. G. (2006). Evidence of breed-dependent differences in susceptibility to porcine circovirus type-2-associated disease and lesions. *Veterinary Pathology*, *43*(3), 281–293.

Opriessnig, T., Yu, S., Thacker, E. L., & Halbur, P. G. (2004). Derivation of porcine circovirus type 2-negative pigs from positive breeding herds. *Journal of Swine Health and Production*, *12*(4), 186–191.

Pértille, F., Da Silva, V. H., Johansson, A. M., Lindström, T., Wright, D., Coutinho, L. L., ... & Guerrero-Bosagna, C. (2019). Mutation dynamics of CpG dinucleotides during a recent event of vertebrate diversification. *Epigenetics*, 14(7), 685-707.

Palgrave, C. J., Gilmour, L., Lowden, C. S., Lillico, S. G., Mellencamp, M. A., & Whitelaw, C. B. A. (2011). Species-specific variation in RELA underlies differences in NF-κB activity: a potential role in African swine fever pathogenesis. *Journal of Virology*, *85*(12), 6008–6014.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, *40*(12), 1413–1415.

Panawala, L. (2017). *Difference Between Lymphocytes and Phagocytes*.

Papatsiros, V. G., Alexopoulos, C., Kritas, S. K., Koptopoulos, G., Nauwynck, H. J., Pensaert, M. B., & Kyriakis, S. C. (2006). Long-term administration of a commercial porcine

reproductive and respiratory syndrome virus (PRRSV)-inactivated vaccine in PRRSV-endemically infected sows. *Journal of Veterinary Medicine. B, Infectious Diseases and Veterinary Public Health*, *53*(6), 266–272.

Park, Y.-J., Yoo, S.-A., Kim, M., & Kim, W.-U. (2020). The Role of Calcium–Calcineurin–NFAT Signaling Pathway in Health and Autoimmune Diseases. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.00195

Parkin, J., & Cohen, B. (2001). An overview of the immune system. *The Lancet*, *357*(9270), 1777–1789.

Parmar, S., Shah, N., Kasarwala, M., Virpura, M., & Prajapati, D. D. (2011). A review on swine flu. *J Pharm Sci Bioscientific Res*, 1(1), 11-17.

Pascall, J. C., Webb, L. M. C., Eskelinen, E.-L., Innocentin, S., Attaf-Bouabdallah, N., & Butcher, G. W. (2018). GIMAP6 is required for T cell maintenance and efficient autophagy in mice. *PloS One*, *13*(5), e0196504.

Patterson, B. K., Seethamraju, H., Dhody, K., Corley, M. J., Kazempour, K., Lalezari, J. P., Pang, A. P., Sugai, C., Francisco, E. B., Pise, A., Rodrigues, H., Ryou, M., Wu, H. L., Webb, G. M., Park, B. S., Kelly, S., Pourhassan, N., Lelic, A., Kdouh, L., … Sacha, J. B. (2020). Disruption of the CCL5/RANTES-CCR5 Pathway Restores Immune Homeostasis and Reduces Plasma Viral Load in Critical COVID-19. *medRxiv : The Preprint Server for Health Sciences*. https://doi.org/10.1101/2020.05.02.20084673

Pepys, M. B., & Hirschfield, G. M. (2003). C-reactive protein: a critical update. *The Journal of Clinical Investigation*, *111*(12), 1805–1812.

Persicke, M., Rückert, C., Plassmeier, J., Stutz, L. J., Kessler, N., Kalinowski, J., Goesmann, A., & Neuweger, H. (2012). MSEA: metabolite set enrichment analysis in the MeltDB metabolomics software platform: metabolic profiling of Corynebacterium glutamicum as an example. *Metabolomics: Official Journal of the Metabolomic Society*, *8*(2), 310–322.

Pertea, M. (2012). The human transcriptome: an unfinished story. *Genes*, *3*(3), 344–360.

Petersen, H. H., Nielsen, J. P., & Heegaard, P. M. H. (2004). Application of acute phase protein

measurements in veterinary clinical chemistry. *Veterinary Research*, *35*(2), 163–187.

Piñeiro, C., Piñeiro, M., Morales, J., Carpintero, R., Campbell, F. M., Eckersall, P. D., Toussaint, M. J. M., Alava, M. A., & Lampreave, F. (2007). Pig acute-phase protein levels after stress induced by changes in the pattern of food administration. *Animal: An International Journal of Animal Bioscience*, *1*(1), 133–139.

Podolska, A., Anthon, C., Bak, M., Tommerup, N., Skovgaard, K., Heegaard, P. M., Gorodkin, J., Cirera, S., & Fredholm, M. (2012). Profiling microRNAs in lung tissue from pigs infected with *Actinobacillus pleuropneumoniae*. *BMC Genomics*, *13*, 459.

Pohl, S., Bertschinger, H. U., Frederiksen, W., & Mannheim, W. (1983). Transfer of Haemophilus pleuropneumoniae and the Pasteurella haemolytica-Like Organism Causing Porcine Necrotic Pleuropneumonia to the Genus Actinobacillus (*Actinobacillus pleuropneumoniae* comb. nov.) on the Basis of Phenotypic and Deoxyribonucleic Acid Relatedness. *International Journal of Systematic Bacteriology*, *33*(3), 510–514.

Pol, J. M., van Leengoed, L. A., Stockhofe, N., Kok, G., & Wensvoort, G. (1997). Dual infections of PRRSV/influenza or PRRSV/*Actinobacillus pleuropneumoniae* in the respiratory tract. *Veterinary Microbiology*, *55*(1-4), 259–264.

Pomorska-Mól, M., Markowska-Daniel, I., Kwit, K., Stępniewska, K., & Pejsak, Z. (2013). C-reactive protein, haptoglobin, serum amyloid A and pig major acute phase protein response in pigs simultaneously infected with H1N1 swine influenza virus and Pasteurella multocida. *BMC Veterinary Research*, *9*, 14.

Pomorska-Mól, M., Markowska-Daniel, I., Kwit, K., Stępniewska, K., & Pejsak, Z. (2013). C-reactive protein, haptoglobin, serum amyloid A and pig major acute phase protein response in pigs simultaneously infected with H1N1 swine influenza virus and Pasteurella multocida. *BMC Veterinary Research*, *9*, 14.

Pomorska-Mól, M., Markowska-Daniel, I., Kwit, K., Stępniewska, K., & Pejsak, Z. (2013). C-reactive protein, haptoglobin, serum amyloid A and pig major acute phase protein response in pigs simultaneously infected with H1N1 swine influenza virus and Pasteurella

multocida. *BMC Veterinary Research*, *9*, 14.

Pos, O., Oostendorp, R. A., van der Stelt, M. E., Scheper, R. J., & Van Dijk, W. (1990). Con A-nonreactive human alpha 1-acid glycoprotein (AGP) is more effective in modulation of lymphocyte proliferation than Con A-reactive AGP serum variants. *Inflammation*, *14*(2), 133–141.

Potempa, L. A., Rajab, I. M., Hart, P. C., Bordon, J., & Fernandez-Botran, R. (2020). Insights into the Use of C-Reactive Protein as a Diagnostic Index of Disease Severity in COVID-19 Infections. *The American Journal of Tropical Medicine and Hygiene*, *103*(2), 561–563.

Pozhitkov, A. E., Tautz, D., & Noble, P. A. (2007). Oligonucleotide microarrays: widely applied—poorly understood. *Briefings in Functional Genomics*, *6*(2), 141–148.

Preynat-Seauve, O., Coudurier, S., Favier, A., Marche, P.-N., & Villiers, C. (2003). Oxidative stress impairs intracellular events involved in antigen processing and presentation to T cells. *Cell Stress & Chaperones*, *8*(2), 162–171.

Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T. L., Smith, S. R., Bamforth, F., … Wishart, D. S. (2011). The Human Serum Metabolome. *PloS One*, *6*(2), e16957.

Putnam, F. W. (1975). *The plasma proteins, Vol. I, 2nd Edit*. New York, San Francisco, London: Acad. Press.

Putri, S. P., Nakayama, Y., Matsuda, F., Uchikata, T., Kobayashi, S., Matsubara, A., & Fukusaki, E. (2013). Current metabolomics: practical applications. *Journal of Bioscience and Bioengineering*, *115*(6), 579–589.

Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., & Dekkers, J. C. M. (2018a). Novel Resilience Phenotypes from a Natural Disease Challenge Model for Wean-to-Finish Pigs. *Iowa State University Animal Industry Report*, *15*(1).

Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., & PigGen Canada. (2018b). Novel Resilience Phenotypes Using Feed Intake Data From a

Natural Disease Challenge Model in Wean-to-Finish Pigs. *Frontiers in Genetics*, *9*, 660.

Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., & PigGen Canada. (2018). Novel Resilience Phenotypes Using Feed Intake Data From a Natural Disease Challenge Model in Wean-to-Finish Pigs. *Frontiers in Genetics*, *9*, 660.

Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., & PigGen Canada. (2018). Novel Resilience Phenotypes Using Feed Intake Data From a Natural Disease Challenge Model in Wean-to-Finish Pigs. *Frontiers in Genetics*, *9*, 660.

Qi Y.-X., Liu Y.-B., & Rong W.-H. (2011). RNA-Seq and its applications: a new technology for transcriptomics. *Yi chuan = Hereditas*, *33*(11), 1191–1202.

Råberg, L., Sim, D., & Read, A. F. (2007). Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science*, *318*(5851), 812–814.

Ramilo, O., Allman, W., Chung, W., Mejias, A., Ardura, M., Glaser, C., Wittkowski, K. M., Piqueras, B., Banchereau, J., Palucka, A. K., & Chaussabel, D. (2007). Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood*, *109*(5), 2066–2077.

Ramjeet, M., Deslandes, V., Gouré, J., & Jacques, M. (2008). *Actinobacillus pleuropneumoniae* vaccines: from bacterins to new insights into vaccination strategies. *Animal Health Research Reviews / Conference of Research Workers in Animal Diseases*, *9*(1), 25–45.

Rausher, M. D. (2001). Co-evolution and plant resistance to natural enemies. *Nature*, *411*(6839), 857–864.

Ravanbakhsh, S., Liu, P., Bjorndahl, T. C., Mandal, R., Grant, J. R., Wilson, M., Eisner, R., Sinelnikov, I., Hu, X., Luchinat, C., Greiner, R., & Wishart, D. S. (2015). Accurate, fully-automated NMR spectral profiling for metabolomics. *PloS One*, *10*(5), e0124219.

Rawat, K. D., Chahar, M., Reddy, P. V. J., Gupta, P., Shrivastava, N., Gupta, U. D., Natrajan, M., Katoch, V. M., Katoch, K., & Chauhan, D. S. (2013). Expression of CXCL10 (IP-10) and CXCL11 (I-TAC) chemokines during Mycobacterium tuberculosis infection and immunoprophylaxis with Mycobacterium indicus pranii (Mw) in guinea pig. *Infection,*

*Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *13*, 11–17.

Rawat, K. D., Chahar, M., Reddy, P. V. J., Gupta, P., Shrivastava, N., Gupta, U. D., Natrajan, M., Katoch, V. M., Katoch, K., & Chauhan, D. S. (2013). Expression of CXCL10 (IP-10) and CXCL11 (I-TAC) chemokines during Mycobacterium tuberculosis infection and immunoprophylaxis with Mycobacterium indicus pranii (Mw) in guinea pig. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *13*, 11–17.

Reap, E. A., & Lawson, J. W. (1990). Stimulation of the immune response by dimethylglycine, a nontoxic metabolite. *The Journal of laboratory and clinical medicine*, 115(4), 481-486.

Reddy, P. P. (2015). *Climate resilient agriculture for ensuring food security* (Vol. 373). Springer.

Reel, S., Reel, P. S., Erlic, Z., Amar, L., Pecori, A., Larsen, C. K., Tetti, M., Pamporaki, C., Prehn, C., Adamski, J., Prejbisz, A., Ceccato, F., Scaroni, C., Kroiss, M., Dennedy, M. C., Deinum, J., Eisenhofer, G., Langton, K., Mulatero, P., … Jefferson, E. R. (2022). Predicting Hypertension Subtypes with Machine Learning Using Targeted Metabolites and Their Ratios. *Metabolites*, *12*(8). https://doi.org/10.3390/metabo12080755

Ren, J., Yan, X., Ai, H., Zhang, Z., Huang, X., Ouyang, J., Yang, M., Yang, H., Han, P., Zeng, W., Chen, Y., Guo, Y., Xiao, S., Ding, N., & Huang, L. (2012). Susceptibility towards enterotoxigenic Escherichia coli F4ac diarrhea is governed by the MUC13 gene in pigs. *PloS One*, *7*(9), e44573.

Restif, O., & Koella, J. C. (2004). Concurrent evolution of resistance and tolerance to pathogens. *The American Naturalist*, *164*(4), E90–E102.

Riffkin, G. G., & Dobson, C. (1979). Predicting resistance of sheep to Haemonchus contortus infections. *Veterinary Parasitology*, *5*(4), 365–378.

Ritchie, H., Rosado, P., & Roser, M. (2017). Meat and dairy production. *Our World in Data*.

Riyaz, L., Butt, M. A., Zaman, M., & Ayob, O. (2022). Heart Disease Prediction Using Machine Learning Techniques: A Quantitative Review. *International Conference on Innovative*

*Computing and Communications*, 81–94.

Robinson, F. R. (1966). *Pathogenesis of a Mycoplasma Infection in Swine*. Texas A & M University.

Rodríguez-Barrueco, R., Latorre, J., Devis-Jáuregui, L., Lluch, A., Bonifaci, N., Llobet, F. J., Olivan, M., Coll-Iglesias, L., Gassner, K., Davis, M. L., Moreno-Navarrete, J. M., Castells-Nobau, A., Plata-Peña, L., Dalmau-Pastor, M., Höring, M., Liebisch, G., Olkkonen, V. M., Arnoriaga-Rodríguez, M., Ricart, W., … Llobet-Navas, D. (2022). A microRNA cluster controls fat cell differentiation and adipose tissue expansion by regulating SNCG. *Advancement of Science*, *9*(4), e2104759.

Rohart, F., Paris, A., Laurent, B., Canlet, C., Molina, J., Mercat, M. J., Tribout, T., Muller, N., Iannuccelli, N., Villa-Vialaneix, N., Liaubet, L., Milan, D., & San Cristobal, M. (2012). Phenotypic prediction based on metabolomic data for growing pigs from three main European breeds. *Journal of Animal Science*, *90*(13), 4729–4740.

Rose-John, S., Winthrop, K., & Calabrese, L. (2017). The role of IL-6 in host defence against infections: immunobiology and clinical implications. *Nature Reviews. Rheumatology*, *13*(7), 399–409.

Rotty, J. D., Brighton, H. E., Craig, S. L., Asokan, S. B., Cheng, N., Ting, J. P., & Bear, J. E. (2017). Arp2/3 complex is required for macrophage integrin functions but is dispensable for FcR phagocytosis and in vivo motility. *Developmental cell*, 42(5), 498-513.

Rowland, R. R. R., Lunney, J., & Dekkers, J. (2012). Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. *Frontiers in Genetics*, *3*, 260.

Roy, B. A., & Kirchner, J. W. (2000). Evolutionary dynamics of pathogen resistance and tolerance. *Evolution; International Journal of Organic Evolution*, *54*(1), 51–63.

Saini, P. K., & Webert, D. W. (1991). Application of acute phase reactants during antemortem and postmortem meat inspection. *Journal of the American Veterinary Medical Association*, *198*(11), 1898–1901.

Saleem, F., Ametaj, B. N., Bouatra, S., Mandal, R., Zebeli, Q., Dunn, S. M., & Wishart, D. S. (2012). A metabolomics approach to uncover the effects of grain diets on rumen health in dairy cows. *Journal of Dairy Science*, *95*(11), 6606–6623.

Samorè, A. B., & Fontanesi, L. (2016). Genomic selection in pigs: state of the art and perspectives. *Italian Journal of Animal Science*, 15(2), 211-232.

Sanford, S. E., & Josephson, G. K. (1981). Porcine Haemophilus pleuropneumonia epizootic in southwestern Ontario: clinical, microbiological, pathological and some epidemiological findings. *Canadian Journal of Comparative Medicine* , *45*(1), 2–7.

Sarathkumara, Y. D., Browne, D. J., Kelly, A. M., Pattinson, D. J., Rush, C. M., Warner, J., Proietti, C., & Doolan, D. L. (2022). The Effect of Tropical Temperatures on the Quality of RNA Extracted from Stabilized Whole-Blood Samples. *International Journal of Molecular Sciences*, *23*(18), 10609.

Sassu, E. L., Bossé, J. T., Tobias, T. J., Gottschalk, M., Langford, P. R., & Hennig-Pauka, I. (2018). Update on *Actinobacillus pleuropneumoniae*-knowledge, gaps and challenges. *Transboundary and Emerging Diseases*, *65 Suppl 1*, 72–90.

Saude, E. J., Slupsky, C. M., & Sykes, B. D. (2006). Optimization of NMR analysis of biological fluids for quantitative accuracy. *Metabolomics: Official Journal of the Metabolomic Society*, *2*(3), 113–123.

Saunders, A. E., & Johnson, P. (2010). Modulation of immune cell signalling by the leukocyte common tyrosine phosphatase, CD45. *Cellular Signalling*, *22*(3), 339–348.

Savard, C., Alvarez, F., Provost, C., Chorfi, Y., D'Allaire, S., Benoit-Biancamano, M.-O., & Gagnon, C. A. (2016). Efficacy of Fostera PRRS modified live virus vaccine against a Canadian heterologous virulent field strain of porcine reproductive and respiratory syndrome virus. *Canadian Journal of Veterinary Research*, *80*(1), 1–11.

Schena, M., Shalon, D., Davis, R. W., & Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, *270*(5235), 467–470.

Schlüter, T., Schelmbauer, C., Karram, K., & Mufazalov, I. A. (2018). Regulation of IL-1

signaling by the decoy receptor IL-1R2. *Journal of Molecular Medicine* , *96*(10), 983–992.

Schmied, J. D., Putz, A., Dekkers, J., Dekkers, M., Dyck, M., Fortin, F., & Others. (2018). Mortality rate and survival of pigs classified by immune response phenotype using the high immune response (HIR$^{TM}$) technology. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 653.

Schmied, J. D., Putz, A., Dekkers, J., Dekkers, M., Dyck, M., Fortin, F., & Others. (2018). Mortality rate and survival of pigs classified by immune response phenotype using the high immune response (HIR$^{TM}$) technology. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 653.

Schroyen, M., Eisley, C., Koltes, J. E., Fritz-Waters, E., Choi, I., Plastow, G. S., Guan, L., Stothard, P., Bao, H., Kommadath, A., Reecy, J. M., Lunney, J. K., Rowland, R. R. R., Dekkers, J. C. M., & Tuggle, C. K. (2016). Bioinformatic analyses in early host response to Porcine Reproductive and Respiratory Syndrome virus (PRRSV) reveals pathway differences between pigs with alternate genotypes for a major host response QTL. *BMC Genomics*, *17*(1), 196.

Schroyen, M., Liu, H., & Tuggle, C. K. (2016). Applications of Systems Biology to Improve Pig Health. In H. N. Kadarmideen (Ed.), *Systems Biology in Animal Production and Health, Vol. 2* (pp. 33–59). Springer International Publishing.

Schulz, L. L., & Tonsor, G. T. (2015). Assessment of the economic impacts of porcine epidemic diarrhea virus in the United States. *Journal of Animal Science*, *93*(11), 5111–5118.

Schweer, W., Schwartz, K., Patience, J. F., Karriker, L., Sparks, C., Weaver, M., Fitzsimmons, M., Burkey, T. E., & Gabler, N. K. (2017). Porcine Reproductive and Respiratory Syndrome virus reduces feed efficiency, digestibility, and lean tissue accretion in grow-finish pigs. *Translational Animal Science*, *1*(4), 480–488.

Segalés, J. (2012). Porcine circovirus type 2 (PCV2) infections: clinical signs, pathology and laboratory diagnosis. *Virus Research*, *164*(1-2), 10–19.

Sellitto, M., Bai, G., Serena, G., Fricke, W. F., Sturgeon, C., Gajer, P., White, J. R., Koenig, S. S.

K., Sakamoto, J., Boothe, D., Gicquelais, R., Kryszak, D., Puppa, E., Catassi, C., Ravel, J., & Fasano, A. (2012). Proof of concept of microbiome-metabolome analysis and delayed gluten exposure on celiac disease autoimmunity in genetically at-risk infants. *PloS One*, *7*(3), e33387.

Serão, N. V. L., Matika, O., Kemp, R. A., Harding, J. C. S., Bishop, S. C., Plastow, G. S., & Dekkers, J. C. M. (2014). Genetic analysis of reproductive traits and antibody response in a PRRS outbreak herd. *Journal of Animal Science*, *92*(7), 2905–2921.

Sha, Y., Phan, J. H., & Wang, M. D. (2015). Effect of low-expression gene filtering on detection of differentially expressed genes in RNA-seq data. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2015*, 6461–6464.

Shadur, B., Asherie, N., Kfir-Erenfeld, S., Dubnikov, T., NaserEddin, A., Schejter, Y. D., Elpeleg, O., Mor-Shaked, H., & Stepensky, P. (2021). A human case of GIMAP6 deficiency: a novel primary immune deficiency. *European Journal of Human Genetics: EJHG*, *29*(4), 657–662.

Shahzad, K., Lopreiato, V., Liang, Y., Trevisi, E., Osorio, J. S., Xu, C., & Loor, J. J. (2019). Hepatic metabolomics and transcriptomics to study susceptibility to ketosis in response to prepartal nutritional management. *Journal of Animal Science and Biotechnology*, *10*, 96.

Shi, H., Paolucci, U., Vigneau-Callahan, K. E., Milbury, P. E., Matson, W. R., & Kristal, B. S. (2004). Development of biomarkers based on diet-dependent metabolic serotypes: practical issues in development of expert system-based classification models in metabolomic studies. *Omics: A Journal of Integrative Biology*, *8*(3), 197–208.

Shimizu, M., Yamada, S., Kawashima, K., Ohashi, S., Shimizu, S., & Ogawa, T. (1996). Changes of lymphocyte subpopulations in pigs infected with porcine reproductive and respiratory syndrome (PRRS) virus. *Veterinary Immunology and Immunopathology*, *50*(1-2), 19–27.

Sidibé, M., Messier, S., Larivière, S., Gottschalk, M., & Mittal, K. R. (1993). Detection of *Actinobacillus pleuropneumoniae* in the porcine upper respiratory tract as a complement to serological tests. *Canadian Journal of Veterinary Research*, *57*(3), 204–208.

Silva, T., Temerozo, J. R., do Vale, G., Ferreira, A. C., Soares, V. C., Dias, S. S. G., Sardella, G., Bou-Habib, D. C., Siqueira, M., Souza, T. M. L., & Miranda, M. (2021). The Chemokine CCL5 Inhibits the Replication of Influenza A Virus Through SAMHD1 Modulation. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.549020

Sim, G. K., Kafatos, F. C., Jones, C. W., Koehler, M. D., Efstratiadis, A., & Maniatis, T. (1979). Use of a cDNA library for studies on evolution and developmental expression of the chorion multigene families. *Cell*, *18*(4), 1303–1316.

Sirinarumitr, T., Zhang, Y., Kluge, J. P., Halbur, P. G., & Paul, P. S. (1998). A pneumo-virulent United States isolate of porcine reproductive and respiratory syndrome virus induces apoptosis in bystander cells both in vitro and in vivo. *The Journal of General Virology*, *79 ( Pt 12)*, 2989–2995.

Skiniotis, G., Boulanger, M. J., Garcia, K. C., & Walz, T. (2005). Signaling conformations of the tall cytokine receptor gp130 when in complex with IL-6 and IL-6 receptor. *Nature Structural & Molecular Biology*, *12*(6), 545–551.

Skovgaard, K., Mortensen, S., Boye, M., Poulsen, K. T., Campbell, F. M., Eckersall, P. D., & Heegaard, P. M. H. (2009). Rapid and widely disseminated acute phase protein response after experimental bacterial infection of pigs. *Veterinary Research*, *40*(3), 23.

Smilde, A. K., Jansen, J. J., Hoefsloot, H. C. J., Lamers, R.-J. A. N., van der Greef, J., & Timmerman, M. E. (2005). ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics* , *21*(13), 3043–3048.

Smith, C. A., O'Maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., Custodio, D. E., Abagyan, R., & Siuzdak, G. (2005). METLIN: a metabolite mass spectral database. *Therapeutic Drug Monitoring*, *27*(6), 747–751.

Soga, T., Sugimoto, M., Honma, M., Mori, M., Igarashi, K., Kashikura, K., Ikeda, S., Hirayama, A., Yamamoto, T., Yoshida, H., Otsuka, M., Tsuji, S., Yatomi, Y., Sakuragawa, T., Watanabe, H., Nihei, K., Saito, T., Kawata, S., Suzuki, H., … Suematsu, M. (2011). Serum

metabolomics reveals γ-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. *Journal of Hepatology*, *55*(4), 896–905.

Song, D., & Park, B. (2012). Porcine epidemic diarrhoea virus: a comprehensive review of molecular epidemiology, diagnosis, and vaccines. *Virus Genes*, *44*(2), 167–175.

Sorensen, N. S., Tegtmeier, C., Andresen, L. O., Piñeiro, M., Toussaint, M. J. M., Campbell, F. M., Lampreave, F., & Heegaard, P. M. H. (2006). The porcine acute phase protein response to acute clinical and subclinical experimental infection with Streptococcus suis. *Veterinary Immunology and Immunopathology*, *113*(1-2), 157–168.

Spek, C. A., Aberson, H. L., Butler, J. M., de Vos, A. F., & Duitman, J. (2021). CEBPD Potentiates the Macrophage Inflammatory Response but CEBPD Knock-Out Macrophages Fail to Identify CEBPD-Dependent Pro-Inflammatory Transcriptional Programs. *Cells* , *10*(9). https://doi.org/10.3390/cells10092233

Sproston, N. R., & Ashworth, J. J. (2018). Role of C-Reactive Protein at Sites of Inflammation and Infection. *Frontiers in Immunology*, *9*, 754.

Steel, D. M., & Whitehead, A. S. (1994). The major acute phase reactants: C-reactive protein, serum amyloid P component and serum amyloid A protein. *Immunology Today*, *15*(2), 81–88.

Straub, R. H. (2017). The brain and immune system prompt energy shortage in chronic inflammation and ageing. *Nature Reviews. Rheumatology*, *13*(12), 743–751.

Stricker, R. L. O., Behrens, S.-E., & Mundt, E. (2010). Nuclear factor NF45 interacts with viral proteins of infectious bursal disease virus and inhibits viral replication. *Journal of Virology*, *84*(20), 10592–10605.

Suankratay, C., Mold, C., Zhang, Y., Potempa, L. A., Lint, T. F., & Gewurz, H. (1998). Complement regulation in innate immunity and the acute-phase response: inhibition of mannan-binding lectin-initiated complement cytolysis by C-reactive protein (CRP). *Clinical and Experimental Immunology*, *113*(3), 353–359.

Suhre, K., Meisinger, C., Döring, A., Altmaier, E., Belcredi, P., Gieger, C., Chang, D., Milburn,

M. V., Gall, W. E., Weinberger, K. M., Mewes, H.-W., Hrabé de Angelis, M., Wichmann, H.-E., Kronenberg, F., Adamski, J., & Illig, T. (2010). Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *PloS One*, *5*(11), e13953.

Sumanth, M. S., Abhilasha, K. V., Jacob, S. P., Chaithra, V. H., Basrur, V., Willard, B., ... & Marathe, G. K. (2019). Acute phase protein, α–1-acid glycoprotein (AGP-1), has differential effects on TLR-2 and TLR-4 mediated responses. *Immunobiology*, 224(5), 672-680.

Sun, Y., Yang, Y., Qin, Z., Cai, J., Guo, X., Tang, Y., Wan, J., Su, D.-F., & Liu, X. (2016). The Acute-Phase Protein Orosomucoid Regulates Food Intake and Energy Homeostasis via Leptin Receptor Signaling Pathway. *Diabetes*, *65*(6), 1630–1641.

Suradhat, S., & Thanawongnuwech, R. (2003). Upregulation of interleukin-10 gene expression in the leukocytes of pigs infected with porcine reproductive and respiratory syndrome virus. *The Journal of General Virology*, *84*(Pt 10), 2755–2760.

Suravajhala, P., Kogelman, L. J. A., & Kadarmideen, H. N. (2016). Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. *Genetics, Selection, Evolution: GSE*, *48*(1), 38.

Sutton, V. R., Brennan, A. J., Ellis, S., Danne, J., Thia, K., Jenkins, M. R., Voskoboinik, I., Pejler, G., Johnstone, R. W., Andrews, D. M., & Trapani, J. A. (2016). Serglycin determines secretory granule repertoire and regulates natural killer cell and cytotoxic T lymphocyte cytotoxicity. *The FEBS Journal*, *283*(5), 947–961.

Swain-Lenz, D., Nikolskiy, I., Cheng, J., Sudarsanam, P., Nayler, D., Staller, M. V., & Cohen, B. A. (2017). Causal Genetic Variation Underlying Metabolome Differences. *Genetics*, *206*(4), 2199–2206.

Taguchi, K. (2013). *Molecular Aspects of Human Alpha-1 Acid Glycoprotein - Structure and Function*. IntechOpen.

Takeuchi, A., Miyamoto, T., Yamaji, K., Masuho, Y., Hayashi, M., Hayashi, H., & Onozaki, K. (1995). A human erythrocyte-derived growth-promoting factor with a wide target cell spectrum: identification as catalase. *Cancer Research*, *55*(7), 1586–1589.

Tanaka, T., Narazaki, M., & Kishimoto, T. (2016). Immunotherapeutic implications of IL-6 blockade for cytokine storm. *Immunotherapy*, *8*(8), 959–970.

Taylor, J. J., Jenkins, M. K., & Pape, K. A. (2012). Heterogeneity in the differentiation and function of memory B cells. *Trends in Immunology*, *33*(12), 590–597.

Tedder, T. F., Boyd, A. W., Freedman, A. S., Nadler, L. M., & Schlossman, S. F. (1985). The B cell surface molecule B1 is functionally linked with B cell activation and differentiation. *Journal of Immunology* , *135*(2), 973–979.

Teijaro, J. R., Walsh, K. B., Rice, S., Rosen, H., & Oldstone, M. B. A. (2014). Mapping the innate signaling cascade essential for cytokine storm during influenza virus infection. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(10), 3799–3804.

Tekkanat, K. K., Maassab, H. F., Cho, D. S., Lai, J. J., John, A., Berlin, A., Kaplan, M. H., & Lukacs, N. W. (2001). IL-13-induced airway hyperreactivity during respiratory syncytial virus infection is STAT6 dependent. *Journal of Immunology* , *166*(5), 3542–3548.

Tellier, R. (2009). Aerosol transmission of influenza A virus: a review of new studies. *Journal of the Royal Society, Interface / the Royal Society*, *6 Suppl 6*(Suppl 6), S783–S790.

Teul, J., Rupérez, F. J., Garcia, A., Vaysse, J., Balayssac, S., Gilard, V., Malet-Martino, M., Martin-Ventura, J. L., Blanco-Colio, L. M., Tuñón, J., Egido, J., & Barbas, C. (2009). Improving metabolite knowledge in stable atherosclerosis patients by association and correlation of GC-MS and 1H NMR fingerprints. *Journal of Proteome Research*, *8*(12), 5580–5589.

Thacker, E. L., Thacker, B. J., & Janke, B. H. (2001). Interaction between Mycoplasma hyopneumoniae and swine influenza virus. *Journal of Clinical Microbiology*, *39*(7), 2525–2530.

The British Columbia Society for the Prevention of Cruelty to Animals (BC SPCA) 2023, Pig farming in Canada, accessed 29 Oct 2023, https://spca.bc.ca/programs-services/farm-animal-programs/farm-animal-production/pigs/.

Ticó, G., Segalés, J., & Martínez, J. (2013). The blurred border between porcine circovirus type 2-systemic disease and porcine respiratory disease complex. *Veterinary Microbiology*, *163*(3-4), 242–247.

Tiziani, S., Lopes, V., & Günther, U. L. (2009). Early Stage Diagnosis of Oral Cancer Using 1H NMR–Based Metabolomics. *Neoplasia* , *11*(3), 269–IN10.

Tobias, T. J. (2014). *Actinobacillus pleuropneumoniae transmission and clinical outbreaks*. Utrecht University.

Tobias, T. J., Raymakers, R. J. M. L., van Nes, A., & van Leengoed, L. A. M. G. (2009). Outbreak of respiratory distress resembling influenza caused by *Actinobacillus pleuropneumoniae* in pigs. *The Veterinary Record*, *164*(13), 402–403.

Tourlomoussis, P., Eckersall, P. D., Waterson, M. M., & Buncic, S. (2004). A comparison of acute phase protein measurements and meat inspection findings in cattle. *Foodborne Pathogens and Disease*, *1*(4), 281–290.

Tseng, G. C., Ghosh, D., & Feingold, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Research*, *40*(9), 3785–3799.

Turner, P. V., Brabb, T., Pekow, C., & Vasbinder, M. A. (2011). Administration of substances to laboratory animals: routes of administration and factors to consider. *Journal of the American Association for Laboratory Animal Science: JAALAS*, *50*(5), 600–613.

Tyner, J. W., Uchida, O., Kajiwara, N., Kim, E. Y., O'Sullivan, M. P., Walter, M. J., Cook, D. N., Danoff, T. M., & Holtzman, M. J. (2004). Chemokine CCL5 signals required for survival during viral infection. *The Journal of Allergy and Clinical Immunology*, *113*(2, Supplement), S49.

Tyner, J. W., Uchida, O., Kajiwara, N., Kim, E. Y., Patel, A. C., O'Sullivan, M. P., Walter, M. J., Schwendener, R. A., Cook, D. N., Danoff, T. M., & Holtzman, M. J. (2005). CCL5-CCR5 interaction provides antiapoptotic signals for macrophage survival during viral infection. *Nature Medicine*, *11*(11), 1180–1187.

Ursell, L. K., Haiser, H. J., Van Treuren, W., Garg, N., Reddivari, L., Vanamala, J., Dorrestein, P.

C., Turnbaugh, P. J., & Knight, R. (2014). The intestinal metabolome: an intersection between microbiota and host. *Gastroenterology*, *146*(6), 1470–1476.

van Baren, N., & Van den Eynde, B. J. (2015). Tryptophan-degrading enzymes in tumoral immune resistance. *Frontiers in Immunology*, *6*, 34.

van den Berg, A., Danuser, J., Frey, J., & Regula, G. (2007). Evaluation of the acute phase protein haptoglobin as an indicator of herd health in slaughter pigs. *Animal Welfare* , *16*(2), 157–159.

Van Reeth, K., Labarque, G., Nauwynck, H., & Pensaert, M. (1999). Differential production of proinflammatory cytokines in the pig lung during different respiratory virus infections: correlations with pathogenicity. *Research in Veterinary Science*, *67*(1), 47–52.

van Velzen, E. J. J., Westerhuis, J. A., van Duynhoven, J. P. M., van Dorsten, F. A., Hoefsloot, H. C. J., Jacobs, D. M., Smit, S., Draijer, R., Kroner, C. I., & Smilde, A. K. (2008). Multilevel data analysis of a crossover designed human nutritional intervention study. *Journal of Proteome Research*, *7*(10), 4483–4491.

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Jr, Hieter, P., Vogelstein, B., & Kinzler, K. W. (1997). Characterization of the yeast transcriptome. *Cell*, *88*(2), 243–251.

Verhoeckx, K. C. M., Bijlsma, S., de Groene, E. M., Witkamp, R. F., van der Greef, J., & Rodenburg, R. J. T. (2004). A combination of proteomics, principal component analysis and transcriptomics is a powerful tool for the identification of biomarkers for macrophage maturation in the U937 cell line. *Proteomics*, *4*(4), 1014–1028.

Vinayavekhin, N., & Saghatelian, A. (2010). Untargeted metabolomics. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, *Chapter 30*, Unit 30.1.1–24.

Vincent, A. L., Ma, W., Lager, K. M., Janke, B. H., & Richt, J. A. (2008). Swine influenza viruses a North American perspective. *Advances in Virus Research*, *72*, 127–154.

Vu, H. L. X., Kwon, B., Yoon, K.-J., Laegreid, W. W., Pattnaik, A. K., & Osorio, F. A. (2011). Immune evasion of porcine reproductive and respiratory syndrome virus through glycan

shielding involves both glycoprotein 5 as well as glycoprotein 3. *Journal of Virology*, *85*(11), 5555–5564.

Wake Forest University Baptist Medical Center. (2008, January 4). *Hydrogen Peroxide Has A Complex Role In Cell Health*. ScienceDaily. www.sciencedaily.com/releases/2008/01/080102134129.html

Walford, H. H., & Doherty, T. A. (2013). STAT6 and lung inflammation. *JAK-STAT*, *2*(4), e25301.

Walker, L. R., Engle, T. B., Vu, H., Tosky, E. R., Nonneman, D. J., Smith, T. P. L., Borza, T., Burkey, T. E., Plastow, G. S., Kachman, S. D., & Ciobanu, D. C. (2018). Synaptogyrin-2 influences replication of Porcine circovirus 2. *PLoS Genetics*, *14*(10), e1007750.

Wang, T., Wang, X., Peng, L., Cong, Y., & Du, C. (n.d.). *Prediction of key genes and metabolic biomarkers in ankylosing spondylitis based on a multi-omics composite network*. Retrieved August 12, 2023, from https://e-century.us/files/ijcem/11/8/ijcem0067893.pdf

Wang, Y., Qu, L., Uthe, J. J., Bearson, S. M. D., Kuhar, D., Lunney, J. K., Couture, O. P., Nettleton, D., Dekkers, J. C. M., & Tuggle, C. K. (2007). Global transcriptional response of porcine mesenteric lymph nodes to Salmonella enterica serovar Typhimurium. *Genomics*, *90*(1), 72–84.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*, *10*(1), 57–63.

Warner, C. M., Meeker, D. L., & Rothschild, M. F. (1987). Genetic control of immune responsiveness: a review of its use as a tool for selection for disease resistance. *Journal of Animal Science*, *64*(2), 394–406.

Warth, B., Raffeiner, P., Granados, A., Huan, T., Fang, M., Forsberg, E. M., Benton, H. P., Goetz, L., Johnson, C. H., & Siuzdak, G. (2018). Metabolomics Reveals that Dietary Xenoestrogens Alter Cellular Metabolism Induced by Palbociclib/Letrozole Combination Cancer Therapy. *Cell Chemical Biology*, *25*(3), 291–300.e3.

Watanabe, A., Yamada, Y., & Yamanaka, S. (2013). Epigenetic regulation in pluripotent stem

cells: a key to breaking the epigenetic barrier. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1609), 20120292.

Welch, M. D., DePace, A. H., Verma, S., Iwamatsu, A., & Mitchison, T. J. (1997). The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly. *The Journal of cell biology*, 138(2), 375-384.

Wen, X., Bian, T., Zhang, Z., Zhou, L., Ge, X., Han, J., Guo, X., Yang, H., & Yu, K. (2017). Interleukin-2 enhancer binding factor 2 interacts with the nsp9 or nsp2 of porcine reproductive and respiratory syndrome virus and exerts negatively regulatory effect on the viral replication. *Virology Journal*, *14*(1), 125.

Wensvoort, G., Terpstra, C., Pol, J. M., ter Laak, E. A., Bloemraad, M., de Kluyver, E. P., Kragten, C., van Buiten, L., den Besten, A., & Wagenaar, F. (1991). Mystery swine disease in The Netherlands: the isolation of Lelystad virus. *The Veterinary Quarterly*, *13*(3), 121–130.

Westermann, A. J., Förstner, K. U., Amman, F., Barquist, L., Chao, Y., Schulte, L. N., Müller, L., Reinhardt, R., Stadler, P. F., & Vogel, J. (2016). Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions. *Nature*, *529*(7587), 496–501.

White, M. (2005). *Pig ailments: recognition and treatment*. Crowood Press.

Whitworth, K. M., Rowland, R. R. R., Ewen, C. L., Trible, B. R., Kerrigan, M. A., Cino-Ozuna, A. G., Samuel, M. S., Lightner, J. E., McLaren, D. G., Mileham, A. J., Wells, K. D., & Prather, R. S. (2016). Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. *Nature Biotechnology*, *34*(1), 20–22.

Whitworth, K. M., Rowland, R. R. R., Ewen, C. L., Trible, B. R., Kerrigan, M. A., Cino-Ozuna, A. G., Samuel, M. S., Lightner, J. E., McLaren, D. G., Mileham, A. J., Wells, K. D., & Prather, R. S. (2016). Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. *Nature Biotechnology*, *34*(1), 20–22.

Whitworth, K. M., Rowland, R. R. R., Petrovan, V., Sheahan, M., Cino-Ozuna, A. G., Fang, Y.,

Hesse, R., Mileham, A., Samuel, M. S., Wells, K. D., & Prather, R. S. (2019). Resistance to coronavirus infection in amino peptidase N-deficient pigs. *Transgenic Research*, *28*(1), 21–32.

Whitworth, K. M., Rowland, R. R. R., Petrovan, V., Sheahan, M., Cino-Ozuna, A. G., Fang, Y., Hesse, R., Mileham, A., Samuel, M. S., Wells, K. D., & Prather, R. S. (2019). Resistance to coronavirus infection in amino peptidase N-deficient pigs. *Transgenic Research*, *28*(1), 21–32.

Wickramarachchi, D. C., Theofilopoulos, A. N., & Kono, D. H. (2010). Immune pathology associated with altered actin cytoskeleton regulation. *Autoimmunity*, 43(1), 64-75.

Wikoff, W. R., Anfora, A. T., Liu, J., Schultz, P. G., Lesley, S. A., Peters, E. C., & Siuzdak, G. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(10), 3698–3703.

Wilkie, B., & Mallard, B. (1999). Selection for high immune response: an alternative approach to animal health maintenance? *Veterinary Immunology and Immunopathology*, *72*(1-2), 231–235.

Wilkinson, J. M., Ladinig, A., Bao, H., Kommadath, A., Stothard, P., Lunney, J. K., Harding, J. C. S., & Plastow, G. S. (2016). Differences in whole blood gene expression associated with infection time-course and extent of fetal mortality in a reproductive model of type 2 porcine reproductive and Respiratory Syndrome Virus (PRRSV) infection. *PloS One*, *11*(4), e0153615.

Wilkinson, J. M., Sargent, C. A., Galina-Pantoja, L., & Tucker, A. W. (2010). Gene expression profiling in the lungs of pigs with different susceptibilities to Glässer's disease. *BMC Genomics*, *11*, 455.

Wilson, T. J., Clare, S., Mikulin, J., Johnson, C. M., Harcourt, K., Lyons, P. A., Dougan, G., & Smith, K. G. C. (2020). Signalling lymphocyte activation molecule family member 9 is found on select subsets of antigen-presenting cells and promotes resistance to Salmonella

infection. *Immunology*, *159*(4), 393–403.

Winkel, V. M., Pavan, T. L. R., Wirthl, V. A. B. F., Alves, A. L. N., & Lucas, S. R. R. (2015). Serum α-1 acid glycoprotein and serum amyloid A concentrations in cats receiving antineoplastic treatment for lymphoma. *American Journal of Veterinary Research*, *76*(11), 983–988.

Winkler, K., Herbig, J., & Kohl, I. (2013). Real-time metabolic monitoring with proton transfer reaction mass spectrometry. *Journal of Breath Research*, *7*(3), 036006.

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M.-A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., … Querengesser, L. (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Research*, *35*(Database issue), D521–D526.

Woillard, J.-B., Labriffe, M., Debord, J., & Marquet, P. (2021). Mycophenolic Acid Exposure Prediction Using Machine Learning. *Clinical Pharmacology and Therapeutics*, *110*(2), 370–379.

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*(4), 569–577.

Wu, J., Zeng, Y. X., & Hirokawa, K. (1999). Signal pathway of mitogen-induced Ca2+-activated K+ currents in young and aged T-cell clones of C57BL/6 mice. *Cellular Signalling*, *11*(6), 391–398.

Wu, Y., Potempa, L. A., El Kebir, D., & Filep, J. G. (2015). C-reactive protein and inflammation: conformational changes affect function. *Biological Chemistry*, *396*(11), 1181–1197.

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics: Official Journal of the Metabolomic Society*, *9*(2), 280–299.

Xiao, S., Jia, J., Mo, D., Wang, Q., Qin, L., He, Z., Zhao, X., Huang, Y., Li, A., Yu, J., Niu, Y., Liu, X., & Chen, Y. (2010). Understanding PRRSV infection in porcine lung based on

genome-wide transcriptome response identified by deep sequencing. *PloS One*, *5*(6), e11377.

Xu, Y., Ritchie, S. C., Liang, Y., Timmers, P. R. H. J., Pietzner, M., Lannelongue, L., Lambert, S. A., Tahir, U. A., May-Wilson, S., Foguet, C., Johansson, Å., Surendran, P., Nath, A. P., Persyn, E., Peters, J. E., Oliver-Williams, C., Deng, S., Prins, B., Luan, J. 'an, … Inouye, M. (2023). An atlas of genetic scores to predict multi-omic traits. *Nature*, *616*(7955), 123–131.

Yang, Q., Vijayakumar, A., & Kahn, B. B. (2018). Metabolites as regulators of insulin sensitivity and metabolism. *Nature Reviews. Molecular Cell Biology*, *19*(10), 654–672.

Yao, Y., Du Jiang, P., Chao, B. N., Cagdas, D., Kubo, S., Balasubramaniyam, A., Zhang, Y., Shadur, B., NaserEddin, A., Folio, L. R., Schwarz, B., Bohrnsen, E., Zheng, L., Lynberg, M., Gottlieb, S., Leney-Greene, M. A., Park, A. Y., Tezcan, I., Akdogan, A., … Lenardo, M. J. (2022). GIMAP6 regulates autophagy, immune competence, and inflammation in mice and humans. *The Journal of Experimental Medicine*, *219*(6). https://doi.org/10.1084/jem.20201405

Ye, Y., Tang, X., Sun, Z., & Chen, S. (2016). Upregulated WDR26 serves as a scaffold to coordinate PI3K/ AKT pathway-driven breast cancer cell growth, migration, and invasion. *Oncotarget*, *7*(14), 17854–17869.

Yoshizaki, K. (2011). Pathogenic role of IL-6 combined with TNF-α or IL-1 in the induction of acute phase proteins SAA and CRP in chronic inflammatory diseases. *Advances in Experimental Medicine and Biology*, *691*, 141–150.

Zhang, C., Wu, Z., Li, J.-W., Zhao, H., & Wang, G.-Q. (2020). Cytokine release syndrome in severe COVID-19: interleukin-6 receptor antagonist tocilizumab may be the key to reduce mortality. *International Journal of Antimicrobial Agents*, *55*(5), 105954.

Zhang, G., Dervishi, E., Dunn, S. M., Mandal, R., Liu, P., Han, B., ... & Ametaj, B. N. (2017). Metabotyping reveals distinct metabolic alterations in ketotic cows and identifies early predictive serum biomarkers for the risk of disease. *Metabolomics,* 13, 1-15.

Zhang, H., Kouadio, A., Cartledge, D., & Godwin, A. K. (2011). Role of gamma-synuclein in microtubule regulation. *Experimental Cell Research*, *317*(10), 1330–1339.

Zhang, H., Lunney, J. K., Baker, R. B., & Opriessnig, T. (2011). Cytokine and chemokine mRNA expression profiles in tracheobronchial lymph nodes from pigs singularly infected or coinfected with porcine circovirus type 2 (PCV2) and Mycoplasma hyopneumoniae (MHYO). *Veterinary Immunology and Immunopathology*, *140*(1), 152–158.

Zhang, L., Zhu, J., Wang, H., Xia, J., Liu, P., Chen, F., ... & Zhang, X. (2021). A high-resolution cell atlas of the domestic pig lung and an online platform for exploring lung single-cell data. *Journal of Genetics and Genomics*, 48(5), 411-425.

Zhao, J., Liu, Y., Wei, X., Yuan, C., Yuan, X., & Xiao, X. (2009). A novel WD-40 repeat protein WDR26 suppresses H2O2-induced cell death in neural cells. *Neuroscience Letters*, *460*(1), 66–71.

Zhao, L., Zhang, H., Kohnen, M. V., Prasad, K. V. S. K., Gu, L., & Reddy, A. S. N. (2019). Analysis of Transcriptome and Epitranscriptome in Plants Using PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing. *Frontiers in Genetics*, *10*, 253.

Zhao, S., Zhu, M., & Chen, H. (2012). Immunogenomics for identification of disease resistance genes in pigs: a review focusing on Gram-negative bacilli. *Journal of Animal Science and Biotechnology*, *3*(1), 34.

Zheng, Y., Li, S., Li, S.-H., Yu, S., Wang, Q., Zhang, K., Qu, L., Sun, Y., Bi, Y., Tang, F., Qiu, H.-J., & Gao, G. F. (2022). Transcriptome profiling in swine macrophages infected with African swine fever virus at single-cell resolution. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(19), e2201288119.

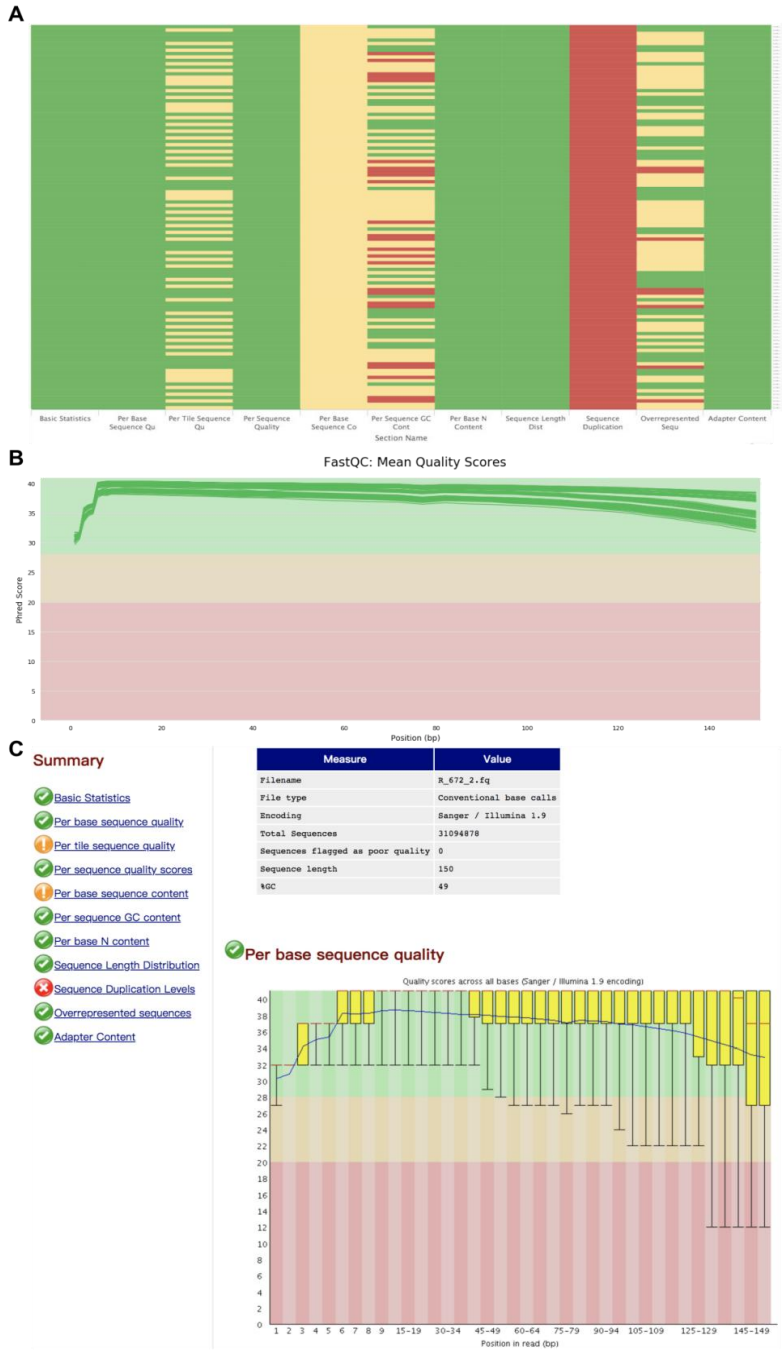
Zhuge, Z.-Y., Zhu, Y.-H., Liu, P.-Q., Yan, X.-D., Yue, Y., Weng, X.-G., Zhang, R., & Wang, J.-F. (2012). Effects of Astragalus polysaccharide on immune responses of porcine PBMC stimulated with PRRSV or CSFV. *PloS One*, *7*(1), e29320.

Zimmerman, J. J., Yoon, K. J., Wills, R. W., & Swenson, S. L. (1997). General overview of

PRRSV: a perspective from the United States. *Veterinary Microbiology*, *55*(1-4), 187–196.

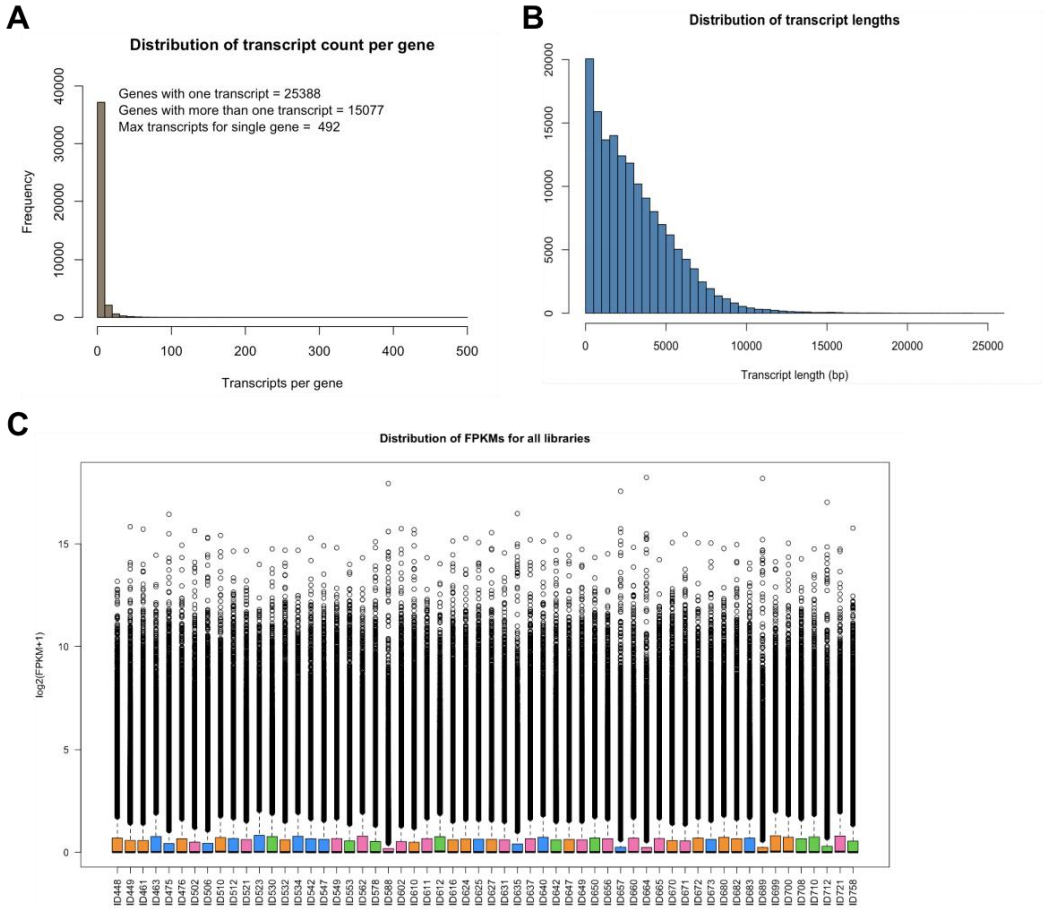
Zuckermann, F. A., Garcia, E. A., Luque, I. D., Christopher-Hennings, J., Doster, A., Brito, M., & Osorio, F. (2007). Assessment of the efficacy of commercial porcine reproductive and respiratory syndrome virus (PRRSV) vaccines based on measurement of serologic response, frequency of gamma-IFN-producing cells and virological parameters of protection upon challenge. *Veterinary Microbiology*, *123*(1-3), 69–85.

# Appendices:




**Appendix 3.1** Quality check of the RNA-Seq reads from the 58 pig whole blood samples.

(A) MultiQC report of aggregated FastQC results with all the indicated check terms from the 558 samples. (B) Mean read quality scores (Phred score) of the 558 samples. (C) Example of the "Per base sequence quality" plot from one sample (pig_672_Read_2)




**Appendix 3.2** Summary of counted reads across all 58 RNA samples from cohort 1 pigs.

**Appendix 3.3** List of names and numbers of the DE genes and transcripts from the transcriptome of 58 pigs (cohort 1) and the 209 pigs (cohort 2).

Top 20 (lowest P-values) genes or transcripts are listed if more than 20 hits were found. Change trends of DE genes or transcripts are indicated as in the first group relative to the second group in each comparison.

| Cohort | Comparisons | Change trend | DE types | Total number | Gene Names |
|---|---|---|---|---|---|
| 1 | Blood1 Resistant vs Others | Up | Genes | 0 | |
| 1 | Blood1 Resistant vs Others | Down | Genes | 1 | *CCDC117* |
| 1 | Blood1 Resilient vs Others | Up | Genes | 1 | *MSTRG.564* |
| 1 | Blood1 Resilient vs Others | Down | Genes | 2 | *MSTRG.22405, CACNG8* |
| 1 | Blood1 Early_dead vs Others | Up | Genes | 1 | *MSTRG.11291* |
| 1 | Blood1 Early_dead vs Others | Down | Genes | 0 | |
| 1 | Blood1 Susceptible vs Others | Up | Genes | 8 | *ZCCHC2, TGM2, CCDC117, THY1, MSTRG.10078, RF01977, RF01978, RF01979* |
| 1 | Blood1 Susceptible vs Others | Down | Genes | 4 | *MSTRG.11828, ADAM19, TAF7, BCL2L1* |
| 1 | Blood1 Resistant vs Others | Up | Transcripts | 100 | *TES, PAFAH2, NDUFS7, NISCH, SLA-5, SPPL2A, RFFL, CCL5, PTK2B, ATOH8* |
| 1 | Blood1 Resistant vs Others | Down | Transcripts | 61 | *MYL9, PFN1, MSTRG.29820, MGAT4B, SMAP2, BCL11B, SRRM2, AXIN1, CCDC117, MSTRG.24962* |
| 1 | Blood1 Resilient vs Others | Up | Transcripts | 115 | *IL4R, MXD1, B4GALT1, NDUFS1, TNRC18, IL6ST, B3GNT3, MSTRG.17360, SLC4A7, MSTRG.18656* |
| 1 | Blood1 Resilient vs | Down | Transcrip | 74 | *UBE2L6, BIRC3, NTAN1, DNAJA3, LYN, PRKAR1A,* |

| | | | | | |
|---|---|---|---|---|---|
| | Others | | ts | | *TMEM106A, MSTRG.27061, PRKACA, PMF1* |
| 1 | Blood1 Early_dead vs Others | Up | Transcripts | 86 | *CLTC, SRRM2, EIF3G, NCOA4, ssc-mir-6782, CKB, SNCG, RNASEL, PSTPIP2, LAPTM5* |
| 1 | Blood1 Early_dead vs Others | Down | Transcripts | 50 | *COPG2, CTNNBL1, SPPL2A, CUL4B, MSTRG.17597, PLEK, RF00139, SNAP23, MSTRG.18000, RERE* |
| 1 | Blood1 Susceptible vs Others | Up | Transcripts | 236 | *MSTRG.32803, MSTRG.21242, CCDC6, UBR4, RNF151, RBM17, ZCCHC2, KLF17, MSTRG.21582, MSTRG.19215* |
| 1 | Blood1 Susceptible vs Others | Down | Transcripts | 83 | *KLF17, MSTRG.3999, SV2A, RNF151, AGK, ATL3, GAPDH, NFKBIZ, ILF2, PTP4A1* |
| 1 | Blood1 Resistant vs Resilient | Up | Genes | 7 | *BRPF3, MSTRG.17793, MSTRG.32328, CACNG8, TGM3, MSTRG.33819, CEBPD* |
| 1 | Blood1 Resistant vs Resilient | Down | Genes | 4 | *MANSC1, CSKMT, CCDC117, MSTRG.8886* |
| 1 | Blood1 Resistant vs Early_dead | Up | Genes | 1 | *ATRN* |
| 1 | Blood1 Resistant vs Early_dead | Down | Genes | 1 | *MSTRG.11291* |
| 1 | Blood1 Resistant vs Susceptible | Up | Genes | 7 | *CEP68, ATRN, IL1R2, GNA13, RGS2, MSTRG.983, MSTRG.11828* |
| 1 | Blood1 Resistant vs Susceptible | Down | Genes | 7 | *CCDC117, BBX, PSTK, THY1, DHCR24, TGM2, PTPRO* |
| 1 | Blood1 Resilient vs Early_dead | Up | Genes | 3 | *M6PR, CXCR5, RF00017* |
| 1 | Blood1 Resilient vs Early_dead | Down | Genes | 2 | *MSTRG.29272, MSTRG.22405* |
| 1 | Blood1 Resilient vs Susceptible | Up | Genes | 17 | *PPP3CA, MSTRG.34160, ZNF572, RF00017, BTG1, KHK, MSTRG.11828, MSTRG.15239, MSTRG.564, MAP3K14* |
| 1 | Blood1 Resilient vs Susceptible | Down | Genes | 11 | *CLTB, MSTRG.33819, CACNG8, MSTRG.29489, MSTRG.33785, MSTRG.31561, AP4E1, MSTRG.29272, ZCCHC2, TGM2* |

| | | | | | |
|---|---|---|---|---|---|
| 1 | Blood1 Early_dead vs Susceptible | Up | Genes | 10 | *RANBP10, INPP5E, MSTRG.9951, FCRL3, MSTRG.21823, MSTRG.11291, ADAM19, SAMD9, TAF7, ZC3HAV1* |
| 1 | Blood1 Early_dead vs Susceptible | Down | Genes | 9 | *ZCCHC2, MTFR1, RF01978, RF01977, RF01979, MSTRG.33819, MSTRG.14994, TGM2, MSTRG.14772* |
| 1 | Blood1 Resistant vs Resilient | Up | Transcripts | 152 | *NR2E1, TES, MGAT4A, CCL5, MBOAT7, NISCH, HECA, PMF1, XPO7, LYN* |
| 1 | Blood1 Resistant vs Resilient | Down | Transcripts | 146 | *MXD1, B4GALT1, BCL11B, B3GNT3, PFN1, LRRC73, XPO1, MSTRG.21158, NDUFS1, SMC4* |
| 1 | Blood1 Resistant vs Early_dead | Up | Transcripts | 125 | *SPPL2A, ZNF444, B4GALT1, TES, SLA-5, ILKAP, MSTRG.18000, RF00571, RFFL, MSTRG.24007* |
| 1 | Blood1 Resistant vs Early_dead | Down | Transcripts | 98 | *SRRM2, MSTRG.21158, MSTRG.25464, MSTRG.24962, SMAP2, MYL9, GSN, RNASEL, ECM1, KIAA1147* |
| 1 | Blood1 Resistant vs Susceptible | Up | Transcripts | 124 | *GAPDH, NISCH, MSTRG.3999, MRPS15, MSTRG.8068, AGK, PCMTD1, USP1, CASP7, ARHGAP4* |
| 1 | Blood1 Resistant vs Susceptible | Down | Transcripts | 222 | *RUFY4, MISP3, CAMTA2, BBX, PSTK, INTS7, AXIN1, LARP4, MSTRG.2870, ANKMY2* |
| 1 | Blood1 Resilient vs Early_dead | Up | Transcripts | 137 | *MXD1, IL4R, SRRM2, ZNF644, CSPG5, CSNK1A1, ZMYM2, ATP2A3, DEGS1, M6PR* |
| 1 | Blood1 Resilient vs Early_dead | Down | Transcripts | 130 | *BIRC3, UBE2L6, CLTC, TCOF1, NTAN1, NIPSNAP3B, DNAJA3, NASP, MSTRG.14953, PRKACA* |
| 1 | Blood1 Resilient vs Susceptible | Up | Transcripts | 205 | *TNFSF8, IWS1, POLR2G, MSTRG.18788, BUD13, B4GALT1, UBXN4, MSTRG.8173, UBASH3A, CAMTA2* |
| 1 | Blood1 Resilient vs Susceptible | Down | Transcripts | 257 | *CNOT8, MBP, PARVG, MSTRG.5591, MSTRG.16177, DNAJA3, GPAT3, TPM3, OARD1, CLTB* |
| 1 | Blood1 Early_dead vs Susceptible | Up | Transcripts | 129 | *REV1, ANKRD13A, KLF17, NFKBIZ, RNF151, TRAV26-2, MSTRG.3999, PPP1R12A, CLTC, CNOT6* |
| 1 | Blood1 Early_dead vs Susceptible | Down | Transcripts | 211 | *ZCCHC2, RREB1, MSTRG.32803, ZC3H4, HIBADH, ABR, CUTA, KLF17, ARHGAP4, FBXL5* |
| 2 | Blood1 Resistant vs Others | Up | Genes | 0 | |

| | | | | | |
|---|---|---|---|---|---|
| 2 | Blood1 Resistant vs Others | Down | Genes | 0 | |
| 2 | Blood1 Resilient vs Others | Up | Genes | 0 | |
| 2 | Blood1 Resilient vs Others | Down | Genes | 0 | |
| 2 | Blood1 Early_dead vs Others | Up | Genes | 0 | |
| 2 | Blood1 Early_dead vs Others | Down | Genes | 0 | |
| 2 | Blood1 Susceptible vs Others | Up | Genes | 0 | |
| 2 | Blood1 Susceptible vs Others | Down | Genes | 0 | |
| 2 | Blood1 Resistant vs Others | Up | Transcripts | 2 | *HBM, ssc-mir-10390* |
| 2 | Blood1 Resistant vs Others | Down | Transcripts | 1 | *FBXL18* |
| 2 | Blood1 Resilient vs Others | Up | Transcripts | 6 | *RPL32, HBM, ssc-mir-10390, RFC1, HBM, SNCG* |
| 2 | Blood1 Resilient vs Others | Down | Transcripts | 3 | *CRTC1, PIK3C2A, MSTRG.24540* |
| 2 | Blood1 Early_dead vs Others | Up | Transcripts | 1 | *ssc-mir-10390* |
| 2 | Blood1 Early_dead vs Others | Down | Transcripts | 1 | *RPS11* |
| 2 | Blood1 Susceptible vs Others | Up | Transcripts | 3 | *HBM, HBM, ELK4* |
| 2 | Blood1 Susceptible vs Others | Down | Transcripts | 1 | *SRGN* |

| | | | | | |
|---|---|---|---|---|---|
| 2 | Blood1 Resistant vs Resilient | Up | Genes | 0 | |
| 2 | Blood1 Resistant vs Resilient | Down | Genes | 0 | |
| 2 | Blood1 Resistant vs Early_dead | Up | Genes | 0 | |
| 2 | Blood1 Resistant vs Early_dead | Down | Genes | 0 | |
| 2 | Blood1 Resistant vs Susceptible | Up | Genes | 1 | *MSTRG.7375* |
| 2 | Blood1 Resistant vs Susceptible | Down | Genes | 0 | |
| 2 | Blood1 Resilient vs Early_dead | Up | Genes | 0 | |
| 2 | Blood1 Resilient vs Early_dead | Down | Genes | 0 | |
| 2 | Blood1 Resilient vs Susceptible | Up | Genes | 0 | |
| 2 | Blood1 Resilient vs Susceptible | Down | Genes | 0 | |
| 2 | Blood1 Early_dead vs Susceptible | Up | Genes | 0 | |
| 2 | Blood1 Early_dead vs Susceptible | Down | Genes | 0 | |
| 2 | Blood1 Resistant vs Resilient | Up | Transcripts | 6 | *MSTRG.19529, PIK3C2A, CRTC1, MSTRG.24540, FBXL18, ssc-mir-10390* |
| 2 | Blood1 Resistant vs Resilient | Down | Transcripts | 6 | *WDR26, ssc-mir-10390, CAT, HBM, HBM, CRTC1* |
| 2 | Blood1 Resistant vs Early_dead | Up | Transcripts | 1 | *RPS11* |

| | | | | | |
|---|---|---|---|---|---|
| 2 | Blood1 Resistant vs Early_dead | Down | Transcripts | 1 | *HBM* |
| 2 | Blood1 Resistant vs Susceptible | Up | Transcripts | 2 | *MAT2B, SNORD38A* |
| 2 | Blood1 Resistant vs Susceptible | Down | Transcripts | 5 | *ACSL1, ELK4, HBM, HBM, RPL23* |
| 2 | Blood1 Resilient vs Early_dead | Up | Transcripts | 9 | *RPL32, PTPRC, HBM, MSTRG.44186, RFC1, PTPRC, ssc-mir-10390, SMG1, SNCG* |
| 2 | Blood1 Resilient vs Early_dead | Down | Transcripts | 2 | *MSTRG.10907, CRTC1* |
| 2 | Blood1 Resilient vs Susceptible | Up | Transcripts | 8 | *TSGA13, RPL32, IFI44L, SPTBN1, GIMAP6, SRGN, MSTRG.44187, HBM* |
| 2 | Blood1 Resilient vs Susceptible | Down | Transcripts | 9 | *PRXL2A, TSGA13, STAT6, MSTRG.41907, MSTRG.10073, CRTC1, ssc-mir-10390, UBC, HBM,* |
| 2 | Blood1 Early_dead vs Susceptible | Up | Transcripts | 2 | *GIMAP6, SRGN* |
| 2 | Blood1 Early_dead vs Susceptible | Down | Transcripts | 6 | *AVIL, ELK4, RPS11, HBM, HBM, UBC* |
| 2 | all groups Blood3 vs Blood1 | Up | Genes | 8 | *MSTRG.11698, VWF, CXCL10, Metazoa_SRP, SERPING1, RPS29, Metazoa_SRP, MSTRG.22239* |
| 2 | all groups Blood3 vs Blood1 | Down | Genes | 6 | *MSTRG.20956, IL1R2, MSTRG.33084, MPG, NPRL3, MSTRG.49035* |
| 2 | all groups Blood4 vs Blood3 | Up | Genes | 21 | *RNaseP_nuc, CST3, MSTRG.43791, MSTRG.43792, MSTRG.43746, MSTRG.43733, MSTRG.43770, MSTRG.34674, MSTRG.43709, MSTRG.19931* |
| 2 | all groups Blood4 vs Blood3 | Down | Genes | 6 | *HSPB1, SERPING1, MSTRG.33488, VWF, LTF, CXCL10* |
| 2 | Resistant Blood3 vs Blood1 | Up | Genes | 11 | *MRPL17, VWF, GZMA, CAMK1, CXCL10, GBP5, MSTRG.11698, Metazoa_SRP, RPS29, Metazoa_SRP* |
| 2 | Resistant Blood3 vs | Down | Genes | 8 | *IL1R2, MSTRG.33084, MSTRG.20956, MPG, NPRL3,* |

| | | | | | |
|---|---|---|---|---|---|
| | Blood1 | | | | *MSTRG.30424, HBZ, MSTRG.30428* |
| 2 | Resistant Blood4 vs Blood3 | Up | Genes | 26 | *SCIMP, CST3, RNaseP_nuc, MSTRG.43733, MSTRG.43770, GZMK, MSTRG.43792, MSTRG.43791, MSTRG.43746, MS4A1* |
| 2 | Resistant Blood4 vs Blood3 | Down | Genes | 6 | *CALD1, VWF, MSTRG.33488, SERPING1, HSPB1, LTF* |
| 2 | Resilient Blood3 vs Blood1 | Up | Genes | 32 | *DDX60, MRPL17, MANF, IDO1, MSTRG.11698, MSTRG.34671, CXCL10, SERPING1, FCGR1A, IL27* |
| 2 | Resilient Blood3 vs Blood1 | Down | Genes | 13 | *CXCR4, C19orf67, SGK1, MSTRG.50671, GALNT10, IL18, MSTRG.35814, DCPS, MSTRG.40549, MSTRG.42312* |
| 2 | Resilient Blood4 vs Blood3 | Up | Genes | 30 | *CST3, SCIMP, FBXO9, MS4A1, CD72, MPG, NPRL3, MSTRG.44117, SGK1, CD79B* |
| 2 | Resilient Blood4 vs Blood3 | Down | Genes | 19 | *MRPL17, LAP3, MSTRG.14749, BNIP5, MSTRG.35506, SERPING1, CXCL10, IDO1, DDX60, FOXS1* |
| 2 | Early_dead Blood3 vs Blood1 | Up | Genes | 15 | *NKG7, Metazoa_SRP, GZMA, MRPL57, CCL5, RPS29, Metazoa_SRP, RBM38, VWF, ACTR10* |
| 2 | Early_dead Blood3 vs Blood1 | Down | Genes | 14 | *RPS26, MSTRG.20956, KPNA4, MSTRG.16304, TEX14, SGK1, SLC25A21, MSTRG.10759, IL1R2, MSTRG.30424* |
| 2 | Early_dead Blood4 vs Blood3 | Up | Genes | 13 | *MSTRG.43733, RNaseP_nuc, CST3, MSTRG.43770, RPL28, MSTRG.34674, MSTRG.19931, RPS29, Metazoa_SRP, MSTRG.30424* |
| 2 | Early_dead Blood4 vs Blood3 | Down | Genes | 5 | *MSTRG.33488, RHEX, TGFB1, RBM38, SRRM2* |
| 2 | Susceptible Blood3 vs Blood1 | Up | Genes | 30 | *SEPTIN5, MSTRG.12911, UNC93B1, MRPL17, CSF1, CXCL10, RHPN2, SERPING1, MSTRG.22239, BNIP5* |
| 2 | Susceptible Blood3 vs Blood1 | Down | Genes | 5 | *MSTRG.33084, MSTRG.20956, MSTRG.50671, IL1R2, MSTRG.49036* |
| 2 | Susceptible Blood4 vs Blood3 | Up | Genes | 47 | *MSTRG.43729, MSTRG.43733, MSTRG.34674, MSTRG.43770, MSTRG.12581, H2AC4, MSTRG.43790, MSTRG.43792, H2BC11, MSTRG.43748* |

| | | | | | |
|---|---|---|---|---|---|
| 2 | Susceptible Blood4 vs Blood3 | Down | Genes | 38 | *CALD1, SEPTIN5, F13A1, UNC93B1, TREML1, LTF, TCIRG1, SERPING1, EVL, MX2* |

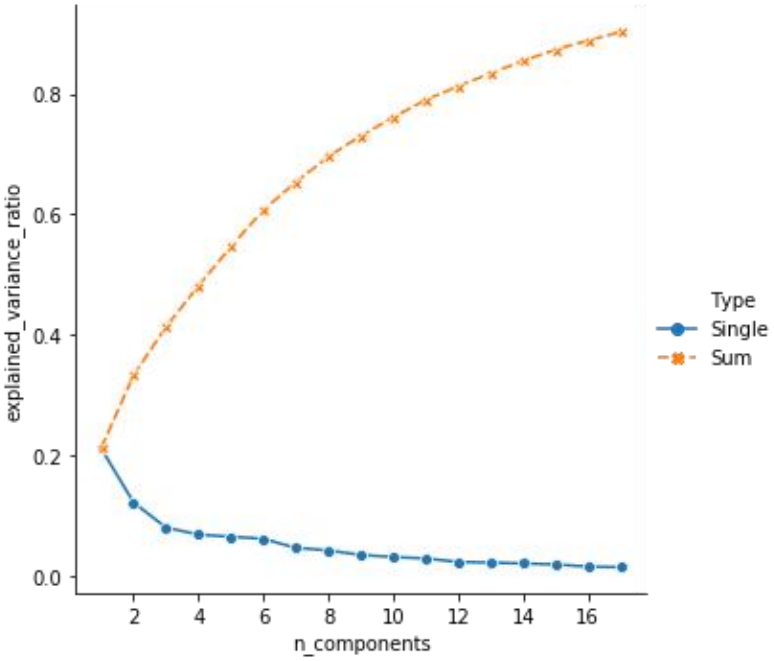
**Appendix 4.1** Significance of ANOVA analysis for batch effect.

| Metabolites | p.value | FDR | Fisher's LSD |
|---|---|---|---|
| Hypoxanthine | 7.37E-52 | 2.80E-50 | 27 - 26; 28 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 28; 27 - 29; 27 - 30; 31 - 27; 28 - 29; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 31 - 29; 32 - 29; 31 - 30; 31 - 32 |
| Dimethylglycine | 1.48E-40 | 2.82E-39 | 27 - 26; 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 30 - 27; 32 - 27; 30 - 28; 32 - 28; 30 - 29; 32 - 29; 30 - 31; 30 - 32; 32 - 31 |
| Choline | 6.93E-34 | 8.77E-33 | 27 - 26; 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 30 - 27; 31 - 27; 32 - 27; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 31 - 29; 32 - 29; 31 - 30; 32 - 30 |
| Tyrosine | 4.93E-29 | 4.68E-28 | 27 - 26; 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 28; 27 - 30; 29 - 28; 31 - 28; 29 - 30; 31 - 30 |
| Succinate | 3.42E-26 | 2.60E-25 | 27 - 26; 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 29; 32 - 27; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 31 - 29; 32 - 29; 32 - 30; 32 - 31 |
| Creatinine | 1.27E-25 | 8.04E-25 | 27 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 28; 30 - 27; 29 - 28; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 31 - 29; 30 - 31; 30 - 32; 31 - 32 |
| Citric acid | 7.56E-25 | 4.10E-24 | 27 - 26; 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 30; 28 - 30; 29 - 30; 32 - 29; 31 - 30; 32 - 30; 32 - 31 |
| Pyruvic acid | 2.85E-23 | 1.36E-22 | 27 - 26; 26 - 29; 31 - 26; 26 - 32; 27 - 28; 27 - 29; 27 - 30; 27 - 32; 28 - 29; 31 - 28; 28 - 32; 30 - 29; 31 - 29; 31 - 30; 30 - 32; 31 - 32 |
| L-Threonine | 3.34E-21 | 1.41E-20 | 26 - 28; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 29 - 27; 30 - 27; 31 - 27; 32 - 27; 29 - 28; 30 - 28; 31 - 28; 32 - 28; 29 - 30; 29 - 31; 29 - 32 |
| L-Phenylalanine | 5.79E-19 | 2.20E-18 | 27 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 27 - 28; 30 - 27; 32 - 27; 29 - 28; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 32 - 29 |
| Isoleucine | 3.10E-18 | 1.07E-17 | 26 - 27; 26 - 28; 26 - 29; 26 - 30; 26 - 31; 26 - 32; 27 - 28; 31 - 28 |
| 3-Methyl-2-oxovaleric acid | 1.89E-17 | 5.98E-17 | 26 - 27; 26 - 28; 26 - 29; 26 - 30; 26 - 31; 26 - 32; 27 - 30; 27 - 32; 28 - 30; 28 - 32; 29 - 30; 29 - 32; 31 - 30; 31 - 32 |
| Glycerol | 1.73E-16 | 5.04E-16 | 26 - 27; 26 - 28; 26 - 30; 26 - 31; 26 - 32; 27 - 28; 27 - 31; 27 - 32; 29 - 28; 30 - 28; 28 - 32; 29 - 31; 29 - 32; 30 - 31; 30 - 32 |
| Creatine | 3.25E-16 | 8.81E-16 | 26 - 27; 26 - 28; 26 - 29; 26 - 30; 26 - 31; 26 - 32; 27 - 30; 27 - 32; 31 - 28; 29 - 32; 31 - 30; 31 - 32 |
| L-Alpha-aminobutyric acid | 1.13E-12 | 2.87E-12 | 26 - 27; 26 - 29; 26 - 30; 26 - 31; 26 - 32; 27 - 29; 27 - 30; 27 - 31; 28 - 29; 28 - 30; 28 - 31; 28 - 32; 32 - 30; 32 - 31 |
| L-Asparagine | 1.84E-12 | 4.36E-12 | 28 - 26; 29 - 26; 30 - 26; 31 - 26; 32 - 26; 28 - 27; 29 - 27; 28 - 30; 28 - 31; 28 - 32; 29 - 30; 29 - 31; 29 - 32 |
| Urea | 2.26E-12 | 5.04E-12 | 26 - 27; 26 - 28; 26 - 29; 27 - 28; 27 - 29; 30 - 28; 31 - 28; 32 - 28; 30 - 29; 31 - 29; 32 - 29 |
| L-Lysine | 5.74E-12 | 1.21E-11 | 27 - 26; 29 - 26; 31 - 26; 32 - 26; 27 - 28; 27 - 30; 32 - 27; 29 - 28; 31 - 28; 32 - 28; 29 - 30; 32 - 29; 31 - 30; 32 - 30 |
| L-Serine | 1.18E-10 | 2.35E-10 | 26 - 27; 26 - 31; 28 - 27; 29 - 27; 30 - 27; 32 - 27; 28 - 31; 29 - 31; 30 - 31; 32 - 31 |

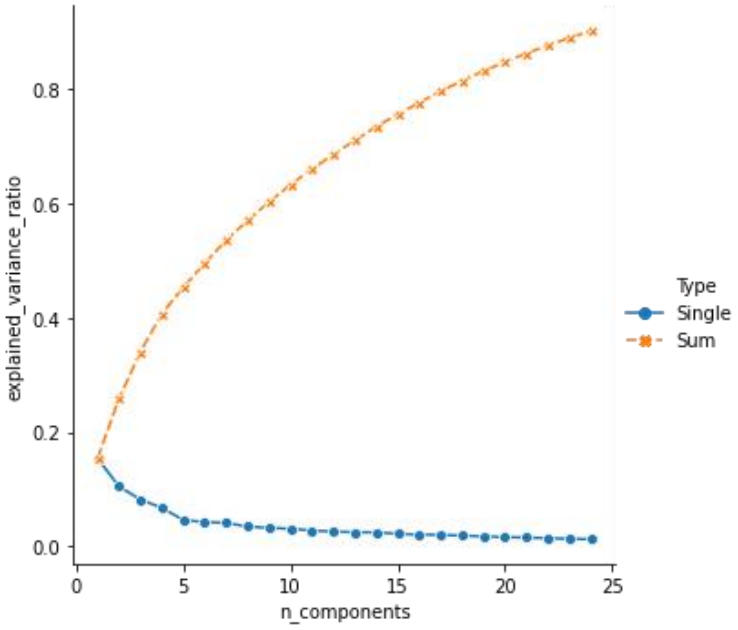
| Metabolites | p.value | FDR | Fisher's LSD |
|---|---|---|---|
| Methionine | 4.65E-10 | 8.83E-10 | 27 - 26; 28 - 26; 32 - 26; 27 - 29; 27 - 30; 27 - 31; 27 - 32; 28 - 29; 28 - 30; 28 - 31; 28 - 32; 32 - 29; 32 - 30 |
| Isobutyric acid | 4.77E-09 | 8.63E-09 | 26 - 28; 26 - 29; 26 - 30; 26 - 32; 27 - 28; 27 - 30; 27 - 32; 29 - 28; 30 - 28; 31 - 28; 29 - 32; 31 - 30; 31 - 32 |
| Formate | 2.73E-08 | 4.72E-08 | 26 - 27; 26 - 28; 26 - 29; 26 - 32; 30 - 27; 31 - 27; 28 - 32; 30 - 29; 31 - 29; 30 - 32; 31 - 32 |
| L-Glutamic acid | 1.12E-07 | 1.84E-07 | 28 - 26; 26 - 29; 26 - 30; 26 - 31; 28 - 27; 27 - 31; 28 - 29; 28 - 30; 28 - 31; 28 - 32; 32 - 29; 32 - 30; 32 - 31 |
| Mannose | 1.05E-06 | 1.66E-06 | 26 - 27; 29 - 26; 26 - 31; 29 - 27; 29 - 28; 29 - 30; 29 - 31; 29 - 32 |
| Aspartate | 1.76E-06 | 2.67E-06 | 29 - 26; 29 - 27; 30 - 27; 29 - 28; 29 - 31; 29 - 32; 30 - 31; 30 - 32 |
| 2-Hydroxybutyrate | 2.99E-05 | 4.36E-05 | 26 - 27; 26 - 28; 26 - 29; 26 - 30; 26 - 31; 26 - 32; 27 - 29; 27 - 30; 32 - 29; 32 - 30 |
| Oxoglutarate | 6.84E-05 | 9.63E-05 | 28 - 26; 28 - 27; 30 - 27; 31 - 27; 28 - 29; 28 - 31; 28 - 32; 30 - 29; 30 - 32; 31 - 32 |
| Acetic acid | 7.51E-05 | 1.02E-04 | 27 - 26; 29 - 26; 30 - 26; 31 - 26; 27 - 28; 27 - 32; 31 - 28; 29 - 32; 30 - 32; 31 - 32 |
| D-Glucose | 1.63E-04 | 2.13E-04 | 29 - 26; 26 - 30; 28 - 27; 29 - 27; 28 - 30; 28 - 31; 29 - 30; 29 - 31; 29 - 32 |
| ornithine | 1.77E-04 | 2.24E-04 | 27 - 26; 28 - 26; 32 - 26; 27 - 31; 28 - 30; 28 - 31; 32 - 31 |
| L-Glutamine | 2.81E-04 | 3.45E-04 | 26 - 27; 28 - 27; 29 - 27; 32 - 27; 28 - 30; 28 - 31; 29 - 30; 29 - 31; 29 - 32 |
| L-Alanine | 2.17E-03 | 2.58E-03 | 27 - 26; 27 - 28; 27 - 29; 27 - 30; 31 - 29; 32 - 29; 31 - 30; 32 - 30 |
| L-proline | 6.22E-03 | 7.17E-03 | 28 - 26; 32 - 26; 28 - 30; 28 - 31; 32 - 30; 32 - 31 |
| L-Lactic acid | 7.25E-03 | 8.11E-03 | 26 - 29; 27 - 29; 28 - 29; 30 - 29; 31 - 29; 32 - 29 |
| Valine | 1.39E-02 | 1.51E-02 | 29 - 26; 29 - 28; 29 - 30; 29 - 31; 29 - 32 |
| Glycine | 1.95E-02 | 2.06E-02 | 30 - 26; 30 - 27; 30 - 28; 30 - 29; 30 - 31; 30 - 32 |

**Appendix 4.2** Plot of explained variance, PCA result for original metabolites concentration(A), PCA results for batch adjusted metabolites concentration(B).

**A**




**B**



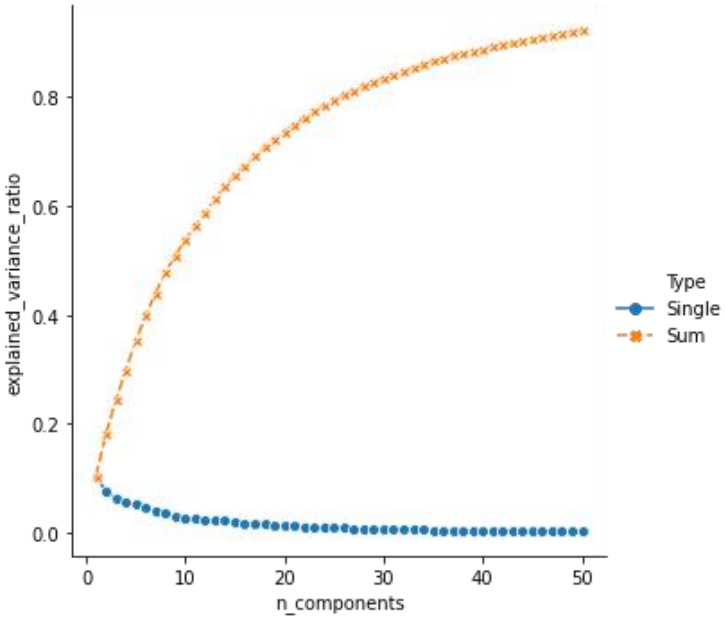
**Appendix 4.3** Results of machine learning for sex prediction of the dataset "MTBLS90: Large-scale non-targeted serum metabolomics in the Prospective Investigation of the Vasculature in Uppsala Seniors".

(A) Accumulated accuracies of each different machine learning algorithm for predicting the subject sex. (B) Plot of PCA components explained data variance. (C) Result of permutation test.

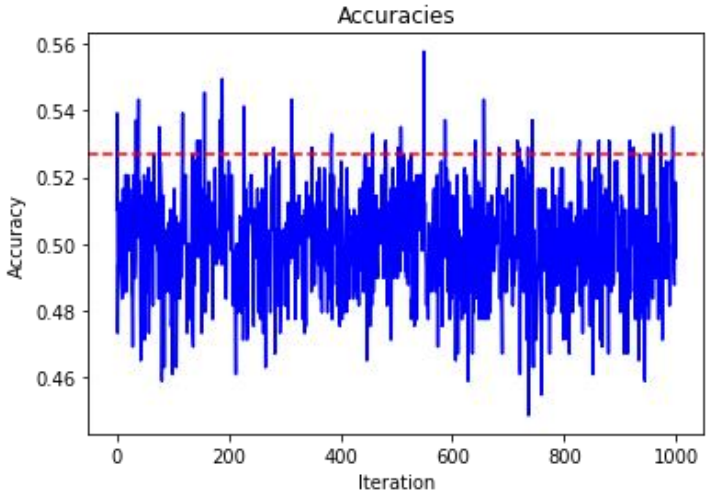
**A**

| Accuracy | con | conPCA26 | conPCA44 |
|---|---|---|---|
| **KNN** | 0.75 | 0.74 | 0.74 |
| **DT** | 0.67 | 0.59 | 0.54 |
| **RF** | 0.79 | 0.69 | 0.72 |
| **SVM** | 0.83 | 0.77 | 0.81 |
| **LR** | 0.71 | 0.77 | 0.78 |

**B**



**C**


Accuracies
Accuracy
0.56
0.54
0.52
0.50
0.48
0.46
0
200
400
600
800
1000
Iteration

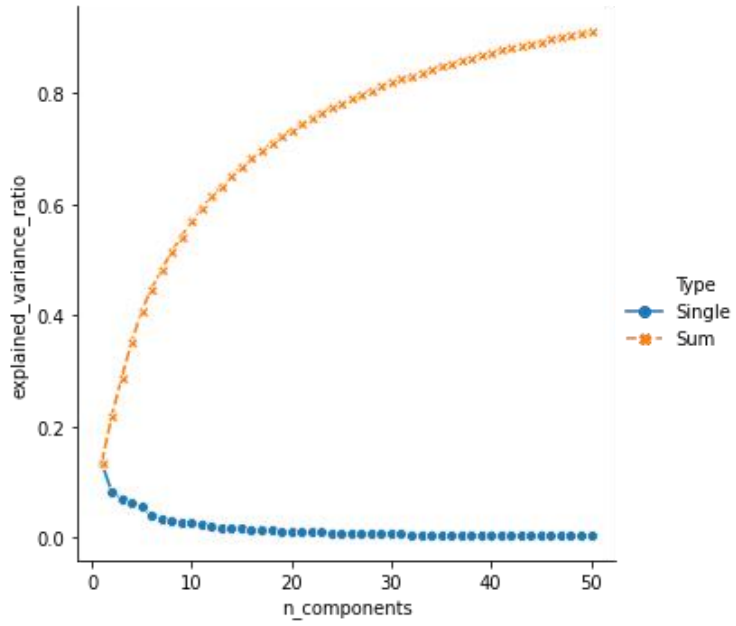
**Appendix 4.4** Results of machine learning for sex prediction of the dataset "ST000369: Investigation of metabolomic blood biomarkers for detection of adenocarcinoma lung cancer (part II)".

(A) Accumulated accuracies of each different machine learning algorithm for predicting the subject sex. (B) Plot of PCA components explained data variance. (C) Result of permutation test.

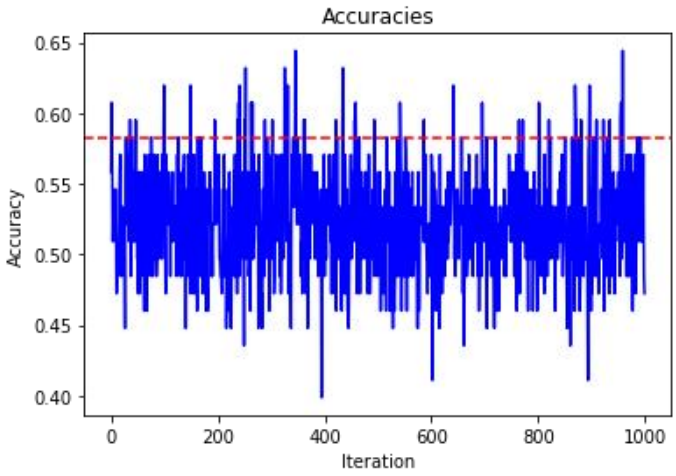
**A**

| Accuracy | con | conPCA28 | conPCA49 |
|---|---|---|---|
| **KNN** | 0.61 | 0.59 | 0.61 |
| **DT** | 0.49 | 0.66 | 0.61 |
| **RF** | 0.63 | 0.66 | 0.71 |
| **SVM** | 0.76 | 0.71 | 0.68 |
| **LR** | 0.83 | 0.68 | 0.76 |

**B**



**C**


Accuracies
0.65
0.60
0.55
0.50
0.45
0.40
Accuracy
0
200
400
600
800
1000
Iteration