

IDENTIFICATION OF THE GENETIC COMPONENT OF HOST SUSCEPTIBILITY TO VIRAL INFECTION IN MICE

Gregory A. Boivin

DEPARTMENT OF HUMAN GENETICS

2015

A thesis submitted to McGill University in partial fulfillment of the
requirements of the degree of Doctor of Philosophy

© Gregory Boivin, 2015

ACKNOWLEDGMENTS

I would like to thank Dr Silvia Vidal and Dr Rob Sladek for all of their wisdom and guidance across the duration of this project. I would like to thank members of the Vidal lab over the years for their help and support with my projects, most notably Dr Sean Wiltshire, Dr Nassima Fodil-Cornu, Dr Julien Pothlichet, Jenn Marton, Gab Leiva, and Mathieu Mancini. I would like to thank the animal technicians that helped with this work, specifically Patricia D'Arcy. I would like to thank several other professors who contributed to the work seen here and to my knowledge and growth during my PhD, specifically Dr Earl Brown from the University of Ottawa, and Dr JC Loredó-Osti from Memorial University. Lastly, I would like to thank my family, who have always given me endless support and help in the pursuit of all of my passions.

TABLE OF CONTENTS

TITLE	
ACKNOWLEDGMENTS	002
TABLE OF CONTENT	003
LIST OF FIGURES.....	007
LIST OF TABLES.....	009
LIST OF ABBREVIATIONS.....	010
ABSTRACT	011
RÉSUMÉ.....	012
PREFACE AND CONTRIBUTIONS OF AUTHORS.....	014
 <u>CHAPTER 1: GENERAL INTRODUCTION</u>	
RATIONALE.....	015
<i>The genetic basis of infectious disease susceptibility</i>	015
<i>Mouse genetic models of infectious disease</i>	016
HYPOTHESIS.....	017
OBJECTIVES.....	017
1. THE AGENTS: RELEVANCE TO HUMAN POPULATION AND PATHOGENESIS	018
SMALLPOX.....	018
<i>Clinicopathological features</i>	019
<i>Poxviruses</i>	021
<i>Cellular stage</i>	022
<i>Animal models</i>	023
<i>Mouse models</i>	024
<i>Poxvirus specific immunity</i>	025
<i>Immune evasion</i>	026
<i>Genetic control of host response</i>	026
INFLUENZA.....	027
<i>Clinicopathological features</i>	028
<i>Antivirals and vaccines</i>	029
<i>Orthomyxoviruses</i>	030
<i>Cellular stage</i>	032
<i>Animal models</i>	033
<i>Host response</i>	034
<i>Immune evasion</i>	036
GENETIC CONTROL OF HOST RESPONSE.....	036

2. THE MOUSE AS A MODEL ORGANISM	038
MOUSE GENETIC RESOURCES FOR COMPLEX TRAIT ANALYSIS	039
GENETIC MAPPING IN F2 AND N2 CROSSES	040
MARKER ASSISTED BREEDING	040
WILD-DERIVED MOUSE STRAINS	041
RECOMBINANT INBRED STRAINS	041
THE COLLABORATIVE CROSS	042
OTHER SPECIALIZED PANELS OF MICE	042
<i>Chromosome substitution strains</i>	042
<i>Genome tagged mice</i>	043
THE PROBLEM OF RESOLVING THE QTL POSITION	043
<i>Advanced intercross lines</i>	043
<i>Heterogeneous stock</i>	044
RECOMBINANT CONGENIC STRAINS	044
3. THE METHOD: FORWARD GENETICS IN MICE	045
QTL IDENTIFICATION: GENETIC MARKERS AND MAPS	045
QTL IDENTIFICATION: STATISTICAL ANALYSIS	046
COMPLEX TRAITS ANALYSIS I: ANOVA	047
COMPLEX TRAITS ANALYSIS II: INTERVAL MAPPING	049
COMPLEX TRAITS ANALYSIS III: MULTIPLE REGRESSION	050
GENOME-WIDE SIGNIFICANCE	051
<i>The Bonferroni correction</i>	051
<i>The Benjamini-Hochberg correction</i>	051
<i>Permutations</i>	051
POPULATION STRUCTURE	052
FIXED AND RANDOM EFFECT MODELS	053
MIXED EFFECTS MODEL	054
QTL IDENTIFICATION: CANDIDATE GENES	055
CANDIDATE GENE IDENTIFICATION: BIONFORMATIC ANALYSIS	055
CANDIDATE GENE IDENTIFICATION: EXPRESSION ANALYSIS	056
EXPRESSION QTL	057
 BRIDGING STATEMENT FROM CHAPTER 1 TO CHAPTER 2	 059

CHAPTER 2: A MAJOR LOCUS ON CHROMOSOME 17 CONFERS INNATE RESISTANCE TO LETHAL MOUSEPOX INFECTION

ABSTRACT	061
INTRODUCTION.....	062
MATERIAL AND METHODS..	064
RESULTS.....	067
DISCUSSION.....	070
BRIDGING STATEMENT FROM CHAPTER 2 TO CHAPTER 3.....	080

CHAPTER 3: MAPPING OF CLINICAL AND EXPRESSION QUANTITATIVE TRAIT LOCI IN A SEX-DEPENDENT EFFECT OF HOST SUSCEPTIBILITY TO MOUSE-ADAPTED INFLUENZA H3N2/HK/1/68

ABSTRACT	082
INTRODUCTION.....	083
MATERIAL AND METHODS.....	086
RESULTS.....	092
DISCUSSION.....	099
BRIDGING STATEMENT FROM CHAPTER 3 TO CHAPTER 4.....	116

CHAPTER 4: MULTIPLE GENETIC LOCI ON MOUSE CHROMOSOME 11 PROVIDE PROTECTION AGAINST INFLUENZA VIRUS INFECTION

ABSTRACT	118
INTRODUCTION.....	119
MATERIAL AND METHODS.....	121
RESULTS.....	126
DISCUSSION.....	131
BRIDGING STATEMENT FROM CHAPTER 4 TO CHAPTER 5.....	145

CHAPTER 5: Pla2g7-DEFICIENCY IS ASSOCIATED WITH DAMPENED INFLAMMATORY RESPONSES AND ENHANCED HOST RESISTANCE TO CHALLENGE WITH MOUSE-ADAPTED INFLUENZA VIRUS

ABSTRACT	147
INTRODUCTION.....	148
MATERIAL AND METHODS..	151
RESULTS.....	156
DISCUSSION.....	160

BRIDGING STATEMENT FROM CHAPTER 5 TO CHAPTER 6.....	173
GENERAL DISCUSSION.....	174
CONCLUSION.....	186
GENERAL REFERENCES.....	188

LIST OF FIGURES

CHAPTER 2: A MAJOR LOCUS ON CHROMOSOME 17 CONFERS INNATE RESISTANCE TO LETHAL MOUSEPOX INFECTION

FIGURE 1: INTER-STRAIN PHENOTYPIC DIFFERENCES IN HOST RESPONSE TO ECTROMELIA VIRUS	074
FIGURE 2: INCREASED SUSCEPTIBILITY TO ECTROMELIA VIRUS MAPS TO CHROMOSOME 17	075
FIGURE 3: H2T23 CANDIDATE GENE IDENTIFIED BY OVERLAPPING CLINICAL AND EXPRESSION QTL	076
FIGURE 4: CONGENIC MOUSE LINE WAS MADE FROM SUSCEPTIBLE BCA70 MICE.....	077
FIGURE 5: H2T23 CANDIDATE GENE VALIDATION.....	078
FIGURE 6: MODEL FOR SUSCEPTIBILITY TO ECTROMELIA VIRUS IN THE MOUSE.....	079

CHAPTER 3: MAPPING OF CLINICAL AND EXPRESSION QUANTITATIVE TRAIT LOCI IN A SEX-DEPENDENT EFFECT OF HOST SUSCEPTIBILITY TO MOUSE-ADAPTED INFLUENZA H3N2/HK/1/68

FIGURE 1: INTERSTRAIN PHENOTYPIC DIFFERENCES IN HOST RESPONSE TO MOUSE-ADAPTED H3N2 CLINICAL ISOLATE HK/68.....	105
FIGURE 2: INCREASED LUNG NEUTROPHILIA, CYTOTOXICITY, AND INFLAMMATION, BUT NOT VIRAL LOAD, CORRELATE WITH SUSCEPTIBILITY.....	107
FIGURE 3: COMPLEX SEX-SPECIFIC GENETIC CONTROL OF HOST RESPONSE	109
FIGURE 4: INCREASED SUSCEPTIBILITY MAPS TO SEX-SPECIFIC CANDIDATE CQTL..	110
FIGURE 5: MALE-SPECIFIC CQTL CONFIRMED THROUGH A SECONDARY F ₂ SCREEN	111
FIGURE 6: PRIMARY CANDIDATE GENES IDENTIFIED BY COMBINING CQTL AND EQTL DATASETS	112
FIGURE 7: CHROMOSOME 17 QTL ASSOCIATED WITH DECREASED SURVIVAL AND INCREASED CYTOKINE EXPRESSION.....	113

LIST OF FIGURES (CONTINUED)

CHAPTER 4: MULTIPLE GENETIC LOCI ON MOUSE CHROMOSOME 11 PROVIDE PROTECTION AGAINST INFLUENZA VIRUS INFECTION

FIGURE 1: BCA80 MICE ARE RESISTANT TO INFLUENZA INFECTION AS COMPARED TO B6 CONTROLS	135
FIGURE 2: RESISTANT BCA80 MICE SHOW ALTERED VIRAL TITER, LUNG AND SPLEEN HISTOLOGY AND LUNG CELL COUNTS	136
FIGURE 3: GENETIC MAPPING OF RESISTANCE TO INFLUENZA VIRUS	138
FIGURE 4: ADDITIONAL RESISTANCE LOCI VALIDATED ON CHROMOSOME 11	139
FIGURE 5: ADDITIONAL CANDIDATE GENES IDENTIFIED BY COMBINING CLINICAL AND EXPRESSION QTL	140
FIGURE 6: GENETIC MAPPING OF RESISTANCE TO INFLUENZA VIRUS IMPLICATES A <i>DE NOVO</i> MUTATION IN THE <i>NF1</i> GENE	141

CHAPTER 5: PLA2G7-DEFICIENCY IS ASSOCIATED WITH DAMPENED INFLAMMATORY RESPONSES AND ENHANCED HOST RESISTANCE TO CHALLENGE WITH MOUSE-ADAPTED INFLUENZA VIRUS

FIGURE 1: HAPLOTYPE STRUCTURE OF THE CONGENIC SEGMENT OF CHROMOSOME 17 FIXED IN B6.A-PLA2G7 MICE	163
FIGURE 2: INCREASED PLA2G7 ACTIVITY IS ASSOCIATED WITH SUSCEPTIBILITY TO INFLUENZA VIRUS	164
FIGURE 3: HISTOLOGICAL ANALYSIS OF RESISTANT <i>PLA2G7</i> ^{-/-} AND WILD-TYPE CONTROLS	165
FIGURE 4: RESISTANT <i>PLA2G7</i> ^{-/-} MICE SHOW ELEVATED CYTOKINE SIGNALING ON DAY 1 POST INFECTION, BUT REDUCED CYTOKINE LEVELS BY DAY 6 POST INFECTION AS COMPARED TO B6 CONTROLS	166
FIGURE 5: DYSREGULATED GENE EXPRESSION BETWEEN <i>PLA2G7</i> ^{-/-} AND CONTROL MICE	168
SUPPLEMENTAL FIGURE 1	169

LIST OF TABLES

CHAPTER 3: MAPPING OF CLINICAL AND EXPRESSION QUANTITATIVE TRAIT LOCI IN A SEX-DEPENDENT EFFECT OF HOST SUSCEPTIBILITY TO MOUSE-ADAPTED INFLUENZA H3N2/HK/1/68

TABLE 1: PCR PRIMERS..... 114

TABLE 2: SIGNIFICANT CQTL IN THE TOTAL, MALE, AND FEMALE

RCS POPULATIONS..... 115

CHAPTER 4: MULTIPLE GENETIC LOCI ON MOUSE CHROMOSOME 11 PROVIDE PROTECTION AGAINST INFLUENZA VIRUS INFECTION

TABLE 1: MARKERS USED FOR CONGENIC AND F₂ MICE..... 142

TABLE 2: LOCATION AND TYPE OF QTL..... 143

SUPPLEMENTAL TABLE 1..... 144

CHAPTER 5: Pla2g7-DEFICIENCY IS ASSOCIATED WITH DAMPENED INFLAMMATORY RESPONSES AND ENHANCED HOST RESISTANCE TO CHALLENGE WITH MOUSE-ADAPTED INFLUENZA VIRUS

TABLE 1: SIGNIFICANTLY ALTERED PATHWAYS BY IPA ON DAY 6

POST-INFECTION..... 170

SUPPLEMENTAL TABLE 1..... 171

SUPPLEMENTAL TABLE 2..... 172

LIST OF ABBREVIATIONS

Actc1 - actin, alpha, cardiac muscle	NA - neuraminidase
Acta2 - actin, alpha 2, smooth muscle	NK - natural killer
ANOVA - analysis of variance	NLR - NOD-like receptor
B6 - C57BL/6	NP - nucleoprotein
BALF - bronchoalveolar lavage fluid	NS-1 - non-structural protein 1
BG - background	NS-2 - non-structural protein 2
C5 - hemolytic complement	OAS - 2'5'oligoadenylate synthetase
CSS - Chromosome substitution strain	OxPL - oxidized phospholipids
DNA - deoxyribonucleic acid	PA - polymerase acid
DSO - donor strain of origin	PAF - platelet activating factor
ECTV - ectromelia virus	PAMP - Pathogen-associated molecular pattern
EEV - extracellular enveloped virion	PB1 - polymerase basic 1
EMMA - Efficient Mixed Model Association	PB2 - polymerase basic 2
eQTL - expression quantitative trait locus	PBS - phosphate buffered solution
FBS – fetal bovine serum	PCA - principal component analysis
Gapdh - Glyceraldehyde-3-Phosphate Dehydrogenase	PCR - Polymerase chain reaction
H&E - hematoxylin and eosin	PFA - paraformaldehyde
HA - hemagglutinin	PFU - Plaque forming units
HPRT - hypoxanthine phosphoribosyl-transferase 1	PKR - Protein kinase R
Ifitm3 - interferon-inducible transmembrane 3	Pla2g7 - Phospholipase A2, Group VII
IMV - intracellular mature virion	PR8 - A/PR/8/34
Itgb1 - integrin, beta 1	qPCR – quantitative polymerase chain reaction
KC - keratinocyte chemoattractant	QTL - quantitative trait locus
LDH - lactate dehydrogenase	RCS - recombinant congenic strains
LDL - low-density lipoprotein	RFLP - restriction fragment length polymorphism
LOD - logarithm of odds	RMP - Resistance to mousepox
LPS - Lipopolysaccharides	RNA - ribonucleic acid
Lyso-PC - lysophosphatidylcholine	RNAse L - ribonuclease L
MCP-1 - monocyte chemoattractant protein 1	RNP - Ribonucleoprotein
MHC - major histocompatibility complex	SSLP - Simple sequence length polymorphism
MP1 - matrix protein 1	TNF - Tumor necrosis factor
MP2 - matrix protein 2	
Mx1 or Mx2 - myxovirus resistance protein	
Myl7 - myosin, light chain 7	

ABSTRACT

Influenza virus causes 250,000 deaths annually worldwide and is the only known pathogen to cause recurrent pandemics (e.g., 1918 – 20 million deaths, 1957 – 2 million deaths, 1968 – 1 million deaths). Ectromelia virus is a mouse model of smallpox, a disease that has killed more individuals than all other infectious diseases combined, and whose viral mechanisms have yet to be fully understood. By studying immunological responses to these infectious diseases using controlled mouse models, it may be possible to discover novel mechanisms of resistance against these pathogens and ultimately identify novel therapeutic targets or immunological insights related to human health. The identification of causative genes underlying mouse QTL remains challenging. In the present thesis, we aimed to bypass the creation of large candidate gene lists by reducing QTL interval size with the use of a unique set of closely related mouse strains known as the AcB/BcA set of recombinant congenic strains (RCS). By combining cQTL from survival data and eQTL from liver or lung tissue of RCS mice, we identified *H2-T23* as a compelling new candidate gene for the control of resistance to ectromelia virus. We validate the likely involvement of *C5*, identify three new QTL on mouse chromosome 11, and identify *Pla2g7* as involved in resistance to influenza infection in mice.

RÉSUMÉ

Malgré des campagnes de vaccinations de part le monde, le virus de la grippe (influenza virus) saisonnière est responsable de 250 000 décès annuels. De plus ce virus est connu pour causer des pandémies tout aussi dévastatrices que régulières (e.g. 20 millions de décès en 1918; 2 million de décès, 1957; 1 million de décès, 1968; 0.3 millions de décès, 2009). L'infection avec le virus ectromelia est un modèle de la variole, une maladie infectieuse qui se distingue pour avoir été la plus létale de l'histoire mais aussi la seule à avoir été éradiquée par l'homme. Cette maladie est causée par un virus capable de manipuler fortement le système immunitaire par des mécanismes dont la compréhension nous ait incomplète. L'étude des réponses immunitaires aux infections virales en utilisant des modèles des lignées des souris consanguines devrait faciliter l'identification de nouveaux mécanismes de résistance de l'hôte contre l'infection par ces pathogènes. Ces découvertes devraient approfondir les connaissances à propos de notre système immunitaire, et éventuellement révéler de nouvelles cibles thérapeutiques contre des agents ravageurs.

Notre approche a consisté tout d'abord à déceler des différences dans la sensibilité à l'infection entre des lignées pures des souris, pour ensuite caractériser la composante génétique de la réponse à l'infection. Cette réponse étant constituée par de nombreux facteurs génétiques (quantitative trait loci or QTL), notre but a été de contourner deux défis majeurs pour l'identification des gènes causaux : la grande taille des intervalles génomiques sous-jacents les régions des QTL, ainsi que les longues listes des gènes y résident. Pour s'y faire, nous avons utilisé une série des lignées recombinantes des souris (RCS) dérivés des souches A/J et C57BL/6 (AcB/BcA) pour des études de liaison génétique. Ces travaux on été combinés aux études fonctionnelles des intervalles en employant des données d'expression génique pour localiser des

QTL qui contrôlent le niveau d'expression de gènes (eQTL) dans le foie (dans le

cas du virus ectromelia) ou le poumon (dans le cas du virus de la grippe). Ce nous a permis d'identifier le gène *H2-T23* comme étant un joueur clé dans la résistance au virus ectromelia. En ce qui concerne le virus de la grippe, nos études de cartographie combinées à la caractérisation immunologique de la réponse à l'infection dans des nouvelles lignées des souris nous ont permis de valider l'implication du gène *C5*, d'identifier deux nouveaux QTL sur le chromosome 11, ainsi que décerner un mécanisme de résistance à l'infection qui dépend de l'expression du gène *Pla2g7*. Nos travaux apportent ainsi de nouvelles pistes pour la compréhension des mécanismes de résistance contre des infections virales d'impacte majeur.

PREFACE AND CONTRIBUTIONS OF AUTHORS

Chapter 2: Sanda Remakus performed all infectious experiments related to ectromelia virus. Gregory Boivin performed the genotyping and statistical analyses and wrote the manuscript. Dr Silvia Vidal, Dr Rob Sladek and Dr Luis Sigal conceived, oversaw and financed the study. To be submitted for publication.

Chapter 3: Julien Pothlichet performed experiments related to the bone marrow derived macrophage extraction and testing, and the ELISA represented in Figure 2E. Gregory Boivin performed the in vivo experiments, the genotyping, the statistical analysis with the guidance of Dr JC Loredó Osti and Dr Rob Sladek, and wrote the manuscript. Dr Earl Brown provided the virus. Dr Silvia Vidal and Dr Rob Sladek conceived, oversaw and financed the study. Published in the Journal of Immunology.

Chapter 4: Mathieu Mancini performed the tissue extraction and FACS analysis. Pablo E. Cingolani performed the whole genome sequencing analysis. Gregory Boivin performed the genotyping, the viral infections and wrote the manuscript. Dr Silvia Vidal and Dr Rob Sladek conceived, oversaw and financed the study. To be submitted for publication.

Chapter 5: Claudia Duerr performed the tissue extraction and FACS analysis. Martin Brandt performed the PLA2G7 activity assay. Colin MacPhee from GlaxoSmithKline provided the *Pla2g7*^{-/-} mice. Gregory Boivin performed the genotyping, viral infections, statistical analysis and wrote the manuscript. Dr Silvia Vidal, Dr Rob Sladek and Dr Jorg Fritz conceived, oversaw and financed the study. To be submitted for publication.

CHAPTER 1: GENERAL INTRODUCTION

RATIONALE

Infectious diseases continue to pose a tremendous toll to human subjects worldwide. Smallpox may be a threat as both a zoonotic disease and biological weapon; while current vaccination and antiviral therapies are not sufficient to prevent more than 30,000 deaths/year due to influenza infection among the elderly, pregnant women and young children, the populations most susceptible to develop influenza pneumonitis during endemic infection. Pandemic influenza strains remain difficult to predict. For example, the 2009 epidemics emerged from pigs rather than from poultry as expected. Solving these problems requires a large number of measures at the societal and individual level. Research offers the possibility to understand the fundamental mechanisms at play during disease. The genetic analysis of infectious diseases is a proven approach to identify key molecules during the infectious process. The identification of such molecules can lead to a better understanding of the host response against infection, and equally as important, can point to mechanisms that may lead to new interventions.

The genetic basis of infectious disease susceptibility: Many human infections can lead to differing outcomes in individuals, often with only a portion of those infected developing severe symptoms. These types of observations (e.g., early remarks of tuberculosis being overrepresented in certain households, as well as later twin and adoption studies [1], [2], [3], [4] indicated that factors intrinsic to individuals could be attributed to their varying levels of susceptibility, and that these factors could be studied in human populations.

Studies of infectious diseases in human populations have been successful in identifying host specific factors that control response to infection [5]. However, it is extremely difficult to recruit cohorts with a large number of human cases and controls collected with standardized protocols at a reasonably similar time after the onset or progression of an infection. While this thesis was ongoing, there

were a limited number of studies that have been vigilant enough to collect such a set for patients with influenza infection (e.g., [6]). Indeed, one of the later challenges in this thesis was finding a human population of influenza patients with which we could test our results, a challenge that generally fell unmet.

If studying the underlying host responses to specific infectious diseases in live human populations is not feasible, one can turn to human cellular models in which host factors can be systematically altered during infection. These models have been successfully applied for infectious diseases (e.g., [7], [8]). However, they are generally restricted to specific aspects of infection (e.g., viral entry or replication), and do not fully encompass the complexity of an infection in a host.

Mouse genetic models of infectious disease: Animal models are one alternative that can adequately represent the complexity of a natural infection, and provide the tools needed to identify and validate new mechanisms of susceptibility to infection diseases. Mouse models of infection have been used to successfully dissect disease phenotypes [9], [10], [11], [12]. Importantly, genes originally found in mice have been shown to play key roles in human susceptibility to infection, providing evidence for the conservation of host immune defense between mouse and human [13]. However, studies in mouse models of infection are not without their problems. Aside from the obvious issues (i.e., that mice are not people), their use in the study of heritable traits generally ends with large portions of the genome being linked or associated to the trait of study. This means that the deliverables of a study are long lists of potential candidate mechanisms instead of specific targets for future therapeutic interventions.

Incredibly, there are over 8000 mouse loci currently linked to the control of physiologic or disease-related phenotypes, with almost 1000 tagged as immune related [14]. The number of loci that have translated into distinct mechanisms is small (e.g., estimated in 2005 to be less than 1% for rodent studies, [15], [16]). As such, increasing the speed and efficiency with which we can find mechanisms underlying diseases using mouse models is extremely important. This is

especially true if we wish to continue to use mouse models to gain insights that outweigh the complex ethical ramifications of their use in severe disease.

HYPOTHESIS

In this thesis, we use a panel of closely related mice to study the host response against two viruses, influenza virus and ectromelia virus. Influenza virus causes 250,000 deaths annually worldwide and is the only known pathogen to cause recurrent pandemics (e.g., 1918 – 20 million deaths, 1957 – 2 million deaths, 1968 – 1 million deaths). Ectromelia virus is a mouse model of smallpox, a disease that is believed to have killed more individuals than all other infectious diseases combined, and whose viral mechanisms have yet to be fully understood.

We propose that the use of inbred mice will serve to clarify disease mechanisms and identify modifiable targets that may impact the host response against viruses of major human concern. Specifically, we hypothesize that the use of the AcB/BcA panel of recombinant congenic strains of mice is an ideal platform to dissect complex traits divergent in A/J and C57Bl/6 inbred mouse strain owing to their potential resolution of QTL localization and availability of expression data for target tissues, provided that population structure and other confounding variables are taken into account.

OBJECTIVES

It is our intention to use modern genetic and statistical techniques within mouse models of these pathogens to facilitate the identification of actual mechanisms of susceptibility or resistance, with the ultimate goal of identifying new immunological mechanisms or targets for potential future therapies against human pathogens of global relevance. The following introductory section is divided into three parts describing: 1, The relevance and background of the pathogens of choice including their clinicopathological features and molecular pathogenesis; 2, The mouse model, including, mouse and genetic resources for

the forward genetic analysis; 3, The methods that can be used to identify disease- or trait-relevant genes using forward genetics.

1. THE AGENTS: RELEVANCE TO HUMAN POPULATION AND PATHOGENESIS

In our research project, we have selected smallpox and influenza mouse models of infection to identify host genes responsible for innate responses associated with protection from virus-induced disease.

Poxvirus and influenza pathogens are NIAID Emerging/Re-emerging and Biodefense Pathogens priority agents. They have different strategies to infect and replicate within their hosts, making them ideal agents to probe the immune system. In addition, although there are available vaccines against both viruses, they are not perfect, as there are safety concerns with respect to the smallpox vaccine and with the efficacy of the influenza virus vaccine.

SMALLPOX

Smallpox is a devastating endemic and epidemic disease of the past caused by variola virus. Variola virus infection had approximately ~30% mortality resulting in billions of deaths throughout human history.

Smallpox was described prior to 1000BC in China, mentioned in Indian Sanskrit texts [17], has coincided with the decline of the Roman Empire (i.e., the plague of Antioch AD 108, over 3 million deaths, [18]), and the decimation of the Aztecs and Incas when introduced by Spanish and Portuguese conquistadors in the early 1500s [17], [19]. During the 1775-1782 North American smallpox epidemic, the spread of virus affected civilians and devastated Native Americans tribes. The disease was responsible for more than 300 million deaths in the 20th century alone [20]. Smallpox affected all levels of society including prominent figures such as Egyptian pharaoh Ramses V [21], Wolfgang Mozart [22], George

Washington [23], Abraham Lincoln [24], and Joseph Stalin [25] among many others.

In modern times smallpox has been eradicated thanks to the work of Edward Jenner who succeeded vaccinating people against the disease, and ultimately, the WHO Global Smallpox Eradication Program. Jenner coined the term vaccination through the inoculation of cowpox (vacca is Latin for cow) to prevent individuals from contracting the more dangerous smallpox in the late 1700s, an era often referred to as the origin of modern immunology. The eradication of variola in 1980 [20] is considered one of the great successes of modern medicine. There is, however, fear that variola virus could be used as a biological weapon. The effect of such an attack would be devastating because most humans are no longer immune to the infection. Furthermore, related animal orthopoxviruses pose a risk as emerging viruses. For example, monkeypox virus, a zoonotic poxvirus endemic to Central and West Africa has caused outbreaks in the US Midwest from imported African pet rodents a few years ago [26]. It is unclear why some people succumbed in the past to smallpox and others did not. The discovery and characterization of host genes and mechanisms that impact the susceptibility to poxvirus infection may shed light into this issue.

CLINICOPATHOLOGICAL FEATURES

Smallpox is caused by infection with variola virus. There are four clinical varieties of Variola virus recognized by the WHO: modified, ordinary, flat, and hemorrhagic [27]. The modified-type is a mild form of the disease with low mortality usually occurring in vaccinated individuals. The majority of unvaccinated individuals manifest ordinary-type smallpox with mortality of ~30%. About 5% of individuals, mainly children, manifest flat-type smallpox, which results in flatter skin lesions and a mortality rate of close to 95%. An additional 3% of individuals, mainly adults and pregnant women, manifest the hemorrhagic form of the disease resulting in mucosal and skin hemorrhage and nearly 100% mortality [20 Jezek Z,

Ladnyi ID. Smallpox and Its Eradication. World Health Organisation: Geneva, 1988].

The non-infectious incubation period of smallpox generally ranges from 7-17 days. Infected individuals then start to experience fever, head and body aches, and occasionally vomiting. The fever can progress to a high level of 101 to 104 degrees Fahrenheit. This phase can last 2 – 4 days, during which the distinctive rash begins to appear, usually first in the mouth and tongue, followed by the face. The rash then develops into sores in the mouth that can erupt to spread the virus. This is when the disease is at its most contagious. By 24 hours after the rash first appears in the mouth it usually has spread to all parts of the body. The onset of the rash is generally associated with a dissipation of the fever. Three days after the onset of the rash raised bumps develop all over the body. By the fourth day of rash, these bumps then develop into fluid filled pustules with depressed centers. The development of pustules generally brings a second stage of fever. The pustules then begin to develop scabs. By two weeks after the onset of the rash, most of the pustules have scabbed. Most scabs will have fallen off by roughly three weeks after the onset of the rash. The individuals are still contagious until all of the scabs are gone, as the scabs themselves also contain infectious virus [28].

Mortality in smallpox was associated with higher levels of circulating virus [29], [30] and a de-regulated immune response, including the production of pro-inflammatory cytokines [31], [31] and sepsis [32], [32]. Virulence factors may contribute to ineffective host responses. Using vaccinia virus, investigators have identified two such factors: smallpox inhibitor of complement enzymes (SPICE), which downregulates the human complement system [33], and high-affinity chemokine-binding protein type II (CKBP-II), which binds chemokine receptors [34]. However, relatively little is known about the systemic pathology of smallpox. Several animal models of poxviruses have been useful in further elucidating the pathogenicity of and immune response against variola virus.

POXVIRUSES

Variola virus belongs to the family *Poxviridae*; these are large double stranded DNA viruses whose genomes vary from 130 to 360kbp in size. The family can be divided into two categories: viruses that are infectious in vertebrates, the Chordopoxvirinae, and those that infect insects, the Entomopoxvirinae. Within the Chordopoxvirinae subfamily is the genus *Orthopoxvirus*, which in addition to variola virus, contains several pathogens that are medically relevant and well studied as models of pathogenesis. These include monkeypox, a zoonotic viral disease principally in tropical regions, vaccinia virus, which was used as the vaccine to eradicate variola, and its close relative cowpox virus, which is maintained in rodents but can infect a broad range of mammals. The genus also contains ectromelia virus (ECTV) the agent of mousepox, the mouse's disease homolog of human smallpox that will be used in this thesis.

The genome of poxviruses is organized such that the genes involved in transcription and virus assembly are located in the central part of the genome, whereas the genes involved in immune evasion and other virus-host interactions are generally located towards the termini [35]. Nearly 90 of the over 150 genes encoded by orthopoxviruses are common to the sequenced members of the chordopoxvirus subfamily, with higher levels of conservation towards the central region of the genome [36]. Genes that are involved in the interaction with the host map to the terminal regions often exhibit lower sequence conservation and variation in copy number. Many of these genes are dispensable for growth in vitro and are referred to as virulence genes.

Poxviruses encode a plethora of virulence molecules that subvert, escape or modulate host responses [37]. An extensive review goes beyond the goal of this thesis. Suffice to say that these modulatory proteins can be classified at least in four groups based on their mode and site of action. Virotransducers interfere with the response to infection intracellularly including the induction of an antiviral state, oxidative burst and apoptosis. Virotealth molecules promote escape of immune

recognition of the infected cell by manipulating the expression of host molecules such as major histocompatibility class I and CD4 molecules. Other viral proteins act extracellularly. Viroreceptors are either expressed at the cell membrane or secreted. Their role is to bind host chemokines and cytokines to counteract their function. Virokines are viral mimics of host cytokines, chemokines and host factors. These factors contribute to host range and virulence of poxvirus.

CELLULAR STAGE

When naturally released from host cells, poxviruses are enveloped and generally described as having a rounded brick shape, covered in surface tubules. Both intracellular mature virions (IMVs) and extracellular enveloped virions (EEVs) are infectious.

The receptor used by poxviruses to access host cells is unknown, but it is generally accepted that the route of entry is through fibroblasts and macrophages in skin abrasions or the respiratory tract. 12 proteins have been identified as important in the fusion of virus and cell [38], [39], although the mechanism and function of these proteins, as well as the structure and function of the fusion complex is not known. Once the fusion between the virion and cell surface has occurred, the virus is enveloped with actin and ezrin projections and internalized [40]. The proteins and their functions required for the entrance of the virus are again unknown.

The virion core contains all the machinery needed to create the viral mRNA, a process that takes place in the cytoplasm [41]. The replication process is generally broken down into three phases, early, intermediate, and late. The early phase comprises the creation of mRNA needed for synthesis of viral DNA and those needed for the intermediate gene transcription. The intermediate genes encode for the genes in late gene expression. The late genes encode for the early transcription factors and other machinery that will be packaged in future virion cores for early replication. DNA replication begins several hours after

infection, and continues until approximately 10000 copies are made per infected cell [42].

After the creation of the mature virions, some are additionally wrapped in membranes [43]. Mature virions exit the cell in a mechanism similar to exocytosis whose precise mechanisms are not known. After exiting the cells at the primary site of infection, variola virus infects multiple organs consisting of the liver, spleen, bone marrow and others, before infecting the skin to produce the distinctive rash associated with the disease. Occasionally, the disease leads to fatality prior to the onset of the rash.

ANIMAL MODELS

There are no animal reservoirs for variola virus in nature. Prior to the adoption of WHO Resolution 52.10 that determined the destruction of existing variola stocks and severely restricted research with the pathogen by 2002, it was clear that most animal species cannot be experimentally infected. Jarling et al [32] exposed cynomolgus monkeys to variola strains using different routes of inoculation. It was not possible to establish infection by the aerosol route, which is infectious for humans, or by the intravenous route at low infectious doses. However, intravenous infection with high viral doses (10^9) recapitulated the classical lesion progression, coagulopathy and toxemia observed in full-blown smallpox.

High viral burdens in target tissues are associated with organ dysfunction and multisystem failure. Virus replication was mostly confined to monocytes and macrophages whereas lymphocytes did not share this high viral load despite severe lymphopenia. Evidence of coagulation cascade activation (increased D dimer levels) corroborated histologic evidence of hemorrhagic diathesis. Depletion of T cell-dependent areas of lymphoid tissues occurred, probably as a consequence of bystander apoptotic mechanisms initiated by infected macrophages. Cytokines, including IL-6 and IFN- γ , contribute to a cytokine storm formerly known as “toxemia.” Microarray analysis on peripheral blood samples

corroborated early up-regulation of IFN responsive genes (e.g., STAT1, STAT2, MX1, MX2, OAS) although IFN α protein was reduced in serum samples. There was also a marked reduction in NF κ B and TNF α signaling when compared to similar pathogenic diseases, suggesting that variola virus may produce products that interfere with this response. Due to research restrictions and the species specificity of variola virus, animal models with other poxviruses have been investigated to understand variola pathogenesis.

Although the reservoir of monkeypox is speculated to be rodents and squirrels [44], [44], both monkeys and humans can be zoonotically infected [45]. Monkeypox was first described as a disease of macaques in Denmark in 1958 [46]. The disease closely resembles smallpox in cynomolgus monkeys, and has been used to test vaccination and antiviral treatments [47], [48], [49], [50]. Comparisons of monkeypox and smallpox have shown that the high pathogenicity of smallpox may be due in part to host complement control of the inflammatory response and the robustness of immune responding in lymph nodes [46].

MOUSE MODELS

The mouse has been frequently used to study the host response against poxviruses. Ectromelia, cowpox and vaccinia viruses are capable of infecting rodents. Of these, ectromelia virus is a natural mouse pathogen and generally reproduces the pathogenic effects of smallpox infection in higher primates and humans [51], [52], [53], [54], [55], [56], [57].

Interestingly, ectromelia and variola viruses are thought to have both originated in rodents before co-evolving with their respective natural hosts. Murine ectromelia virus infection has helped identify that resistance to severe poxvirus infection is associated with type-1 cytokine responses (including TNF α , Interferon gamma, and interleukin 2) and natural killer cell and CD8 $^{+}$ cytotoxic T lymphocyte responses [58], [59], [60], [61].

POXVIRUS SPECIFIC IMMUNITY

The natural route of infection of smallpox is through aerosolized droplets of saliva. With ECTV, the main route is through abrasions in the skin – either through contact with contaminated surroundings or with another infected animal (e.g., through saliva from a bite). Infected Langerhan's cells (antigen presenting cells of the skin and mucosa) migrate from the superficial site of infection to the lymph node where they present viral antigens to naïve T and B cells. The immune response to ECTV is different from that of influenza virus, with the most integral cellular components being CD8+ T cells and NK cells. Strong IFN γ induction leading to cytolytic CD8+ T cell and NK cell based immunity are associated with resistance to mousepox infection [58], [62], [63].

Mice lacking CD8+ T cells have up to 100% mortality when challenged with ECTV. The antiviral effects of this cell population could be demonstrated even before traditional MHC class I activity was detected, suggesting multiple immune resistance mechanisms [61].

In both the CD8+ T cell and NK cell compartments, granule-mediated apoptosis is a key mechanism of resistance to ECTV. Antigen presentation via MHC class I and the T cell receptor results in the activation of CD8+ T cells. This in turn results in the synthesis of many cytokines, especially TNF α and IFN γ , and the release of perforin and granzyme. Mice lacking perforin show an LD₅₀ that is 6 logs lower as compared to unaffected mice upon challenge with ECTV [64]. Mice lacking either granzyme A or granzyme B also show increased mortality and spread of virus, with the susceptibility compounded when both were non-functional [53]. CD8+ T cells and NK cells can also impact resistance to ECTV through inducing apoptosis via TNF α associated Fas/Fas-L signaling.

IMMUNE EVASION

Approximately a quarter of the gene products of poxviruses are thought to be involved in immune evasion [65]. Poxviruses can inactivate a range of host immune signaling molecules including IFN γ [66], [67], TNF α [68], [69], IL-18 [70], and IL-1 [71], [72], [73] through the release of soluble secreted proteins. Generally, these bind host immune molecules directly to limit or abolish their signaling capabilities. Poxviruses are also able to downregulate the host immune response through modulating levels of caspase 1 [74], [75], [76], caspase 8 [77], [74] and granzyme B [78] through the action of various serine protease inhibitors. They can further interfere with apoptosis by altering the release of cytochrome C (a pivotal step in the apoptotic cascade) [79], competitively inhibit the action of PKR [80], and alter the activation state of innate immune cells during infection [81]. In addition, there are many other immune and inflammatory host systems targeted by poxvirus proteins that have yet to be fully elucidated [82], [83].

The large portion of the poxvirus genome devoted to immune evasion allows a multitude of options for evasion of the host immune system. Understanding new host genes involved in the complex interaction between poxviruses and their hosts could lead to increased understanding of immune responding, and ideally new therapeutic avenues for current pathogenic viral infections.

GENETIC CONTROL OF HOST RESPONSE

The mouse has served as a model to dissect the genetic basis of resistance to poxviruses. Laboratory mouse strains show divergent levels of resistance against ectromelia virus [84]. C57BL/6, C57BL/10 and 129 are relatively resistant as compared to susceptible A/J, DBA/2 and BALB/c mice. There have been at least four genetic loci associated with these genetic differences. Resistance to mousepox-1 (*Rmp1*) maps to the NK cell gene complex on chromosome 6 [85], with other mousepox resistance loci mapping near the complement gene C5 on

chromosome 2 (*Rmp2* [85]), the MHC on chromosome 17 (*Rmp3* [86]), and the selectin gene complex on chromosome 1 (*RMP4* [85]).

CD94, mapping to the NK cell gene complex on chromosome 6 has been shown to be essential to C57BL/6 resistance to ectromelia virus through the protective NK cell recognition of infected cells [87]. However, the genetic components of resistance to RMP2-4 remain to be identified. Additionally, there may be more loci associated with resistance to poxvirus infection that have not yet been identified. By further examining the genetic basis of resistance to poxvirus infection in mice, we hope to validate or identify new QTL, genes, or gene interactions that lead to resistance to poxvirus infection and increased knowledge of the immunological pathogenicity of this notorious human disease.

INFLUENZA

Influenza is a highly contagious infectious disease caused by influenza virus. Influenza-like symptoms have been reported as early as 412 BC by Hippocrates and others [88]. Since these early reports, there have been many documented flu like pandemics over the centuries. The Russian Flu of 1889, which caused an estimated 6 million deaths, was the first such outbreak to be thought to be conclusively caused by influenza [89].

Unlike orthopoxviruses, seasonal `viruses do not cause systemic disease, as they are usually confined to the respiratory tract. Despite the availability of antivirals and vaccines, seasonal influenza virus-associated deaths in the US per year have recently ranged from 5,000 to 52,000 [90]. The number of influenza virus-associated deaths usually are higher during pandemic years, with the extreme case of the 1918 influenza pandemic that caused more than 20 million deaths world-wide [91]. Although secondary bacterial infections and cardiovascular complications account for a significant proportion of these deaths, primary severe viral pneumonia is also common among the cases of severe disease. In fact, primary viral pneumonia has been the most common finding in

severe cases of human infections in the 2009 pandemic H1N1 virus [92], [93]. The host immune response is known to play a significant role in the pathogenesis of influenza infections. On the one hand, a weak immune response as seen in immunocompromised or elderly individuals can lead to unimpaired viral replication and severe disease [92]. On the other hand, an exacerbated response, associated with hypercytokinemia and enhanced infiltration of immune cells in the lung, can result in immunopathology and acute respiratory distress [94], [95], [96].

The incidence of recurring influenza pandemics is partly due to the broad host range and the segmented nature of the influenza genome. Influenza A viruses are thought to have derived from wild waterfowl, and have been isolated from pigs, birds, horses, seals, and whales, among other animals [97], [98]. The subtypes of the virus are divided based on two proteins on the surface of the virus: hemagglutinin (18 varieties) and neuraminidase (11 varieties) [99]. Concurrent infection of a host with different influenza strains can result in a reassortment event, or a swapping of viral segments between strains, causing the creation of new influenza species to which the human immune system is naïve. Together, the broad host range of influenza virus and its capacity for antigenic shift are the basis of novel influenza strains capable of epidemics in human populations.

CLINICOPATHOLOGICAL FEATURES

In seasonal influenza infections, viral shedding increases sharply between 0.5 and 1 day after challenge, and peaks on day 2 post-infection. Generally, respiratory symptoms peak by day 2 post infection and are negligible by day 7 post infection, while nasal and systemic symptoms (i.e., fever, headache, muscle ache and fatigue) peak by day 3 and dissipate by day 9 post infection [100].

Influenza strains that are highly pathogenic in humans, (e.g., H5N1), have been reported to have a longer incubation period than seasonal strains, lasting between 2-5 days. Lower respiratory symptoms are extremely common early in

the infection, while nasal and upper respiratory symptoms are not always present. More severe early systemic symptoms have been reported, including abdominal and pleuritic pain, and bleeding from the nose and gums [101]. Almost all patients develop pneumonia, with radiographic abnormalities (e.g., interstitial infiltrates and lobular consolidation) present by roughly 7 days after the onset of fever. The disease then progresses to acute respiratory distress syndrome with renal and occasional cardiac dysregulation. The fatality rate of highly pathogenic influenza has been reported at over 50% [101].

After influenza infection of the respiratory tract, alveolar macrophages and dendritic cells are among the first cells activated. Activated macrophages are strong producers of cytokines after infection. In line with this, depletion of macrophages during influenza infection was shown to result in reduced cytokine production (e.g., TNF α and IL-6). [102]. Once activated, dendritic cells mature and migrate to the lung associated lymph nodes and present peptide antigens to naïve T cells. These naïve T-cells then differentiate into effector cells. During this early response, dendritic cells are also activated and carry antigens from the pathogen to promote the expansion of antigen specific cells, CD4 and CD8 T cells, and B cells. Depletion of lung resident DCs was shown to reduce type 1 IFN and CD8+ cells after infection [103]. This adaptive, antigen-specific immune response is responsible for eliminating the pathogen and for establishing memory responses that prevents re-infections. Infected lung epithelial cells also play a role in initiating the immune response, and produce monocyte chemoattractant protein 1 (MCP-1), which in turn recruits monocytes that can differentiate into inflammatory macrophages and monocyte derived dendritic cells [104].

ANTIVIRALS AND VACCINES

Current strategies to increase resistance to influenza virus are mainly based on antiviral drugs or vaccines. Several antivirals are licensed for use with influenza: The M2 ion channel inhibitors (amantadine and rimantadine), and the NA inhibitors (zanamivir and oseltamivir, [105]). There are other newer drugs

currently in development (e.g., DAS181 cleaves sialic acid linked receptors that influenza virus use to attach to host cells, T-705 is a viral polymerase inhibitor, [106], [107], [108]). However, drugs targeting influenza are susceptible to the development of resistance through viral mutations [109], [110], [111], [112]. As it stands today, these antiviral drugs provide a useful addition to influenza vaccines, but vaccination remains the cornerstone of influenza prevention.

Vaccines against influenza have been available since the 1930s. [113]. Currently, they are designed to provide immunity against multiple types of influenza circulating in the human population (i.e., trivalent or quadrivalent, with influenza A/H1N1, H3N2 and 1 or 2 strains of Influenza B). The vaccines are generally made either using inactivated or live attenuated virus. Inactivated virus vaccines are created by growing the virus in fertilized chicken eggs. The virus is removed and purified from the eggs, inactivated and broken down with detergents and disinfectants, and then subsequently injected into individuals to stimulate an immune response. Live attenuated vaccines created through the process of adapting influenza virus to grow in cold environments of around 25°C, preventing the virus from replicating within the human body at 37°C. When the vaccine in use matches the circulating strain, the efficacy in healthy adults under 65 years of age is typically 80%-90% [113]. However, vaccines appear less effective in elderly individuals [114] and vaccine compositions have not always been good matches for circulating strains [115]. Even with the numerous vaccines and drugs available, influenza remains the vaccine-preventable disease that causes the highest mortality rate in the USA [116].

ORTHOMYXOVIRUSES

Influenza A virus is a member of the Orthomyxoviridae family, which is defined by viruses that have a negative-sense, single-stranded, and segmented RNA genome. Negative-sense indicates that the packaged genetic information is complementary to the mRNA. There are six different members of the

Orthomyxoviridae family: the *Influenza A*, *B*, and *C* viruses, *Thogotovirus*, *Isavirus*, and *Quaranilvirus*.

Influenza viruses are pleomorphic with round virions roughly 100nm and elongated virions occasionally over 300nm in diameter [117]. They are wrapped in a lipid membrane derived from the host cell. The influenza A genome consists of approximately 13.5 Kbp, sectioned into 8 negative sense RNA segments. All influenza A and B type strains, as well as infectious salmon anemia virus contain the 8 segment structure. Influenza C viruses contain 7 segments, while Thogoto virus and Quaranihil virus contain 6 [118].

The Influenza A virus genome encodes 11 proteins from the 8 negative sense segments. Their functions are briefly listed below. Three of the 11 proteins project from the viral surface. The hemagglutinin (HA) protein binds to host cells. The neuraminidase (NA) protein cleaves sialic acid to allow progeny virus to exit the host cell and prevent progeny virus particles from aggregating. The matrix protein 2 (M2) is a proton-selective ion channel that acidifies the viral core after entry into the host cell to initiate viral replication. The matrix protein 1 (M1) lies just beneath the envelope and has multiple functions including aiding the encapsidation of RNA-nucleoprotein cores into the membrane envelope to aid in viral assembly and budding.

The core of the virus particle is made up of the ribonucleoprotein complex, consisting of the polymerase basic 1 (PB1), polymerase basic 2 (PB2), polymerase acid (PA), and the nucleoprotein (NP) proteins that package and protect the viral genetic segments. There are also two non-structural proteins (NS1 and NS2). NS1 prevents polyadenylation of host mRNA to evade antiviral immune responses. NS2 (also called nuclear export protein) mediates the export of the RNP complex from the host nucleus after assembly [119]. Lastly, PB1-F2, which is expressed from alternate open reading frames within the PB1 gene, has been shown to increase apoptotic and inflammatory responses in host cells, increasing viral pathogenicity [120].

Interestingly, Influenza A viruses increase the coding capacity of their genomes via both splicing and use of alternative open reading frames. The M and NS genes each give rise to a spliced mRNA encoding the M2 and the NEP/ NS2 proteins, respectively [121].

CELLULAR STAGE

Influenza viruses bind to sialic acid receptors on epithelial cells in the host respiratory tract. HA proteins express preference for sialic acids with different linkages. Human viruses preferentially bind to N-acetylneuraminic acid attached to the penultimate galactose sugar by an $\alpha 2,6$ linkage (SA $\alpha 2,6$ Gal), whereas avian viruses mostly bind to sialic acid with an $\alpha 2,3$ linkage. This has bearing on disease severity, as in humans, $\alpha 2,6$ receptors are more prevalent in the upper respiratory tract, while $\alpha 2,3$ receptors are more prevalent in the alveoli in the lower respiratory tract. Interestingly mice, as well as birds, only express the $\alpha 2,3$ sialic acid receptor, which may be a main determinant in why many human influenza strains are not infectious in mice without adaptation [122].

The virion enters the host cell via endocytosis [123]. Once in the acidic environment of the endosome, the HA, which was cleaved, undergoes a conformational change and that enables it to interact with the endosomal membrane [123]. The M2 protein then acidifies the viral core, which dissociates the RNP from the M1 protein. Several altered HA proteins then form a pore through which the RNP, accessory proteins and RNA-dependent RNA polymerase are then transported to the nucleus of the cell through nuclear localization signals [124]. Transcription and polyadenylation of viral RNA then proceeds, dependent on host polymerase II activity.

Viral RNA are then either kept in the nucleus, or exported to the cytoplasm, where they are translated into proteins that are either folded, glycosylated, and secreted onto the cell surface through the Golgi apparatus, or transported back to the nucleus to be packaged with negative sense RNA to form new virion core

complexes. M1 and NEP are involved in this nuclear export of RNPs [125], [126]. The final packaging of progeny virions is not currently well understood. Influenza viruses then assemble and bud from the infected cells.

After exiting the cells at the primary site of infection, variola virus infects multiple organs consisting of the liver, spleen, bone marrow and others, before infecting the skin to produce the distinctive rash associated with the disease. Occasionally, the disease leads to fatality prior to the onset of the rash.

ANIMAL MODELS

Highly pathogenic H5N1 and 1918 H1N1 pandemic strains as well as seasonal strains have been studied in macaques. Macaques infected with highly pathogenic influenza show dysregulated cytokines and chemokines as compared to seasonal influenza [96], [127]. These studies seem to point to the overactive inflammatory and cell death response leading to increased fatality in the highly pathogenic strains of influenza.

Ferrets are often used as a model of influenza infection due to their similar lung physiology to humans. This includes a similar distribution of sialic acids throughout their respiratory tract [128], [129]. While the level of immunological reagents available is not as developed as for other animal models, ferret antiviral and proinflammatory mediators have also been shown to be dysregulated in response to highly pathogenic influenza infection [130], [131]. Ferrets have also been used as a model for vaccination studies [132], [133].

Mice have also been extensively used as a model of influenza infection. More background with regards to influenza infection in mice will be brought up in more detail later in the introduction.

HOST RESPONSE

Influenza virus enters the endosomal compartment as part of its replication cycle. Some virions are likely degraded in the endosome, exposing the viral RNA that is now recognized by TLR3 (due to the presence of dsRNA areas in the genomic RNA segments) and TLR7/8 [134]. Activation of TLR3 in macrophages and dendritic cells results in IRF3 and NF- κ B activation, and these transcription factors cooperate in the induction of IFN β . Activation of TLR7/MyD88 in plasmacytoid dendritic cells results in IRF7 and NF- κ B activation, and in secretion of high levels of IFN β and IFN α .

Viral RNA products are also recognized by RIG-I and the other two RLRs (MDA-5 and LGP2) in the cytoplasm. A critical molecular motif that is exploited by RIG-I to distinguish viral RNA from cellular RNA is the presence of a 5'-triphosphate [135], [136]. This is a common feature of the RNA genomes of most negative strand RNA viruses, including influenza virus. This process results in a conformational change that exposes the CARD domains, which become ubiquitinated by the action of E3 ligases such as TRIM25 [137]. Ubiquitin then promotes the interaction of RIG-I with the downstream adaptor MAVS and subsequent activation of IRF-3 and NF- κ B, leading to type I IFN production.

There are more than 300 genes induced by IFN, and not all of them have been characterized with respect to their antiviral activity [138]. This Type I IFN response can further elicit mechanisms of resistance against flu, such as PKR and OAS, viperin, and tetherin. Type I interferons generally use the JAK-STAT signaling pathway to initiate downstream responses. Typically, this leads to the nuclear translocation of activated STAT proteins, which bind to specific regions of DNA (interferon stimulated response elements) and begin transcription of antiviral genes such as 2'5'oligoadenylate synthetase (OAS) ribonuclease L (RNase L), and the myxovirus resistance proteins (Mx1, Mx2 in mice).

These genes then go on to perform immune functions. PKR binds dsRNA, reducing viral production. OAS can be activated by dsRNA in the cytosol, which

in turn activates RNase L. Activated RNase L degrades viral and cellular ssRNAs, inhibiting protein synthesis and viral growth. Mx proteins were identified as potent inhibitors of influenza virus replication and are thought to block essential viral components through mediating vesicle trafficking [139].

Another mechanism of intracellular recognition of influenza virus infection is through the nucleotide-binding oligomerization domain receptors (NOD-like receptors or NLRs). As a group, these receptors are sensors of PAMPs and cell stress and initiate downstream signaling pathways involved in inflammation, autophagy (i.e., the breakdown of improperly functioning cellular components) or cell death [140].

The NALP (i.e., NLRP1-NLRP14) and the IPAF (i.e., IPAF and NAIP) families are thought to participate in the formation of complexes called inflammasomes that interact with caspases (enzymes that break down proteins at specific target sites) to cleave several proinflammatory signaling molecules (e.g., IL1B and IL18) into their active forms. Mice lacking NLRP3 or NLRX1 have shown increased susceptibility and reduced immune responding to influenza virus infection [141], [142]. The development of inflammasomes has also been shown to play a central role in influenza virus immunity [143].

Neutralizing antibodies are generally directed against HA, NA or M2 surface proteins of influenza virus [144]. Depleting B-cells in mice has resulted in increased levels of mortality in response to influenza infection. In addition, passive transfer of influenza HA-specific antibodies to immune-compromised mice has shown protection from lethal IAV challenges [144].

The high genetic variability of Influenza A viruses, coupled with the action of NS1 can allow the virus to constantly be adapting, altering, or escaping the host immune response and antiviral drugs. Identifying new host genes involved in this complex interaction could be useful in ameliorating the effects of this and other pathogenic viral infections.

IMMUNE EVASION

The influenza virus non-structural protein (NS1) has evolved multiple mechanisms to interfere with the IFN response during viral infection. First, NS1 has an RNA binding domain that preferentially binds dsRNA [145]. As dsRNA molecules are known activators of RLR sensors and of PKR and OAS, NS1 competes with these cellular proteins for RNA binding, reducing their activation. Second, NS1 inhibits the export of cellular mRNAs to the cytoplasm [146], [147], [148]. NS1 also interferes with RIG-I functioning, both through direct interaction and by interfering with TRIM-25 [149]. NS1 can also block host gene expression by interfering with newly synthesized host pre-mRNAs [146], having a major impact on genes newly synthesized after influenza infection. NS1 has been shown to interfere with the function of (PKR) and (OAS) as well as various cytokines and chemokines after infection [150], [151], [152]. This retardation of the host immune response can last as long as 2 days post infection, even while the host lung is infected to high titers [153]. The deletion of NS1 from influenza A viruses results in infections with higher type 1 IFN responses and reduced mortality rates in animal models [154].

The high genetic variability of Influenza A viruses, coupled with the action of NS1 can allow the virus to constantly be adapting, altering, or escaping the host immune response and antiviral drugs. Identifying new host genes involved in this complex interaction could be useful in ameliorating the effects of this and other pathogenic viral infections.

GENETIC CONTROL OF HOST RESPONSE

Several inbred and wild-derived strains of mice are naturally resistant to influenza virus (e.g., AG2, SL/NiA, T9, and CAST/Ei). They share a Type I interferon mediated protection from disease through the action of Mx1 (i.e., myxovirus resistance 1, [155]). Studies have shown that this protective effect appears to be from an early block of viral replication [156]. Although Mx1 was discovered in the

mouse, its human equivalent, MXA, has been shown to provide defense against the introduction of avian influenza A viruses into the human population [156]. A similar path of discovery was found for interferon-inducible transmembrane 3 (IFITM3) associated resistance to influenza. Mice lacking *Ifitm3* show increased pathogenicity and viral pneumonias when challenged with influenza. This led to the discovery of an enriched minor *IFITM3* allele in hospitalized individuals with influenza that reduced influenza virus restriction in vitro [6]. These studies illustrate that novel genetic determinants of resistance to influenza found in the mouse can indeed translate to the human population.

The majority of inbred lines of mice, however, do not have a functional copy of *Mx1*. Even considering this fact, many phenotypic differences have been seen between *Mx1*-negative inbred mice in response to influenza infection, indicating that there may be more genetic loci remaining to be discovered.

Over the course of this thesis, there have been several mouse studies that have identified novel genetic correlates of resistance in the mouse. Three studies have been completed using the BXD panel of mice, created from resistant B6 and susceptible DBA/2 progenitors. Strong QTL on chromosomes 5 and 19, as well as on chromosomes 2, 16 and 17 were identified using A/PR/8/34 (PR8), and following mice for weight loss and survival. Interval mapping (single marker regression) was performed, followed by 2000 permutation tests for significance [157]. Previously identified expression QTL from basal lung tissue were then identified within the regions of linkage [158].

While using the highly pathogenic H5N1 strain, QTL for influenza resistance were identified on chromosomes 2, 7, 11, 15 and 17 [159]. A natural premature stop codon in the DBA/2 population was found in the hemolytic complement gene on chromosome 2. No causal genes were identified for the other QTL. Interestingly, the regions found on chromosomes 2 and 17 were non-overlapping in these two studies. Additional work with the BXD panel also identified a QTL controlling cytokine levels (e.g., TNFa, IFNa, and CCL2) on chromosome 6 [160].

A study using the pre-collaborative cross recombinant inbred panel of mice and the PR8 strain of influenza identified QTL on chromosomes 1, 7, 15, and 16 [161]. The study used an analytical technique called "bagging", which uses bootstrap sampling to generate multiple models from smaller randomly selected populations and then averages these models to get a model that should have increased overall predictive accuracy [162].

There is still a lot of work to be done to validate these QTL and to identify their underlying genetic mechanisms. The identification of more genes, pathways, and networks involved in influenza resistance should lead to new avenues of therapy, hopefully aiding the hundreds of thousands of individuals who succumb to complications of this infection annually worldwide.

2. THE MOUSE AS A MODEL ORGANISM

The *Mus* genus was established when mice and rats diverged from a common ancestor roughly 6 million years ago. The etymology of the word "mouse" is thought to be from the Sanskrit word *mush*, meaning to steal. The cohabitation of the mouse with humans has been speculated to originate as early as the dawn of human farming, where human settlements allowed the mouse access to shelter and food.

Roughly 10,000 years ago, *Mus musculus* diverged into separate populations. The *domesticus* to the west of India [163], the *musculus* to the north [164], [165], and the *castaneus* to the east [166]. The *domesticus* and *musculus* populations spread through western and eastern Europe respectively, with the *domesticus* following the western European population to colonize the Americas and Australia. In the 1800s, the mouse became a fancy pet among the British, Chinese and Japanese. Mice were traded to develop new breeds with varying coat colours and increasing docility to humans [166]. Eventually, crossing of *musculus* and *castaneus* in Japan formed the *Mus musculus molossinus* subspecies [167]. The breeding and sale of mice by Miss Abbie Lathrop and

others in the early 20th century brought many common mouse lines (e.g., C57BL/6) into the world of science. Currently, there are over 400 inbred lines of mice listed on the Mouse Genome Informatics Group at The Jackson Laboratory (a leading mouse supplier). These strains, however, have since been shown to contain limited and non-randomly distributed genetic diversity mainly of *Mus musculus domesticus* origins [168].

Despite the restricted genetic diversity that does not fully encompass the *Mus musculus* genus, these inbred lab strains have advantages for scientific research. They are generally docile and able to be housed in small quarters with a controlled simplified diet. Most notably, decades of study have also provided detailed information on their development, reproduction, physiology, behavior, and genetics. Their commercial supply further ensures that they are generally readily available from controlled specific pathogen free facilities, reducing potential confounding variables in their study.

MOUSE GENETIC RESOURCES FOR COMPLEX TRAIT ANALYSIS

Inbred lab strains of mice have been useful in understanding the underlying cause of many traits, including resistance to infection [169]. In particular, inbred strains of mice vary in their susceptibility to influenza and ectromelia viruses. The viral titers, inflammatory responses and mortality rates of the inbred strains have been shown to vary significantly across inbred strains for both influenza [170], [171], [172] and ectromelia [173], [174], [175], [176]. Interesting corollaries exist with the distributions of susceptibility of the inbred strains to both viruses. For example, in both cases the C57BL/6 strain of mice is found to be relatively resistant to infection whereas the A/J strain is found to be relatively susceptible.

GENETIC MAPPING IN F_2 AND N_2 CROSSES

Generally, to ascribe strain differences to a particular genetic region, inbred lines with divergent phenotypes of interest are crossed together to create progeny with genetic contributions from both parent lines. The progeny are tested, and their phenotypic and genetic profiles are compared to identify specific genetic regions that segregate with the original phenotype of interest. Two classic examples of this are the F_2 intercross and N_2 backcross. In an F_2 population, mice from two inbred lines are crossed together and their progeny (i.e., F_1 or first filial generation) are intercrossed to produce a population that ideally contains a distribution at each genetic marker used in the cross (i.e., 25% homozygous for parent 1, 50% heterozygous, 25% homozygous for parent 2 for non-sex chromosomes). In an N_2 backcross, the F_1 generation is crossed back to one of the original parental strains, ideally giving a distribution at each marker of 50% heterozygous and 50% homozygous for the parental strain used in the backcross. By comparing the genetic makeup and the phenotype of interest in a population of F_2 or N_2 mice, it is possible to identify markers in regions of the mouse genome that are linked to the studied phenotype. For these studies to be effective and comprehensive, the markers should be evenly spaced across all chromosomes, providing adequate coverage of the genome.

The inter-crossing of inbred strains has resulted in genetic regions of linkage with many varied phenotypes from basic physiology, to behavioural traits, to resistance to diseases, to many more categories, and has successfully lead to the identification of the genes that are directly responsible for controlling traits (for examples, see [177], [178], [179]).

MARKER ASSISTED BREEDING

Another strategy to genetically map a given phenotype is to use marker assisted breeding in multiple generations of breeding. Generally, additional markers (e.g. mutations or small genomic repeats) are identified within the region of genetic

linkage. The markers are used to produce subsequent mouse generations with smaller and smaller set of alleles from one parent isolated amongst alleles from the other parent. This marker-assisted breeding continues until only one gene or mutation is left within the isolated region from one parent, indicating the likely involvement of that gene. Much of the genetic basis for resistance to ectromelia has come from this type of study. For example, four regions of genetic linkage to mousepox resistance (RMP1-4, [85]) were identified through serial crosses of C57BL/6 and DBA/2 mice.

WILD-DERIVED MOUSE STRAINS

A strategy to increase the genetic diversity offered by the mouse model has been to incorporate mice obtained from the wild in different regions of the globe. This process has also been successful in crossing classic inbred laboratory mice with wild-derived mice, resulting in linkage from a broad sampling of the diversity within the entire *Mus musculus* genus (e.g., [180]). Wild-derived lines of mice originate from mice captured from the wild, and do not have the same constrained genetic diversity of the classic inbred strains. An example of this increased diversity is the presence of a functional copy of *Mx1* in wild-derived mice conferring increased resistance to influenza virus infection. This functional allele is absent in inbred strains of mice except for A2G. Interestingly, A2G was developed from illegitimate breeding of the A line of mice with an unknown non-inbred mouse, capturing the functional copy of *Mx1*, and thus increasing resistance to influenza virus [139].

RECOMBINANT INBRED STRAINS

Entire panels of mice have been created to aid in identifying and understanding new regions of genetic linkage. The panel most similar to the standard F_2 is the recombinant inbred panel of mice. They are created from strict brother sister matings of over 20 generations from original F_1 pairs. The end result is a set of

inbred lines of mice each representing a distinct F_2 genotype from an original set of parental strains. Increased control of extraneous environmental or infectious phenotypes can be achieved by testing multiple animals with identical genotypes. Additionally, having a line of mice instead of a single mouse per F_2 genotype can result in greater insights overall, through testing multiple related phenotypes and sharing resources (e.g., expression datasets in multiple tissues and streamlined analysis tools, see www.genenetwork.org). Many of the genetic regions of linkage for influenza resistance have come from the use of the BXD (i.e., C57BL/6J $\times$ DBA/2J) recombinant inbred panel.

THE COLLABORATIVE CROSS

The most ambitious mouse panel undertaken to date has been the Collaborative Cross. The Collaborative Cross is a multi-strain mouse panel that has attempted to take 8 parental strains of inbred mice, and breed them in a recombinant inbred fashion to produce many genetically diverse homozygous lines (i.e., by intercrossing the 4 F_1 mice, then intercrossing those G_1 progeny to obtain many genetically diverse G_2 families to be bred within their lines to homozygosity). A study testing several pre-Collaborative Cross lines (i.e., lines that had not completely bred to homozygosity) with the PR8 strain of influenza identified QTL on chromosomes 1, 7, 15, and 16 [161].

OTHER SPECIALIZED PANELS OF MICE

Other panels of mice also exist to aid in the identification of genetic linkage. Chromosome substitution strains (CSS) consist of a panel of 21 strains, corresponding to the 19 autosomes and two sex chromosomes, with each strain containing a single chromosome from the donor strain substituting for the corresponding chromosome of the host strain [181], [182]. By testing multiple mice from each strain for expression of a phenotype, it can be determined if the phenotype is linked to any particular chromosomes.

Genome tagged mice are created from backcrossing a donor strain to a host strain for multiple generations to achieve a set of congenic mouse lines. Each line is homozygous for the donor strain at a section of a chromosome and is homozygous for the host strain generally everywhere else across their genome. Several dedicated mouse lines exist for each chromosome, with the full panel together accounting for the entire set of donor autosomes [183]. When tested, these panels result in smaller areas of linkage than the CSS mice, but are still not ideal.

THE PROBLEM OF RESOLVING THE QTL POSITION

Generally all of these strategies encounter difficulties when trying to reduce a genetic linkage region to a specific gene or causal polymorphism. The sheer size of the linkage regions can often result in many compelling potential candidate genes and polymorphisms. Often, years of forward genetic studies are required to produce congenic mice with a small enough genomic interval size to prove the causal nature of a gene controlling a trait. A large proportion of studies never follow through to the identification of a causative polymorphism. This difficulty can be compounded in the case of complex traits (i.e., when multiple genes are involved in a phenotype), where multiple regions of linkage to related phenotypes can lie in close proximity to one another on the genome, making traditional forward genetics difficult.

Several panels of mice have been created in an attempt to reduce the size of genetic regions linked to phenotypes, aiding in the identification of potential causative genes and in the dissection of complex traits. Advanced intercross lines of mice are created by crossing two inbred mouse strains and then subsequently randomly intercrossing each generation until the genetic segments of each mouse in the panel have been sufficiently reduced in size (usually this is at least the F_4 generation). This breeding procedure can result in three to five fold reduction in the size of genetic segments linked to the phenotype of interest, but

does not benefit from the eventual creation of clonal mouse lines for further study [184].

Heterogeneous stock mice are created in a similar way to the advanced intercross lines, but begin with 8 progenitor strains instead of 2. The resulting heterogeneous population of mice again drastically reduces the size of genetic segments linked to phenotypes (theoretically as high as 30 fold [184], [185]) and has been used to map hundreds of traits [186]. The increases in complexity of these crosses, however, brings along confounding population structures and difficulties in statistical analyses [187].

RECOMBINANT CONGENIC STRAINS

Another breeding method used to reduce the region sizes of genetic linkage is the creation of recombinant congenic strains (RCS) by backcrossing F_1 mice to one of the two progenitor inbred strains for two generations. This results in many families of mice that contain roughly 12.5% genetic information from one parental strain randomly interspersed throughout the genetic background of the other. Each of these families is then bred to homozygosity through at least 14 generations of brother sister mating.

This strategy results in benefits similar to those of the recombinant inbred lines of mice. Mice with identical genetic makeup can be phenotyped across multiple time points and multiple traits, allowing a deeper initial study of traits and their interactions. Additionally, analysis techniques can be streamlined and databases related to multiple traits can be curated for future study (e.g., expression analyses of multiple tissues). RCS also benefit from a reduced size of genetic linkage compared to recombinant inbred panels of mice or traditional F_2 populations. This reduced interval size can simplify the identifying causal genes or polymorphisms, especially in the context of complex traits.

Our research group has developed a panel of RCS mice from C57BL6/J and A/J progenitors called the AcB/BcA panel [188], [189]. A/J and C57BL6/J show

variation in their physiology as well as their response to many naturally and experimentally induced pathologies [190], www.informatics.org). The AcB/BcA panel created from these progenitors has been used to study the genetic correlates of many infectious, inflammatory, metabolic and neoplastic phenotypes [191], [192], [193], [194], and has expression datasets for both lung and liver available for analysis [195].

Over the course of this thesis, we used the AcB/BcA panel of RCS mice to identify new causative genes underlying complex infectious disease phenotypes through the identification of small intervals of genetic linkage inherent in the panel and extracting compelling causal candidates through the analysis of gene expression within these intervals.

3. THE METHOD: FORWARD GENETICS IN MICE

The ultimate goal of forward genetics is linking a change in phenotype with a defined change in genotype. This first requires a phenotype to be associated with a given region of the mouse genome. This process is named quantitative trait locus (QTL) mapping and requires genetic markers to chart the mouse genome and statistical analysis to identify regions significantly associated with the phenotype. Once a QTL is validated several methods are used to extract genomic and functional information that can illuminate identification of candidate genes.

QTL IDENTIFICATION: GENETIC MARKERS AND MAPS

Originally, the mapping of traits by genetic linkage relied on traits themselves as genetic markers (e.g., physical characteristics of fruit flies such as body color, eye color wing shape, etc.). These phenotypes soon became difficult to expand upon, and biochemical phenotypes were used additionally (e.g., yeast nutrient

requirements for growth). In humans, it began possible to map traits using blood type and variation in human leukocyte antigens (the HLA system) [196].

Marker maps with a reasonable density were not created until the advent of restriction fragment length polymorphisms (RFLPs). Using various RFLPs one could identify polymorphisms between individuals whose DNA would be cleaved by the restriction enzyme and those that would not. Simple sequence length polymorphisms (SSLPs) were then identified as small areas of DNA that were repeated a different number of times in different individuals. Microsatellites, short SSLPs, were used extensively in genotyping studies for their even spacing across the genome. In the earliest uses of the AcB/BcA panel of RCS mice (the mouse panel used in the present thesis), the original marker map consisted of approximated 600 such microsatellite markers spaced evenly across the mouse genome. The advent of single nucleotide polymorphisms allowed the creation of dense marker maps.

More recently, the genetic architecture of the classic inbred strains of mice was identified in great detail [168]. The use of advanced genotyping techniques has increased the density of marker maps that can be used in genetic studies, resulting in smaller regions of linkage, and fewer candidate genes per region of linkage. Over the course of the present thesis, we upgrade the marker map of our mouse panel to include SNPs from the Mouse Diversity Array (Affymetrix), which contains information on over 600,000 SNPs polymorphic among known mouse strains, of which over 100,000 are polymorphic between our progenitor B6 and AJ strains.

QTL IDENTIFICATION: STATISTICAL ANALYSIS

When testing genetic markers for evidence of linkage, it is important to establish if the results that we find are meaningful. This is based on a statistical analysis of our findings, which requires that we formulate null and alternate hypotheses

reflecting two possible outcomes with our results. The first possible outcome, the null hypothesis, is that the distribution of our phenotype between our different genetic groups at a marker is no different from chance. In this case, there would be no evidence of linkage of genotype with phenotype. The second possible outcome, the alternative hypothesis, would be that the distribution of the results cannot be reasonably attributed to chance, and is thought to represent linkage between genotype and phenotype. To determine whether we are willing to say the distribution of our results differs from chance, statistical tests often attribute a p-value to the results. A p-value is a number between 0 and 1 that can be thought of as the proportion of times the results could be attributed to the null hypothesis. In our example, it would be the proportion of times the distribution of our phenotype among our different genetic groups could be attributed to chance. Historically, if the p-value is less than .05, indicating a 1 in 20 chance or less that the distribution of results could be attributed to the null hypothesis, then the result is deemed statistically significant.

Many early genetic studies of infectious diseases were set up simply, with all subjects belonging to one of two groups for a trait of interest (e.g., presence or absence of a disease). To identify whether any markers of a dichotomous (two outcome) genetic study has evidence of linkage, one could look at the number of individuals that have each genotype and each phenotype in the study for each marker. If the genotype and phenotype values that we observe for a marker are different from what we would expect from chance (i.e., individuals from one phenotypic group having an overrepresentation of one genetic group at a marker), then we can claim statistical significance. This type of analysis of observed versus expected values can be completed with a chi-square (which gives an approximated p-value) or Fisher's exact (which gives an exact p-value) test.

COMPLEX TRAITS ANALYSIS I: ANOVA

There are many interesting traits that can't be described in simple dichotomous terms (e.g., blood pressure, viral levels after infection). The genetic regions linked

with these continuous variables are called quantitative trait loci (QTL) and need to be statistically analyzed in a different way. The most common of the statistical tests looks at how spread out the phenotypic values are within and between each genetic group. If the amount of variance between genetic groups is much higher than within each group while taking into account the number of individuals in each group, the genotype at that marker can be said to be significantly linked to the phenotype of interest. This is called an analysis of variance (ANOVA), and is sometimes referred to as marker regression in genetic studies. For this test to be reliable, the observations have to be independent from one another (i.e., one observation cannot have influenced another), and the variance within each genetic group should be statistically equivalent and normally distributed (i.e., the values are symmetrical around the mean, with values further away from the mean being less likely than those close to the mean).

Sometimes, it is helpful to think of ANOVA in the form of a linear relationship between the phenotype and genotype. For example, the phenotype at a marker can be modeled as $Y = B_0 + B_1X + E$, where B_0 is the average phenotype, X is the effect of genotype with B_1 (sometimes just called the beta value) representing the strength of the effect of the genotype, and E is the variance in the dataset. It is possible to see that when the effect of genotype (B_1 , the beta value) on the phenotype is zero, then the phenotypic distribution of the entire population can be simply described with an average and a description of the variance from that average. It can be very useful to think of the beta value along with the p-value when looking at relationships between variables. An effect can be strongly significant (very low variance within groups), but not be very relevant (i.e., a low beta value, with the average phenotype of each genotype very similar to each other). Ideally, it is best to find results with high statistical significance, and high validity.

ANOVA is generally quite robust, and can account for the presence of covariates, or variables that may influence the association of genotype with a phenotype of interest (e.g., sex, experiment, environmental effects, etc). However, the strength

of the linkage results directly depends on how many markers are available to test. That is, if you don't have a marker that is close enough to a gene that is influencing a trait (or that otherwise adequately represents the genetic makeup of the population at that gene), then the ANOVA results may not represent the true strength of linkage in the dataset. ANOVA also requires that genotypes be present for all individuals when testing at a marker, so there is no way of interpreting and including missing data points. This can be an issue when dealing with incomplete data sets.

COMPLEX TRAITS ANALYSIS II: INTERVAL MAPPING

Interval mapping was created to address these issues [197]. Instead of only testing for linkage directly at each marker in a study, interval mapping uses a genetic map to infer possible linkage at areas between two markers. Generally, if an individual has identical genotypes at two adjacent markers, they will be highly likely to have the same genotype at all locations in between those markers. If the individual has different genotypes at two adjacent markers, then a probability can be assigned for having each different genotype at each spot between the two markers. This probability depends on the individual's genotype at each marker, the distance between each marker, and the number of genetic recombinations that are likely to have happened at each point between the markers. We can then work out the probability, or likelihood, that the hypothetical genetic distribution is linked to the phenotype for any location between markers; and the location with the highest probability, or the maximum likelihood, of linkage can be identified.

The statistical term to indicate the strength of this likelihood is called the logarithm of odds (LOD) score. It is the logarithm of the ratio of the probability of observing a set of data assuming that the genotype is linked to the phenotype (a positive beta value in the linear model), to the probability of observing a set of data assuming there is no linkage (a beta value of zero). Essentially, the LOD score describes how much more likely the data are to have originated in the

presence of linkage than in its absence. As an example, a LOD score of 3 means that the chance of finding a distribution of results, given that there is a QTL at the specified position being tested, is 1,000 times more likely than if there is no QTL.

There are several reasons why interval mapping is preferred over ANOVA to identify genetic linkage. Firstly, it is able to approximate the peak level of linkage across the entire genome and is not exclusively restricted to marker positions. This adds the benefit of giving a better estimate of the strength of linkage and an approximate position of the region of linkage, usually determined as the region within 1.5 LOD of the maximal linkage peak. Secondly, as interval mapping can interpret effects between genotyped markers, it is also able to deal with missing data better than ANOVA. Instead of removing individuals that do not have a genotype at a particular marker, you can simply use the next closest marker for the calculation. The disadvantage of interval mapping compared to ANOVA is increased complexity resulting in increased computation time. The time needed to complete a statistical test can become an important factor in choosing a test when running large complex sets of data, as can happen in genome wide studies.

COMPLEX TRAITS ANALYSIS III: MULTIPLE REGRESSION

Interval mapping and analysis of variance make use of a single-QTL model. Often, interesting phenotypes exist where the genetic component is known to involve more than one factor (e.g., in many infectious diseases, [5]). The multiple regression method has been developed to approximate interval mapping; in a genome wide analysis, multiple regression requires less computational complexity and provides greater power to detect the existence of multiple genetic regions of linkage. Further, it can separate linked QTL better, and can identify interactions between regions of linkage [198].

Multiple regression is an extension of ANOVA. While the ANOVA example we discussed above had one genotype effect included in its linear model (i.e., B_1X),

multiple regression includes the effects of genotypes at multiple markers (e.g., B_1X_1 , B_2X_2 , etc). Additional markers can be added into the model until the effect of adding another marker does not increase the prediction of the phenotype a significant amount. Generally, this test can be accomplished by looking at the difference between the model fitted phenotype and the actual phenotype for each individual in a dataset. The differences are squared (to remove the positive or negative direction of the difference) and then added up to get a residual sum of squares for a model. The closer predicted phenotypes are to the true values, the lower the residual sum of squares.

GENOME-WIDE SIGNIFICANCE

When testing for linkage at multiple locations across the genome, relying on single marker levels of significance can result in false positives (i.e., regions of linkage that appear significant, but are in fact due to chance). As the number of markers with different distribution of alleles among the population increases, it becomes more likely that the phenotypic distribution will result in a statistically significant result at least one marker.

To tease out these spurious (i.e., fake or false) results from true regions of linkage, we need to account for significance at a genome wide level. Many different p-value corrections have been proposed in an attempt to account for multiple testing and avoid type I errors, including the Bonferroni correction, Benjamini-Hochberg correction and permutation analysis. The Bonferroni correction multiplies the p-value for each test by the number of tests made. The Benjamini-Hochberg correction sorts and ranks the p-values of a test, giving the smallest value a rank of 1, the second smallest a rank of 2, and so on. The p-values are then multiplied by the number of tests performed and divided by their assigned rank to give the final adjusted p-value.

Permutations are a common statistical method used to adjust for the testing of multiple markers to attribute genome wide significance. To see how often

genome wide results could be attributed to chance, the phenotype is randomized with respect to the genotype for all individuals in the dataset, and the genome wide screen for linkage is performed on the randomized dataset. This process is repeated many times, while recording a test statistic for the maximum level of linkage for each randomization. This set of maximum linkage scores can be used to give an approximate cutoff for a genome wide significance (e.g., the level where 10 of 1000 chance permuted test statistics would be considered significant, equivalent to a p-value of 0.01) or an approximate p-value to an actual genome wide result (e.g., the number of permuted maximum test statistics that were higher than the actual test statistic, divided by the number of permutations).

POPULATION STRUCTURE

Additionally, genome wide studies can sometimes result in false-positive regions of linkage for reasons other than multiple testing. Indeed, it is possible that individuals with similar phenotypes may simply be more genetically related to one another. That is to say, they may have a genetic structure that is more “identical by descent”. This non-random population structure can introduce errors into tests of linkage, which assume that the observed phenotypes are independent from one another. If there is a factor controlling the phenotype (e.g., population structure) that is not included in a model, the model may not function well. This confounding effect of population structure is especially important to consider for certain inbred mouse populations. It is well known that population structure can cause illusory correlations between phenotypes and genotypes, leading to an elevated false-positive rate [199]. If the phenotype varies between subpopulations, then any marker allele that is in high frequency in an overrepresented subpopulation may be linked to the phenotype [200], [201].

FIXED AND RANDOM EFFECT MODELS

Several methods have been developed to correct for the confounding effects of population structure. Generally, these can be broken down into fixed or random effect models.

Fixed effect models consider only the data available in the sampled population as a whole and attempt to identify subpopulations from within the dataset. A solution derived from a fixed effect model was to use statistical tests that do not adhere to assumptions of independence to identify the population structure effect in a dataset and rescale the linkage results according to how confounded the population was found to be (i.e., lowering all significance levels assuming the identified population structure effects all markers equally). This method of “genomic control” trades off the statistical power available to find increasingly small effects in the dataset for the ability to identify the levels of rescaling needed, and therefore was not completely effective. Another solution was to assign individuals to subpopulations based on relatedness. Linkage was then examined within these subpopulations. This “structured association” was somewhat effective in controlling population structure but could not completely capture complex patterns of relatedness and used computationally intensive algorithms in the derivation of subpopulations, making it less than ideal with large or complex populations. Another way of attempting to identify subpopulations is to use principal component analyses (PCA) [202], [202]. This statistical method converts possibly correlated variables into a set of linearly uncorrelated variables called “principal components”. The use of PCA can result in the assignment of subpopulations with much less computation intensity than other methods.

Random effects models on the other hand treat the dataset as a randomly selected part of a larger whole population. Random effect models study changes in variance of the phenotype that that can be attributed to population variance. A mapping strategy based on population structure was developed by assuming the relatedness of individuals in a group was proportionally contributing to the

phenotypic variance in the trait of interest, with increasing relatedness of individuals resulting in increasing phenotypic similarity [203]. Methods of linkage analysis that use this “identity by descent” model of genetic inheritance originated in the study of multiple sibling pairs in humans [204]. These have been expanded to be able to accommodate larger and more complex populations and to be able to estimate the size of this background effect through the analysis of the relatedness of every possible pair of tested individuals at multiple points across the genome [205], [206], [207], [208], [209], [210]. These techniques are computationally intensive and do not fully control for the effect of population structure, making their use less than ideal in large or complex populations.

MIXED EFFECTS MODEL

More recently, a mixed-model strategy was introduced that incorporated both the techniques of structured association (using population subgroups) and covariance (using pair-by-pair relatedness to estimate phenotypic variation) to control for population structure [211]. This method produced very promising results, but was computationally intensive due in part to its use of structured association, and in part to its non-intuitive derivation of relatedness. The model of relatedness was non-linear (i.e., the formula contained exponents), which meant that solving for maximum likelihood took much longer than the regular models proposed above. In addition, the complex nature of the non-linear model meant that occasionally the maximum likelihood that was obtained through the model was underrepresented and thus not ideal.

By using computationally less intensive methods, including a simplified mathematically correct version of relatedness, and phylogenetic trees to control population structure, the practical application of mixed models to the mapping of complex genetic traits can be greatly increased. The Efficient Mixed Model Association method [212] has implemented these simplified techniques, and is able to obtain an accurate model of maximum likelihood while taking orders of magnitude less time than original mixed models while still maintaining excellent

control of population structure. This is the mixed-model that we use to derive the regions of genetic linkage in our current study.

QTL IDENTIFICATION: CANDIDATE GENES

After obtaining a valid QTL, it can be difficult to narrow down the causative genetic segment into a single gene or polymorphism. Traditional forward genetic studies employ a strategy of continually breeding new generations of congenic animals created from progenitors containing the region of linkage, and a background control, to systematically reduce the size of the genetic region associated with the phenotype. After many generations of breeding have passed, it is possible to isolate the allele from the causal polymorphism or gene in the control background, giving conclusive evidence of its involvement in the phenotype of interest, and also giving an excellent animal model to study the direct effects of the polymorphism or gene in future work. This technique can be extremely long, and is often not fully completed. Several bioinformatics approaches have been created in an attempt to help in this identification.

CANDIDATE GENE IDENTIFICATION: BIOINFORMATIC ANALYSIS

28 mouse strains have had their genomes fully sequenced by Sanger (www.sanger.ac.uk/resources/mouse/genomes). One strategy to reduce the potential candidates in a region of linkage is to consider only those genes in the interval that have polymorphisms between the progenitor strains. Occasionally, a gene can be identified that has a polymorphism leading to a likely change in the function of the protein encoded by the gene (e.g., the transcription of the gene ends early through a polymorphism inducing a premature stop codon). These genes are highly likely to have altered functions, and as such can be considered good candidates for involvement in underlying linkage. Software programs also exist that attempt to infer the damaging nature of the polymorphism between two strains (e.g., PolyPhen2, SIFT, PANTHER) when the resulting change is not as

drastic as a premature stop of transcription. These tools use conservation of amino acid residues in sequence alignments, evolutionary relationships and physical considerations of proteins to infer altered protein function. These tools can be helpful in the identification of candidate genes within a region of linkage, but are not foolproof (and generally they aren't all that reliable as of yet). Additionally, bioinformatics databases exist that list the function and location of expression of genes (KEGG - www.genome.jp/kegg/pathway.html, GeneCards - www.genecards.org, BioGPS - www.biogps.org, ImmGen - www.immgen.org). Identifying genes in a region of linkage that are in pathways associated with the phenotype of interest, and expressed in relevant cell types can also lead to the identification of strong candidate genes in a region of linkage.

Knowing the polymorphism content, location of expression, and the pathway involvement of all of the genes in a region of linkage can drastically reduce the list of potential candidate genes underlying regions of linkage, and can lead to increased identification of causal polymorphisms underlying QTL. However, when using established pre-existing panels of crossed mouse lines, it is possible that there may be genetic mutations that have arisen naturally in the lines of mice in the panel that differentiate them from the control mice housed elsewhere and used in the derivation of these bioinformatics datasets. *De novo* mutations do arise in panels of inbred mice, and can be the causal variant underlying regions of linkage in mouse studies (for example, see [10]). So while publicly available bioinformatics datasets can be useful, they may not adequately represent the full genetic diversity of all populations.

CANDIDATE GENE IDENTIFICATION: EXPRESSION ANALYSIS

Another strategy used to identify potential altered function among the genes in a region of linkage is to study their levels of expression directly. Gene expression profiling of model systems has been used to analyze the genetic basis of complex phenotypes [213], [214], [215]. This strategy takes advantage of the

heritability of gene expression profiles to find genetic variants linked to changes in gene expression [216], [217], [218], [219], [220].

EXPRESSION QTL

Expression QTL, or eQTL, reside on genomic regions that contribute to the variation of gene expression. Regulatory variants are known to affect clinical phenotypes, and have been proposed to contribute strongly to the evolution of complex traits [221], [222]. The identification of *cis*-regulatory polymorphisms (i.e., polymorphisms altering expression levels of genes in their direct vicinity) has been shown to be possible [223], [224], [225]. Methods for identifying strong *trans*-regulatory polymorphisms (i.e., that control expression levels of genes situated at far distances or on different chromosomes from the causative polymorphism) have proven more difficult.

Gene expression profiles have previously been analyzed in the RCS [195]. Using Affymetrix MU74A microarrays, RNA expression levels in lung tissue of 30 RCS mouse strains were obtained in duplicate [189]. Using a multiple regression model to measure the effect of both the genotype of the RCS line at each gene, and the majority background of the RCS strain (either B6 or A/J), evidence for over 1500 *cis*-regulatory polymorphisms was found. Unfortunately, these results were confounded by the presence of polymorphisms between the progenitor strains in the Affymetrix probe sets. In this thesis, we plan to redo the *cis*-regulatory analysis of gene expression in the RCS using custom software to remove the confounding polymorphisms.

The strategy of co-localizing expression and clinical QTL has proven successful in quickly identifying candidates in regions of linkage [225]. We plan on using this technique by identifying *cis*-regulated genes from RCS lung tissue in regions of linkage to influenza resistance or susceptibility. We will also identify *cis*-regulated genes from liver tissue in regions of linkage to ectromelia virus resistance or susceptibility. Ideally, this strategy will lead to the identification of genes

underpinning the resistance of certain inbred strains of mice to influenza and ectromelia viruses.

BRIDGING STATEMENT FROM CHAPTER 1 TO CHAPTER 2

This introduction has reviewed the two storied and impactful viruses studied in this thesis. We have described the host interactions with these viruses that result in productive viral infection, the host immunological functions that become activated to clear infection, and the host factors that can influence resistance and lower pathogenic damage associated with infection. We have also reviewed the history of mouse models in genetic studies, and a brief overview of the statistics that can be used to identify new genes or genetic regions associated with resistance to infections, with particular emphasis on the control of population structure and its potential influence on spurious results.

Over the course of the 4 papers to follow, we create a new marker map to be used with the recombinant congenic strains of mice and identify a different and successful statistical method to identify novel regions of linkage in our panel of inbred mouse strains. We verify a strong QTL associated with ectromelia virus infection and present an extremely strong candidate gene (*H2-T23*) through eQTL studies that is potentially responsible for the drastic differences of resistance seen between our susceptible and resistant strains of mice. We then verify the involvement of this gene through complementation studies. We also identify 2 QTL associated with influenza susceptibility (on chromosomes 2 and 17), and several QTL on chromosome 11 associated with resistance to influenza infection. We confirm the likelihood that the *hemolytic complement* (*C5*) gene on chromosome 2 is likely controlling this susceptibility locus, and identify a novel susceptibility gene *Pla2g7* on chromosome 17 through eQTL analyses. We verify the importance of *Pla2g7* on influenza resistance through knockout and congenic studies. We further go on to identify a *de novo* mutation in the *Nf1* gene as a candidate underlying resistance to influenza infection on chromosome 11. The impact and implications of these results are also discussed.

CHAPTER 2: A MAJOR LOCUS ON CHROMOSOME 17 CONFERS INNATE RESISTANCE TO LETHAL MOUSEPOX INFECTION

Gregory A. Boivin¹, Sanda Remakus², Luis Sigal², Robert Sladek^{1,3,4}, and Silvia M. Vidal¹

¹ Department of Human Genetics, McGill University, Montreal, Quebec, Canada

² Immune Cell Development and Host Defense Program, Research Institute of Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111, USA.

³ McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada

⁴ Department of Medicine, Experimental Medicine Division, McGill University, Montreal, Quebec, Canada

Running title: *H2-T23 Controls Resistance to Ectromelia Virus*

ABSTRACT

Although eradicated in 1980, smallpox caused significant damage to human populations, infecting upwards of 50 million people annually until the 1950s and the advent of world-wide vaccination and eradication programs. Mouse models of smallpox have been developed using ectromelia virus (ECTV). There is a large amount of data suggesting host genetic factors influence resistance of inbred strains of mice to ECTV. Previous QTL for ectromelia virus resistance have been identified on mouse chromosomes 1, 2, 6, and 17. The identification of causative genes underlying mouse QTL remains challenging. In the present study, we aimed to bypass the creation of large candidate gene lists by reducing QTL interval size with the use of a unique set of closely related mouse strains known as the AcB/BcA set of recombinant congenic strains (RCS). By combining cQTL from survival data and eQTL from liver tissue of RCS mice, we identified *H2-T23* as a compelling new candidate gene on chromosome 17.

INTRODUCTION

Poxviruses are large double stranded DNA viruses. The most notable virus of this family is variola virus, which causes smallpox. Although eradicated in 1980, it caused significant damage to human populations, killing ~30% of those infected, and leaving over three quarters of survivors with deep-pitted scars. Smallpox infected upwards of 50 million people annually until the 1950s and the advent of vaccination [226]. The dissection of the pathogenesis of this ancient scourge is still not fully understood, and remains relevant due to potential threats of variola virus associated bioterrorism and zoonotic poxvirus infection [227].

Host genetics are important in susceptibility to infection with many viruses [228], [229], [230], [231]. Due to the constraints of using the highly pathogenic human variola virus in the laboratory, murine models have been developed to dissect aspects of the disease. Both clinical and expression based complex disease phenotypes have been mapped using mouse models [9], [10], [11], [12]. Additionally, genes found using mouse models have proven important in human susceptibility to infection. This provides validation of the mouse model as relevant to human health, as well as providing evidence for the conservation of host immune defense [13].

Smallpox has been modeled in the mouse using Ectromelia Virus (ECTV) [232]. There is a large amount of data suggesting host genetic factors influence resistance of inbred strains of mice to ECTV [233]. For example, A/J, DB1 and BALB/c mice exhibit increased susceptibility to ECTV as compared to C57BL/6, C57BL/10, and 129 strains [234]. Susceptible mice generally show significantly increased viral titers in the spleen and liver, and succumb to infection within the first two weeks. Four mouse genetic regions linked to resistance to ECTV have been identified using mouse models. *Rmp1* maps to the NK gene complex on Mouse Chromosome 6 [63]. *Rmp2* maps to near the complement component C5 on mouse chromosome 2 [86]. *Rmp3* maps to the MHC on mouse chromosome

17 [86]. Finally *Rmp4* maps near the selectin gene complex on mouse chromosome 1 [235]. Although these loci provide a strong validation of the importance of host genetics in ECTV outcome, the genes underlying their altered resistance or susceptibility remain to be identified. Evidence has also been published showing that the NK-cell surface molecule CD94 is required for resistance to ectromelia virus, and that these molecules preferentially recognize the major histocompatibility complex MHC class 1b molecule Qa-1 [87].

The identification of causative genes underlying mouse QTL for complex phenotypes remains challenging because of many factors, including large interval sizes contained within most mouse QTL. In the present study, we aimed to bypass the creation of large candidate gene lists by reducing QTL interval size with the use of a unique set of closely related mouse strains known as the AcB/BcA set of recombinant congenic strains (RCS). The RCS were derived by inbreeding the second backcross generation of A/J and B6 progenitor inbred mice [189]. Each AcB/BcA strain inherits only a limited portion (1/8) of the genes from either A/J or B6. The genetic composition of these mice provides smaller candidate intervals than traditional F₂ crosses, and the expression profiles of the set of mice have been well characterized in multiple tissues. Expression QTL (eQTL) have been used in the RCS to identify genes that are differentially *cis*-regulated across the AcB/BcA set [236]. eQTL have also been used in similar fashion to identify novel genes involved in numerous traits, including body fat, lipid levels and bone density [12].

In this study, we used the AcB/BcA panel of RCS mice to dissect the differential susceptibility to ECTV between A/J and C57BL/6 inbred strains of mice. We identified a major locus on mouse chromosome 17, and identify *H2-T23* as a compelling new candidate by combining cQTL from survival data and eQTL from liver tissue of RCS mice.

MATERIALS AND METHODS

Animals and Ethics

All recombinant congenic, F₂, A/J, and C57BL/6 (B6) mice were maintained at McGill University animal facilities in compliance with the Canada Council on Animal Care as regulated by the McGill University Animal Care Committee. Recombinant congenic mice (N=417) of the AcB/BcA set were derived from two successive backcrosses (N₃) to either A/J (AcB) or B6 (BcA) parental mice [189]. In this breeding scheme, each AcB strain inherits a different set of discrete congenic segments including ~12.5% genes from the B6 genome and ~87.5% genes from A/J, and reciprocally for each BcA strain. Congenic mice were created from repeated backcrossing of BcA17 mice to B6 controls for over 7 consecutive generations. Progeny from reciprocal (BcA70 × B6) F₂ crosses, congenic (BcA70 × B6) crosses, and F₁ complementation (congenic × *H2-T23^{-/-}*) were generated with B6 mice ordered from Jackson Laboratories, and bred in McGill or in Fox Chase Cancer Center. For infections, all mice were transferred to the Fox Chase Cancer Center. The Institutional Animal Care and Use Committee of the Fox Chase Cancer Center approved all experimental animal protocols.

Infections and Viral Titer Analyses

For survival analysis, mice were infected subcutaneous with 10⁴ pfu ECTV Moscow Strain. Mice were observed daily for signs of disease (lethargy, ruffled hair, weight loss, skin rash, eye secretions) for at least 17 days post infection. Mice were euthanized if they show clinical signs or a condition score under 2 (mouse is becoming thin and bones are prominent).

Genotyping and qPCR analyses.

Genomic DNA was prepared from tails using Proteinase K and serial phenol/chloroform extractions followed by ethanol precipitation. The recombinant congenic strains of mice were genotyped using a map consisting of 1215

markers representative of 112,016 polymorphic SNPs from the Jackson Diversity Chip and Illumina MD linkage panel and 625 additional microsatellite markers genotyped in house [236]. The reciprocal (BcA70 × B6) F₂ crosses (N=249) were genotyped with a selection of 68 markers spanning the informative A/J segments of BcA70 mice using Sequenom iPLEX Gold technology. Genotyping was performed at the McGill University and Genome Quebec Innovation Centre (Montreal, Canada). qPCR was performed on uninfected tissue using *H2-T23* previously published primers on AB Biosystems StepOne Plus real time PCR machine (software version 2.2.2).

Statistical and bioinformatic analyses

All statistical analyses were performed using the freely available package R [237]. To test for significant cQTL in the RCS dataset, we used a mixed model statistical test designed to correct for genetic relatedness in mouse models known as Efficient Mixed Model Association [212]. Sex was used as a covariate in all analyses due to previously published reports showing influence of sex on ECTV outcomes.

To identify eQTL, we used the procedure outlined in Boivin et al 2011 [236]. Briefly, uninfected liver expression data from 44 microarrays (Affymetrix, 15 BcA strains, 7 AcB strains, two mice per strain) were analyzed using Custom CDFv13 to reorganize oligonucleotide probes based on the latest genome and transcriptome information. The reorganized probe sets were mapped to genomic locations based on the mouse genome (Build 37) and positioned with respect to individual RCS genotypes. 22,119 probes were mapped to sections of the genome for each AcB and BcA strain. We inferred A/J or B6 strain of origin for each gene based on the genotype of surrounding markers with a call rate of 96.9%. Expression values were normalized using RMA. To assign significance between differentially expressed genes and genotypes, ANOVA was conducted on a per gene basis using the linear model $\text{Expression} \sim \text{DSO} + \text{BG} + \text{DSO} \times \text{BG}$, where background (BG) and donor strain of origin (DSO) were coded as binary

phenotypes corresponding to A/J or B6. The cutoff for genome-wide significance was computed using the Benjamini Hochberg correction [238 Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289–300].

For survival analysis of F_2 mice, a Kaplan-Meyer based test of significance was conducted per marker using sex as a covariate. Genome wide significance was computed using bootstrapping (i.e., sampling the phenotype with replacement).

RESULTS

Distribution of Survival of AcB/BcA Mice

We identified marked differences between A/J (Mean survival time 7.8 +/- 0.2 SEM, n = 28) and B6 (Mean survival time 21.5 +/-0.5 SEM, n = 28) mice after ECTV challenge (Figure 1A, $p < 0.0001$).

Since A/J and B6 mouse strains showed marked differences in their response to infection, we used the AcB/BcA panel of recombinant congenic strains (RCS) to investigate the genetic control of survival upon infection with ECTV. Twenty-seven RCS were challenged with a dose of 10^4 pfu ECTV Moscow Strain and monitored for survival over 22 days. All AcB strains (i.e., those whose genetic makeup is mostly A/J in origin) have survival profiles similar to the A/J parental strain. For the BcA strains (i.e., those whose genetic makeup is mostly B6 in origin) only BcA69, BcA70 and BcA74 (and to a lesser extent BcA84 and BcA86) have survival profiles that deviate from the parental B6 strain (Figure 1B). The small number of deviant strains lends to the hypothesis of a single gene or locus controlling a large proportion of the variance in survival in our model.

Mapping cQTL using the AcB/BcA Mice

Genome-wide linkage analysis using an Efficient Mixed Model Association (EMMA) test using sex as a covariate [212] identified two clinical quantitative trait loci (cQTL) on chromosomes 7 (77.9 Mb to 80.3 Mb) and 17 (29.4 Mb to 41.2 Mb) associated with decreased survival after challenge with ECTV (Figure 2A). Sex was included as a covariate because of previous literature identifying possible sex effects in ECTV. To confirm these cQTL, we generated an F_2 mouse population from susceptible BcA70 and resistant B6 mice. The chromosome 7 and 17 QTL are both contained in the A/J segments of BcA70. F_2 mice were infected as above with a subcutaneous dose of 10^4 pfu ECTV and followed for 22 days for survival. Sex was once again used as a covariate in the analysis. In F_2 mice, chromosome 17 was the only segment to reach significance after

correction for genome-wide testing (Figure 2B). This single QTL is in line with the initial distribution of survival of RCS mice and confirms the locus on chromosome 17.

Primary candidate genes identified using eQTL

To identify primary candidate genes within our chromosome 17 cQTL, we performed linkage analysis using microarray expression data from uninfected liver tissue of 44 mice from 22 strains of the AcB/BcA set. We hypothesized that genes showing highly significant *cis*-regulated expression differences between A/J and B6 genotypes within cQTL would be primary candidates for further validation. Benjamini Hochberg correction of ANOVA derived p-values identified a set of 1172 probes representing 784 genes from the microarray analysis showing genome-wide significance (Figure 2C).

Cis-eQTL were co-localized with the chromosome 17 cQTL identified above (Figure 3A). *H2-T23* (36.1 Mb, $p=3 \times 10^{-14}$) is the most significant of the eQTL located within the cQTL on chromosome 17 (Figure 3A) and is within the minimal A/J genetic segment shared by BcA69, BcA70, and BcA74 (Figure 3B). To confirm the *H2-T23* eQTL, we performed a qPCR on RNA from uninfected tissue from resistant B6, as well as susceptible A/J and BcA70 mice (Figure 3C). *H2-T23* is the gene encoding Qa1, the natural ligand of CD94, which has been implicated in protective NK cell killing of cells infected with ECTV [87], making this a strong candidate for the control of decreased survival to ECTV in our model.

Validation of *H2-T23*

Only the B6 control RNA amplified a measurable amount of *H2-T23* by qPCR ($p=.02$, Figure 3C). This result is in line with an ancestral SNP previously identified in A/J mice producing a premature stop codon in the *H2-T23* gene. To confirm the role of *H2-T23* in susceptibility to ECTV, we created congenic mice from BcA70, isolating the A/J congenic segment (0-38.5Mb) surrounding the *H2-T23* region on chromosome 17 on a B6 background through the backcrossing of

congenic mice to B6 controls for over 7 consecutive generations of breeding (Figure 4). 15 male and 15 female congenic mice were tested for susceptibility against 5 B6 wild type male mice.

The mice were infected as before with 10^4 pfu ECTV and followed for 21 days post infection. As remarked in previous literature, there was a marked sex difference between male and female congenic mice, with ~40% of males and ~80% of females surviving infection, as compared to 100% of the resistant B6 controls ($p < 0.01$, Figure 5A). The survival of male congenic mice was remarkably similar to the survival of male F_2 mice containing the homozygous A/J allele at *Trim39* (the closest marker to *H2-T23*, Figure 4B), suggesting that the susceptibility in both cases is driven by the same mechanism. To identify whether the premature stop codon in *H2-T23* was the underlying cause of the susceptibility to the F_2 and congenic mice, we created F_1 mice from congenics and *H2-T23*^{-/-} mice. 16 male F_1 congenic $\times$ *H2-T23*^{-/-} mice were tested against wild type B6 controls with 10^4 pfu ECTV as above. The F_1 mice displayed significantly increased susceptibility, with roughly 30% of mice surviving to resolve infection as compared to resistant wild type controls ($p < 0.01$, Figure 5C). Additionally, the severity of susceptibility was in line with the homozygous congenic and the homozygous A/J mice at *Trim39* from the F_2 , and not in line with *Trim39* het mice (Figure 5B), providing compelling evidence for *H2-T23* as the gene of interest underlying the chromosome 17 QTL.

DISCUSSION

Study Rationale

The goal of this study was to identify potential candidate genes involved in the differential susceptibility of A/J and B6 mice to ECTV. We used the AcB/BcA set of RCS strains created from A/J and B6 founders to identify regions of chromosome 7 and 17 as potentially linked to ECTV susceptibility. The chromosome 17 cQTL (29.4 to 41.2 Mb) was confirmed in an F₂ cross. Using microarrays from livers of AcB/BcA mice, we identified *H2-T23* as an eQTL within our chromosome 17 cQTL, making it a strong candidate for further study.

A/J and B6 mice have been shown to harbour large differences in susceptibility upon infection to ECTV [239]. In the present study, we found A/J mice to have significantly reduced survival as compared to B6 mice after ECTV infection and have greater viral titers in liver and spleen tissue by day 6 post ECTV infection.

Mapping

To identify host genetic factors influencing ECTV outcome in B6 and A/J mice, we made use of a set of mouse strains derived from A/J and B6 mice, the AcB/BcA RCS. Following infection of the 29 RCS, only 3 (BcA69, BcA70, and BcA74) showed highly deviant phenotypes (i.e., these 3 strains are ~87.5% B6 genetically, but show survival times that are closer to A/J than B6). This strain distribution suggests the susceptibility in the AcB/BcA panel is controlled by a single genetic locus. This finding is against previous research showing that the genetic control of viral infections is complex involving multiple genes. However, it shows a distinct advantage of using the RCS to identify novel loci involved in the host response to viral infections. By limiting the genetic diversity in our population to two parental mouse strains, we reduce the number of potential genetic loci involved in our cross and have increased statistical power to identify the loci unique to A/J and B6 mice.

We statistically analyzed the genetic and phenotypic information from the RCS. Two genetic loci, on chromosomes 7 (78.0 Mb – 80.3 Mb) and 17 (29.4 Mb – 41.2 Mb), were significantly linked to survival after ECTV infection. Sex was used as a covariate in this analysis, but results were extremely similar if the covariate was omitted. We created an F₂ population from susceptible BcA70 and resistant B6 mice to confirm the presence of these loci. Only the chromosome 17 QTL was significantly linked to survival in this population. Again, sex was used as a covariate, and results looked very similar if the covariate was included in the analysis or not, with the only significant linkage in the F₂ data set being on chromosome 17. The lack of validation of the chromosome 7 locus is due to the chromosome 7 and 17 regions having the same genetic strain distributions. We did not have enough RCS strains to differentiate between the genetic and phenotypic distributions at the chromosome 7 and 17 loci. Because the chromosome 7 QTL was not reproduced in the F₂ population and had identical genetic and phenotypic distributions to the chromosome 17 locus, it was most likely a false positive identified in our initial analysis. Our results reiterate the importance of secondary F₂ crosses to validate QTL identified in primary screens, and confirm our chromosome 17 locus (29.4 Mb – 41.2 Mb) as linked to ECTV outcome.

Using eQTL to identify candidate genes underlying cQTL

To identify primary candidates in our chromosome 17 region, we performed a *cis*-eQTL analysis to identify differentially expressed genes from the livers of RCS mice. This strategy has been used to identify genes involved in complex diseases [12]. In the present paper, we identified *H2-T23* as a highly significant eQTL, and co-localized to our chromosome 17 cQTL, making it an excellent candidate for further study. We confirmed the eQTL by qPCR, and identified its most plausible cause as a premature stop codon identified previously in the gene through sequencing of the inbred lab mouse strains by Sanger. Interestingly, this is the second instance of eQTL/cQTL colocalization in the AcB/BcA mouse model identifying a premature stop codon as a likely compelling candidate for a complex

genetic disease. Previously, *hemolytic complement* (*C5*) was identified through this method as a candidate in influenza infection [236]. In this case, the *C5* gene did not have any expression in A/J progenitor mice, leading to the identification of the eQTL. The *C5* mutation has been validated for its role in susceptibility to influenza virus [159]. The present paper adds more positive evidence for the use of co-localization of eQTL and cQTL while searching for genes underlying differences in complex genetic diseases identified between inbred lines of mice.

From uninfected tissue, we can find intrinsic differences between our progenitor strains that determine host responses to infection. Because of this, the eQTL dataset presented here could prove useful for many future complex genetic studies involving A/J and B6 mice, and for naturally occurring or induced disease states.

Validating *H2-T23*

H2-T23 is the gene encoding Qa1^b in mice, a highly conserved MHC class 1-like molecule. CD94 forms receptors with NKG2A, -C, and -E on the surface of natural killer (NK) cells. It has been shown that Qa1^b can provide resistance to B6 mice against ECTV when interacting with CD-94-NKG2E on NK cells that are being simultaneously stimulated by NKG2D [18]. There is also evidence for Qa1 interacting with CD94-NKG2A molecules on CD-8 + T cells regulating cytotoxic T cell function [240]. For these reasons, we hypothesized that the natural premature stop codon present in *H2-T23* in A/J mice was the underlying cause of susceptibility in our present study. To confirm this, we created congenic mice consisting of the A/J fragment surrounding *H2-T23* on chromosome 17 (0-38.5Mb), on a fully B6 genetic background. Survival analysis of these congenic mice confirmed that the locus identified in our initial RCS and subsequent F₂ cross was linked to susceptibility. We then created F₁ mice from our congenics and *H2-T23*^{-/-} mice. As expected, we saw that the F₁ mice retained the susceptibility seen in our congenic and F₂ mice homozygous for the A/J allele at *H2-T23*. Importantly, mice heterozygous for *H2-T23* in our F₂ cross show a clear

intermediate susceptibility as compared to both congenic and F_2 mice homozygous for the A/J allele of $H2-T23$ (Figure 5). Additionally, previous reports have shown similar findings for $H2-T23^{+/-}$ mice, adding support to the complementation shown in our study (Figure 5C). With this data, we present the model shown in Figure 6.

It is important to note that due to the impact of the many potential confounding genes in the MHC on our results, we do not know exactly the full standalone effect of $H2-T23$ in our model. Our F_2 data presented do look like a continuous phenotype (Figure 5B), implying that our complementation study (breeding $H2-T23^{+/-} \times$ congenic mice) worked successfully in restoring a fully susceptible strain of mice (Figure 5C). However, it is always possible that there are impacts of multiple genes in our congenic fragment that are contributing to our results, so further experimentation in the future is needed to unequivocally confirm and validate $H2-T23$ in this context.

Interestingly, while sex did not influence the mapping of our loci, we did notice a strong sex effect in our congenic (Figure 4B) and F_2 mice (Data not shown), with males being more susceptible than females to ECTV infection. This reiterates the importance of including sex in any genetic analyses when dealing with ECTV.

The present study shows that the AcB/BcA mouse panel can be a powerful tool when dissecting complex genetic phenotypes. By combining smaller cQTL intervals with baseline eQTL information from tissue of interest, it is possible to quickly and efficiently identify potential candidate genes involved in susceptibility to globally relevant diseases and pathogens. Additionally, the current paper provides the groundwork for future studies outlining the interaction of Qa1^b with both NK cells and CD-8⁺ T cells. Future work must be done to understand how the polymorphisms between A/J and B6 in $H2-T23$ alter the interactions of infected cells with both NK and T-cell populations.

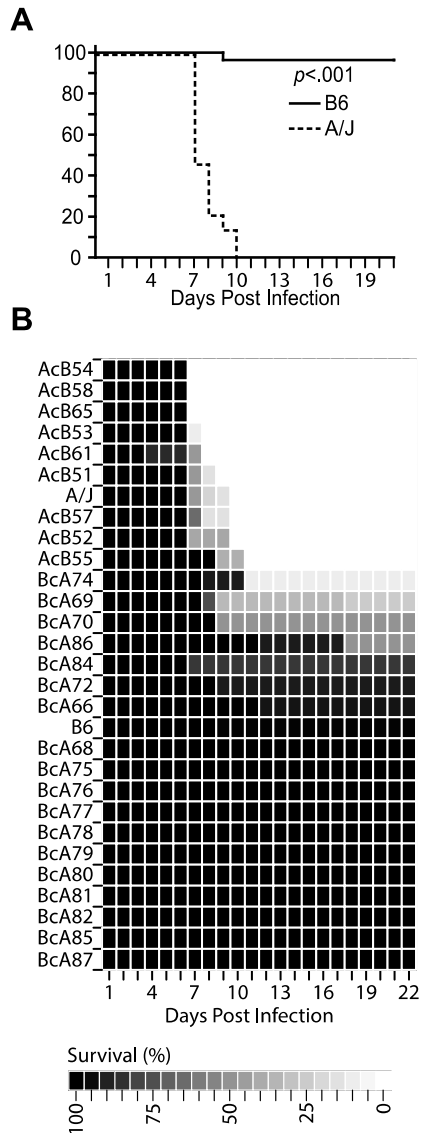


Figure 1: Inter-strain phenotypic differences in host response to ectromelia virus. A/J ($n = 28$) and B6 ($n = 28$) mice were infected with ECTV and followed for infection (A, $p < 0.001$, Kaplan-Meier). In total 417 RCS mice (7-12/strain) were screened for susceptibility with ectromelia virus. The survival distribution is indicative of a single or a small number of genetic loci (B).

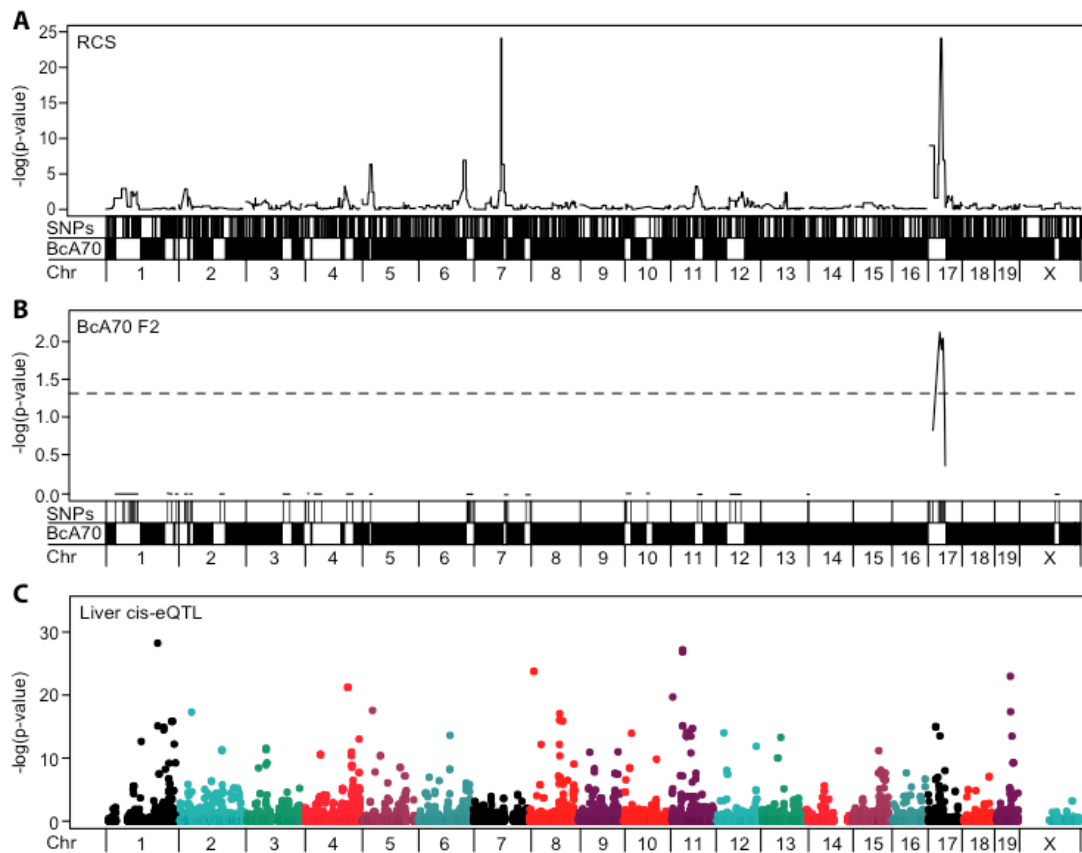


Figure 2: Control of survival to ECTV challenge maps to Chromosome 17.

Genome-wide linkage analysis was done in the RCS using a binary measure of susceptibility. The negative log genome wide p-values are shown (A). F₂ mice created from susceptible BcA70 and B6 control mice were tested, and validated the chromosome 17 QTL (B). Expression levels of genes were identified in uninfected liver tissue using microarrays. *Cis*-regulated expression differences were identified by ANOVA (C).

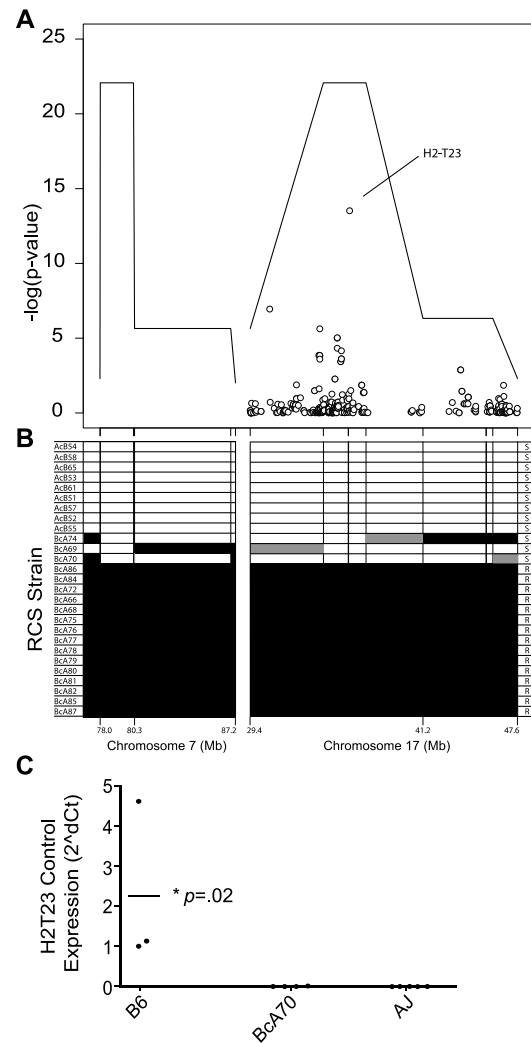


Figure 3: *H2-T23* candidate gene identified by overlapping clinical and expression QTL. The eQTL (negative log p-values) are shown overlaid onto the cQTL identified through the survival analysis of RCS mice (A). The distribution of RCS mice with A/J (white), B6 (Black) and unknown (grey) alleles underneath the peak markers are shown (B). The resistance or susceptibility of each strain is indicated by an R or S in the right side of the figure (B). The chromosome 7 peak that did not validate in the F₂ cross has the same genetic makeup across the RCS strains; therefore there is not enough power in the RCS data set to separate this peak from the validated chromosome 17 region in the initial RCS screen. The eQTL for *H2-T23* is confirmed via qPCR with group sizes of 3-5 male mice (C, ANOVA).

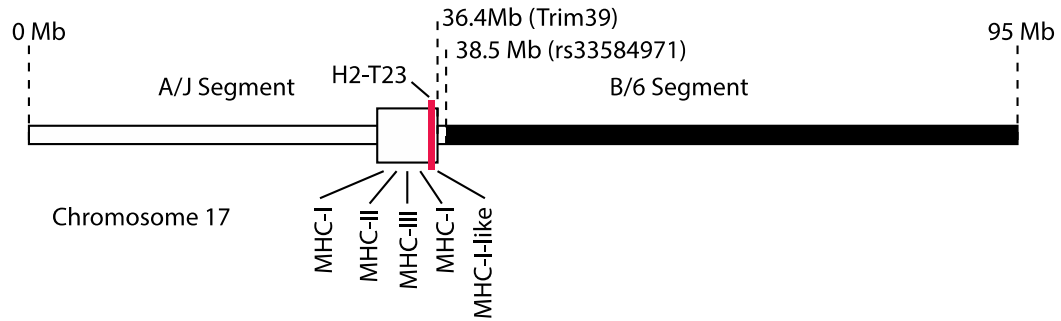


Figure 4: Haplotype structure of the congenic segment derived from the BcA70 mouse strain. The position of the break points of the H2 congenic mice are shown. The A/J region of chromosome 17 from BcA70 was isolated through 8 consecutive backcrosses of BcA70 to B6. The distal crossover is between Trim39 (36.4Mb) and rs33584971 (38.5Mb). *H2-T23*, the gene of interest, lies from ~36.16-36.17Mb.

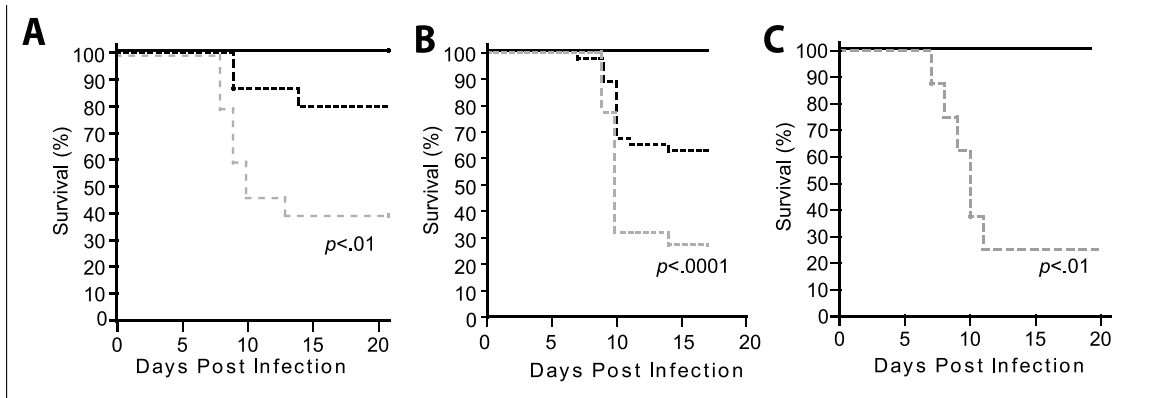


Figure 5: *H2-T23* candidate gene validation. The survival of male *H2* congenic male mice (grey dashed, n = 15), female congenic mice (black dashed, n = 15) and male B6 controls (black solid, n = 5) is shown (A). The distribution of survival of the male homozygous AJ (AA, grey dashed, n = 22), heterozygous (AB, black dashed, n = 46) and homozygous B6 (BB, black solid, n = 18) mice at the marker *Trim39* (the closest marker to *H2-T23*) in the BcA70 × B6 F₂ is shown (B). The survival of male F₁ mice created from H2 congenic mice crossed to *H2-T23* mice is shown (grey dashed, n = 16) compared to wild type male B6 controls (black solid, n = 5, Kaplan-Meier) (C).

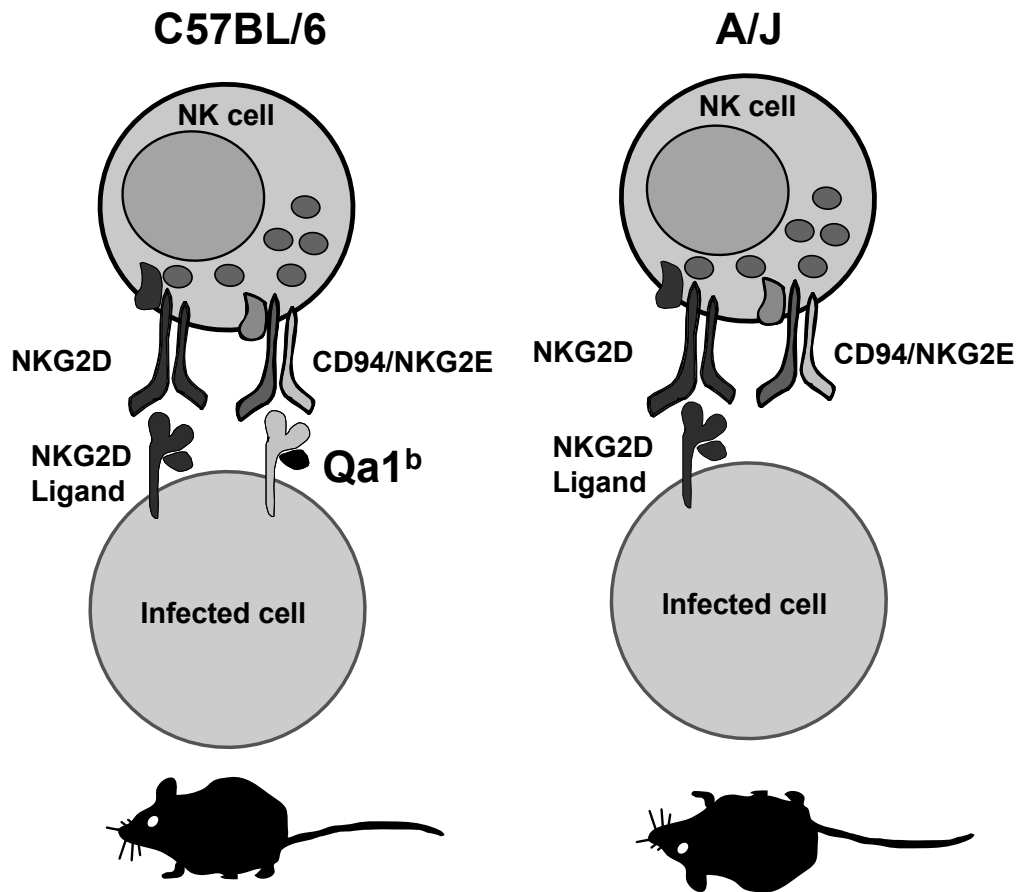


Figure 6: Model for Susceptibility to ECTV in the Mouse. In the absence of QA1^b, NK cells may not be able to recognize infected cells through CD94/NKG2E receptors, which leads to uncontrolled viral replication and eventual death.

The previous chapter has verified a strong QTL associated with ectromelia virus infection and presented an extremely strong candidate gene (*H2-T23*) through eQTL studies that is potentially responsible for the drastic differences of resistance seen between our susceptible and resistant strains of mice. We then went on to verify the likely involvement of this gene through complementation studies.

In the following chapters, we switch virus models from ectromelia virus to influenza virus. We go on to identify 2 QTL associated with influenza susceptibility (on chromosomes 2 and 17), and several QTL on chromosome 11 associated with resistance to influenza infection. We confirm the likelihood of the hemolytic complement (*C5*) gene on chromosome 2 as likely controlling this susceptibility locus, and identify a novel susceptibility gene *Pla2g7* on chromosome 17 through eQTL analyses. We verify the importance of *Pla2g7* on influenza resistance through knockout and congenic studies. We further go on to identify a de-novo mutation in the *Nf1* gene as a candidate underlying resistance to influenza infection on chromosome 11. The impact and implications of these results are then discussed.

CHAPTER 3: MAPPING OF CLINICAL AND EXPRESSION QUANTITATIVE TRAIT LOCI IN A SEX-DEPENDENT EFFECT OF HOST SUSCEPTIBILITY TO MOUSE- ADAPTED INFLUENZA H3N2/HK/1/68.

Gregory A. Boivin^{1,2}, Julien Pothlichet^{1,2}, Emil Skamene², Earl G. Brown^{3,4}, J. Concepción Loredó-Ostí⁵, Robert Sladek^{1,6,7}, and Silvia M. Vidal^{1,2,8}

¹ Department of Human Genetics and ²Centre for the Study of Host Resistance, McGill University, Montreal, Quebec, Canada

³ Department of Biochemistry, Microbiology and Immunology, and ⁴Emerging Pathogens Research Centre, Faculty of Medicine, University of Ottawa, Ontario, Canada

⁵ Department of Mathematics and Statistics, Memorial University of Newfoundland, St Johns, Newfoundland and Labrador, Canada

⁶ McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada

⁷ Department of Medicine, Experimental Medicine Division, McGill University, Montreal, Quebec, Canada

⁸ Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada

Running title: Male-specific locus controls mouse influenza susceptibility

ABSTRACT

Seasonal influenza outbreaks and recurrent influenza pandemics present major challenges to public health. By studying immunological responses to influenza in different host species, it may be possible to discover common mechanisms of susceptibility in response to various influenza strains. This could lead to novel therapeutic targets with wide clinical application. Using a mouse-adapted strain of influenza (A/HK/1/68-MA20 [H3N2]), we produced a mouse model of severe influenza that reproduces the hallmark high viral load and overexpression of cytokines associated with susceptibility to severe influenza in humans. We mapped genetic determinants of the host response using a panel of 29 closely related mouse strains (AcB/BcA panel of recombinant congenic strains) created from influenza-susceptible A/J and influenza-resistant C57BL/6J (B6) mice. Combined clinical quantitative trait loci (cQTL) and lung expression QTL (eQTL) mapping identified candidate genes for two sex-specific QTL on chromosomes 2 and 17. The former includes the previously described *Hc* gene, a deficit of which is associated with the susceptibility phenotype in females. The latter includes the phospholipase gene *Pla2g7* and *Tnfrsf21*, a member of the TNFR superfamily. Confirmation of the gene underlying the chromosome 17 QTL may reveal new strategies for influenza treatment.

INTRODUCTION

Influenza epidemics are the leading viral cause of mortality in the industrialized world [241], [242]. Because of the broad host range and segmented genome of the influenza virus, concurrent infection of the same host with different influenza strains can result in reassortment and adaptation, with the production of new, highly pathogenic viruses (such as avian H5N1, swine-origin 2009 H1N1) from which host populations are not protected by conventional vaccination or pre-existing Abs. Such reassortment and adaptation events are the classic basis of global influenza pandemics, which can kill millions of people worldwide [243], [244], [245]. Predicting the next pandemic influenza strain is difficult, rendering the production of appropriate pre-emptive vaccines unlikely. The emergence of highly virulent influenza strains, as well as resistance to antiviral compounds, underscores the need to better understand the determinants of disease severity. This, in turn, has great potential to reveal novel targets for preventive and therapeutic strategies against this globally relevant pathogen.

Host genetic factors play an important role in the onset, progression, and outcome of infection with many viruses [228], [229], [230], [231], [246]. With regard to influenza, many genetic risk factors have been documented, such as sex [247], blood type [248], [249], ethnicity [250], and familial history of susceptibility to influenza [251], [252]. However, linkage and association studies showed that the genetic component of susceptibility to infection is usually complex (i.e., involving more than one factor influencing susceptibility) [5]. Genetic heterogeneity, incomplete penetrance, sex, environment, and gene–gene interactions can all make complex traits difficult to analyze in human populations [253]. Genetic mapping with mouse models of infection has been used to successfully dissect both clinical and expression-based complex disease phenotypes [9], [10], [11], [12]. Importantly, genes originally found in mouse mapping studies were later shown to play key roles in human susceptibility to infection, providing evidence for the conservation of host immune defense [13].

The mouse has been useful in dissecting influenza pathogenesis. Murine models have been used to identify novel markers of pathogenicity [254], [255], genes involved in the host response to influenza [256], and vaccination strategies [257]. Patterns of early and sustained cytokine expression associated with susceptibility to influenza in higher primates and humans have also been characterized in mice [258], [127]. Further, studies in mice led to the identification of IFN-stimulated genes, such as *Mx1*, that encode innate antiviral proteins [259]. The *Mx1* gene product protects mice against lethal challenge with highly virulent influenza strains [260], [261]. More recently, a genetic-mapping study in female mice revealed five quantitative trait loci (QTL) associated with resistance to a highly pathogenic H5N1 virus [159]. *Hc*, located in one of these five QTL, was shown to decrease viral load and increase survival time after influenza infection [159].

Identifying the causative genes underlying mouse QTL for complex phenotypes remains challenging. A major reason is the difficulty of resolving QTL into sufficiently small genomic intervals to make gene identification possible. In the current study, we aimed to overcome this problem by using a unique set of closely related mouse strains in which the genetic composition and genetic expression are well characterized [195].

The AcB/BcA reciprocal series of recombinant congenic strains (RCS) was derived by inbreeding the second backcross generation of A/J and B6 progenitor inbred mice [189]. In this breeding scheme, each strain inherits only a limited portion (one eighth) of the genes from either A/J or B6. The relatively small size of the congenic segments fixed in individual RCS facilitates the search and testing of candidate genes. To identify primary candidates within genetic regions of linkage, we used lung microarray data to identify expression QTL (eQTL) for transcripts that are differentially *cis*-regulated across the AcB/BcA set. The colocalization of clinical QTL (cQTL) and eQTL in the RCS provides a strong tool

for the identification of new genes involved in the host response to influenza infection.

In the current study, we examined host genetic control of the mouse-adapted influenza strain A/HK/1/68-MA20 (H3N2) [262], which induces an overwhelming inflammatory response and death in A/J mice. Combined cQTL and lung eQTL mapping identified candidate genes for two sex-specific QTL on chromosomes 2 and 17.

MATERIALS AND METHODS

Animals and ethics

Inbred A/J, B6, 129X1/SvJ, and BALB/c mice were purchased from The Jackson Laboratory (Bar Harbor, ME). Recombinant congenic mice ($n = 398$) of the AcB/BcA set were derived from two successive backcrosses (N_3) to either A/J (AcB) or B6 (BcA) parental mice, as previously reported [189]. Briefly, in this breeding scheme, each AcB strain inherits a different set of discrete congenic segments, including ~12.5% genes from the B6 genome and ~87.5% genes from A/J, and reciprocally for each BcA strain. Progeny ($n = 177$) from reciprocal (BcA70 $\times$ B6) F_2 crosses were generated in-house. Experimental protocols were in accordance with the institutional guidelines of the Canadian Council on Animal Care. Mice were maintained at McGill University animal facilities in compliance with the Canada Council on Animal Care, as regulated by the McGill University Animal Care Committee.

Cell lines and virus strains

MDCK (ATCC CCL-34) and mouse fibroblast L-929 (ATCC CCL-1) cell lines were maintained, respectively, in DMEM and RPMI 1640 (Wisent) medium supplemented with penicillin (100 U/ml), streptomycin (100 μ g/ml), and FBS (10%). Influenza viruses A/PR8/34 (H1N1) and A/HK/1/68-MA20 (H3N2) were grown in 10-day-old embryonated hen's eggs. A/HK/1/68-MA20 is a mouse-adapted influenza virus strain derived from a H3N2 strain that was clinically isolated in Hong Kong during the 1968 pandemic, as previously reported [262]. Briefly, A/HK/1/68-MA20 was prepared by serially passing lung homogenates from mice intranasally infected with the wild-type clinical isolate for 20 rounds of infection. The experimental adaptation led to an increase in virulence representing a $10^{3.5}$ change in the LD_{50} and the selection of 11 mutations, several of which are in common with the virulent human H5N1 isolate A/HK/156/97 [262]. Titration of infectious virus was determined by a plaque assay, as described previously [263].

Infection of mice, definition of phenotype, and tissue collection

For each RCS, at least six male mice and six female mice were infected. A/J and B6 mice were included as controls in each infection. The F₂ mice were infected in groups of 46–65 mice, together with control mice from the parental strains. All RCS and F₂ mice were 84 ± 15 d old at the time of infection. For the susceptibility screen, mice were monitored daily for 2 wk following intranasal inoculation with a weight-adjusted dose of 10⁴ PFU influenza virus per 22 g body weight. At the experimental end point, we measured the semiquantitative lung-consolidation score, which was defined as: 0, no signs of hemorrhagic consolidation; 1, small or sporadic loci involving <10% of lung tissue; 2, 10–20% lung involvement; 3, 20–50% lung involvement; 4, 50–70% lung involvement; and 5, 70–100% lung involvement. Clinical signs (weight loss, labored breathing, lack of grooming, and low motility) were recorded daily. Mice presenting respiratory distress were humanely sacrificed. Human intervention at this decision point may minimally influence our definition of susceptibility and survival. For gene-expression analysis, male mice were inoculated with a dose of 10³ PFU influenza virus per 22 g body weight at 1, 3, or 7 d prior to sacrifice and tissue harvest.

For further phenotype analysis, male mice were euthanized at different time points postinfection by CO₂ in preparation for three different experiments. In the first experiment, the trachea was exposed and cannulated immediately after death. The lung lobes were inflated under constant pressure (25 cm H₂O) with 10% neutral buffered formalin (Sigma-Aldrich), excised, and embedded in paraffin for histopathology. In the second experiment, mouse tracheas were cannulated, and four volumes of 0.5-ml cold sterile saline were instilled through the tracheal cannula, withdrawn, and pooled to recover bronchoalveolar lavage fluids (BALF), as previously described [264]. In the third experiment, lungs were perfused with cold saline, flash frozen, and stored at –80°C until processing for viral titer determination or RNA isolation.

Histopathology, immunohistochemistry, and image analysis

Paraffin-embedded lung samples were sent to the Morphology Unit in the Department of Pathology and Laboratory Medicine, University of Ottawa. Tissue blocks were excised in 5- μ m-thick sections and stained with hematoxylin and eosin (H&E) for light microscopic examination. A board-certified pathologist blindly categorized each lung section according to a histopathological score based on the number and distribution of inflammatory cells within the lung parenchyma, as well as on noninflammatory changes, such as evidence of bronchiolar epithelial injury and repair. Sections were also analyzed using custom color-recognition software designed to identify white space (open airspaces); hematoxylin-associated nuclei and eosin-associated collagen, muscle fiber, and proteinaceous infiltrates. Briefly, each pixel in the histological image was filtered into one of three preset binary conditions (white, blue, and red) based on 8-bit color space. Other sections were stained with primary Ab, rabbit serum against A/HK/1/68-MA20 virus, and secondary Ab conjugated to Alexa Fluor 555 (Molecular Probes) to detect viral Ag, as described previously [265].

BALF cell content and determination of lactate dehydrogenase activity

BALF washes recovered from each mouse were centrifuged at 1000 rpm for 10 min. Supernatants were frozen and stored for protein analyses, whereas the cell pellets were resuspended in sterile PBS. The total number of leukocytes in BALF was counted with a hemocytometer, and the cells were spun onto frosted microscope slides (Fisher) at an approximate concentration of 3×10^5 cells/ml. The percentage of total leukocytes, consisting of eosinophils, lymphocytes, macrophages, and neutrophils, was blindly determined from counts of 200 cells in a cytospin sample stained with Diff-Quick (Dade Behring). The level of lactate dehydrogenase (LDH) activity in the BALF supernatant was determined by commercial colorimetric assay (Roche). The absorbance was read at 490 nm and reported as OD in BALF.

Quantitative PCR

Total RNA was extracted from whole lung using RNeasy columns (QIAGEN) and transcribed into cDNA using M-MLV reverse transcriptase with random hexamers (Invitrogen), according to the manufacturer's instructions. Real-time quantitative PCR (qPCR) was performed using Platinum SYBR Green SuperMix-UDG (Invitrogen) together with experimental or control primers. Experimental primers targeted a panel of cytokines and chemokines and were designed to span exon junctions with the help of primer3 (Table I). Target transcripts were normalized to the control housekeeping genes *Gapdh* or *Hprt*. Samples were run in duplicate, with five mice per condition. Reactions were performed using the PTC200 Thermal Cycler with Chromo4 Continuous Fluorescence Detector (MJ Research), and expression was analyzed using Opticon Monitor 3 software (MJ Research). Relative mRNA expression was calculated by subtracting the mean ΔCT of the control samples from the ΔCT of the infected samples ($\Delta\Delta CT$). The amount of target mRNA, normalized to the endogenous reference, was calculated as $2^{-\Delta(\Delta CT)}$.

Ex vivo macrophage preparation, infection, and analysis

Bone marrow-derived macrophages were isolated from male mice, as previously described [266]. Briefly, marrow was extracted by flushing with 2 ml RPMI 1640. RBCs were lysed, and the remaining cells were resuspended in complete media (RPMI 1640, 10% FBS, and 30% L-929 supernatant). Cells were maintained in culture for 5 d and supplemented with additional L-929 supernatant at days 2 and 4 prior to plating. Cells were seeded at 4×10^5 cells in 500 μ L Opti-MEM (Invitrogen) medium/well in 24-well plates and incubated overnight. After two washes with 200 μ L Opti-MEM, cells were infected with different amounts of influenza virus. Cell supernatants were harvested 18 h postinfection, and TNF- α , MCP-1, and keratinocyte chemoattractant (KC) were quantified by ELISA assays, per the manufacturer's instructions (R&D Systems). Maximal LDH activity was determined using a cytotoxicity assay (Roche) and used to quantify the relative

number of B6 and A/J cells. Results are expressed as cytokine concentration normalized to maximal LDH activity.

Genotyping

Genomic DNA was prepared from tail biopsies using Proteinase K and serial phenol/chloroform extractions, followed by ethanol precipitation, as previously described [189]. The RCS of mice were genotyped using 110,567 polymorphic markers from the Mouse Diversity Genotyping Array (Affymetrix), 1,449 polymorphic markers from the MD linkage panel (Illumina), and 625 microsatellite markers [189]. Polymorphic markers with 95% call rates and two or more informative strains (i.e., AcB strains with a B genotype or BcA strains with an A genotype) were used to identify haplotype blocks. Informative markers (the first and last markers of each haplotype block) were included in the final marker map. We selected a set of 1,215 markers that spanned the entire mouse genome (Build 37) and covered the relevant break points in the AcB/BcA panel of mice. The reciprocal (BcA70 × B6) F₂ crosses ($n = 177$) were genotyped with a selection of 68 markers spanning the informative A/J segments of BcA70 mice using Sequenom iPLEX Gold technology. Genotyping of the F₂ was performed at the McGill University and Genome Quebec Innovation Centre).

Statistical and bioinformatics analyses

All of the statistical analyses were performed using the freely available package R [237]. Because of the large background effect identified in the AcB/BcA panel of mice, we used a mixed-model statistical test designed to correct for genetic relatedness in mouse models, known as Efficient Mixed Model Association (EMMA, <http://mouse.cs.ucla.edu/emma/>). Because flexible time-to-events models are not implemented in EMMA, we transformed the continuous variable “survival,” representing time of sacrifice, into the binary variable “susceptibility” that indicated whether an animal was sacrificed during the experimental period. One locus mapping for the time-to-sacrifice phenotype for F₂ mice was analyzed

using a parametric survival regression at the markers. The significance was assessed via Bonferroni correction controlling the genome-wide type 1 error.

To identify eQTL, we reanalyzed lung-expression data previously obtained by Lee et al. [195] on MGU74Av2 microarrays (Affymetrix) for 54 mice (13 BcA, 12 AcB, B6 and A/J mice in duplicate) using Custom CDF version 12 to reorganize oligonucleotide probes based on the latest genome and transcriptome information. Probe sets that were deemed problematic (e.g., those containing strain-specific single nucleotide polymorphisms) were removed. The remaining probe sets were mapped to genomic locations based on the mouse genome (Build 37) and positioned with respect to individual RCS genotypes. Probes mapping to multiple locations were removed from the analysis. As a result, 11,294 of the initial 12,488 probes were mapped to sections of the genome for each AcB and BcA strain. We inferred an A/J or B6 strain of origin for each gene based on the genotype of surrounding markers with a call rate of 96.7%. If a gene was in between A/J and B6 markers for a given RCS, it was coded as NA. Expression values were normalized using the Robust Multiarray Analysis for Affymetrix gene chips.

To define the association between differentially expressed genes and genotypes, ANOVA was conducted on a per-gene basis using the linear model $\text{Expression} \sim \text{DSO} + \text{BG} + \text{DSO} * \text{BG}$, where background (BG) and donor strain of origin (DSO) were coded as binary phenotypes corresponding to A/J or B6. The cutoff for genome-wide significance was computed using the Benjamini–Hochberg correction [238]. Expression datasets have been deposited at National Center for Biotechnology Information/Gene Expression Omnibus under accession number GSE35888 (<http://www.ncbi.nlm.nih.gov/geo/>).

RESULTS

Interstrain phenotypic differences in host response to mouse-adapted H3N2 clinical isolate

In a preliminary screen to identify differential host responses to influenza virus infection, we used two well-characterized influenza strains: the prototype A/PR/8/34 (PR8) strain and A/HK/1/68-MA20 (in this article HK/68), a mouse-adapted influenza virus strain derived from a pandemic H3N2 clinical isolate from Hong Kong. HK/68 has 11 mutations that were selected through mouse adaptation, including several mutations in common with a highly pathogenic human H5N1 isolate [262]. We monitored survival in male mice over 14 d following infection with intranasal inoculation of serial dilutions of PR8 and HK/68 viruses of four inbred mouse strains, 129X1, A/J, B6, and BALB/c, which are known to have differences in their response to many pathogens. The mice presented marked interstrain differences in response to HK/68; A/J and 129X1 were relatively susceptible compared with B6 and BALB/c mice. These differences were not as pronounced in response to PR8 (data not shown). For this reason, we used the HK/68 strain for the remainder of the study.

More detailed clinicopathological observations of A/J and B6 mice infected with HK/68 revealed that, at a dose of 10^4 PFU, A/J mice had a significantly lower survival (12.5%, $n = 40$) compared with B6 mice (Fig. 1A; 93.9%, $n = 33$, $p < 0.0001$). In addition, A/J mice showed altered gross lung morphology compared with B6 mice, represented by severe hemorrhagic consolidation (Fig. 1B, $p < 0.0001$). Likewise, A/J mice had an earlier onset and more sustained clinical signs of infection, such as labored breathing, poor grooming, and low motility in their home cages (Fig. 1C). Finally, histological examination of H&E-stained pulmonary tissue showed bronchial epithelial necrosis and perivascular lymphocytic infiltrates in B6 mice, whereas A/J mice presented more extensive, disseminated interstitial inflammation and luminal protein infiltration (representative slides are shown in Fig. 1D). Ten images representing H&E

slides from each experimental condition (i.e., A/J uninfected, B6 uninfected, A/J infected, and B6 infected) were analyzed using color-recognition software. Consistent with the pathologist's findings, we identified a significant increase in eosin-associated collagen, muscle fiber, and proteinaceous infiltrates in infected A/J mice (Fig. 1E, $p < 0.001$). Thus, in terms of clinicopathological phenotypes, A/J mice are significantly more susceptible than B6 mice to influenza HK/68 infection.

Increased lung neutrophilia, cytotoxicity, and inflammation, but not viral load, correlate with susceptibility

To further characterize the differential susceptibility observed between A/J and B6 mice, we examined molecular, cellular, viral, and inflammatory phenotypes in BALF and lungs at days 1, 3, and 7 post-infection. LDH levels in BALF are a measure of cytotoxicity and cell lysis of damaged cells and serve as a marker of lung pathology [267], particularly during infection with highly pathogenic influenza strains [268]. By day 3 post-infection, both mouse strains showed evidence of LDH activity in BALF. However, by day 7, LDH levels were significantly higher in A/J mice (Fig. 2A, *left panel*, $p < 0.001$). The two strains had a comparable total number of infiltrating cells (Fig. 2A, *middle panel*, $p = 0.3317$). LDH activity correlated with recruitment of neutrophils in BALF, which were significantly more numerous in A/J mice by day 7 (Fig. 2A, *right panel*, $p < 0.001$). Additionally, A/J mice showed significantly fewer macrophages on day 3 ($p < 0.01$) and day 7 ($p < 0.05$) compared with B6 mice (Fig. 2A, *right panel*). No differences were found for eosinophil or lymphocyte populations (Fig. 2A, *right panel*). These results seem to indicate that the immune response in A/J animals is primarily neutrophil oriented by day 7 postinfection, as opposed to the predominantly macrophage-oriented immune response in resistant B6 mice.

A/J mice showed a small, but significant, increase in lung virus load compared with B6 mice (6.8×10^6 versus 6.2×10^6 PFU/g, respectively) on day 3 post-infection (Fig. 2B, $p < 0.001$). However, the kinetics of viral replication were

otherwise comparable, reaching a maximum of 10^7 PFU/g on day 1 and decreasing to $\sim 10^5$ PFU/g by day 7 post-infection in both strains. This observation was confirmed using immunohistochemistry analysis to detect HK/68, which showed that the localization of the virus in the lung was similar in A/J and B6 mice (Fig. 2C).

As determined by qPCR using oligonucleotide primers (Table I), A/J mice had significantly higher lung cytokine and chemokine expression levels compared with B6 mice mostly at day 3 post-infection (Fig. 2D, *Ifn- β* , $p = 0.0394$; *Il-10*, $p = 0.0032$; *Il-6*, $p = 0.0009$; *Tnf- α* , $p = 0.0003$; *Mcp-1*, $p < 0.0001$; *Kc*, $p = 0.0025$; *Il-9*, $p = 0.0158$). The primers used for the qPCR reactions are shown in Table I. Because these differentially expressed molecules are mainly produced by macrophages, and to explore whether over-reactive inflammation is a trait intrinsic to A/J macrophages, we derived bone marrow macrophages from A/J and B6 mice to confirm our results *ex vivo*. In agreement with the *in vivo* results, post-infection with influenza virus, macrophages from A/J mice had increased cytokine and chemokine expression compared with B6 mice at both the transcript (data not shown) and protein levels (Fig. 2E; MCP-1, $p < 0.0001$; TNF- α , $p < 0.001$; KC, $p < 0.001$), with no significant differences in viral replication (data not shown). Thus, susceptibility to HK/68 in A/J mice was characterized by increased expression of primarily macrophage-produced lung inflammatory mediators at day 3 post-infection, leading to neutrophilia and increased cytotoxic markers in BALF on day 7 post-infection.

Complex sex-specific genetic control of host response

Because the A/J and B6 mouse strains showed marked differences in their response to infection, we used the AcB/BcA panel of RCS to investigate the genetic control of susceptibility to the influenza HK/68 strain. Twenty-nine RCS were challenged with a dose of 10^4 PFU/22 g of influenza HK/68 and monitored for 14 d. Survival was a continuous variable identifying the day of sacrifice during the experimental period (days 1–14) (Fig. 3A). Susceptibility was a binary

variable identifying whether the animal was sacrificed during the experimental period that minimally changed the strain distribution from survival (Fig. 3B) and was used for the statistical analysis.

The distribution of the AcB/BcA panels of strains is illustrated in Fig. 3. We observed a significant effect of strain background (AcB versus BcA, $p < 0.0001$), as well as a significant effect of strain ($p < 0.0001$) and a significant sex $\times$ strain interaction for the BcA70 and BcA72 mouse strains (BcA70, $p < 0.05$; BcA72, $p < 0.05$). The continuous distribution of both survival time and susceptibility across the panel of 29 strains suggests complex genetic control of the susceptibility to influenza virus in A/J and B6 mice.

Increased susceptibility maps to sex-specific candidate cQTL

We assigned a binary measure to each mouse according to its survival during the 14-d infection period (0 = sacrificed, 1 = survived). Using this simplified measure of survival, we performed genome-wide linkage analysis and identified two cQTL on chromosomes 2 (24–38 Mb, $p < 0.01$) and 17 (37–48 Mb, $p < 0.01$) associated with increased susceptibility to influenza HK/68 infection (Fig. 4A). Because of the significant sex $\times$ strain interaction observed in our population, we also analyzed the females and males separately. Interestingly, the chromosome 2 locus reached genome-wide significance only in the female population (Fig. 4B), whereas the locus on chromosome 17 reached genome-wide significance only in males (Fig. 4C). Table 2 shows the cQTL identified while analyzing the full RCS cohort, the male RCS mice, and the female RCS mice.

To confirm our novel locus on chromosome 17, we generated an F₂ mouse population from susceptible BcA70 and resistant B6 mice. BcA70 males were significantly more susceptible than BcA70 females in our initial RCS screen. The informative A/J genomic segments of BcA70 cover the linked region of chromosome 17, making the strain ideal to use in the cross (Fig. 4C). F₂ mice were infected as above and followed for 14 d to obtain measures of time to

sacrifice or survival. Because the chromosome 17 locus is male specific and because a strain $\times$ sex interaction is present in the BcA70 F₂ mice ($p = 0.0097$), we analyzed the female and male F₂ mice separately. When looking at the survival of mice per genotype at each marker, we confirmed linkage of susceptibility with chromosomes 4 (Fig. 5A) and 17 (Fig. 5B), as noted by the lower survival of the AA genotype at these marker positions. Although the chromosome 4 region was detected in male and female mice, the chromosome 17 region was only detected when we analyzed the male F₂ mice. The distribution of effects confirms the complex nature of host resistance to influenza virus and more specifically confirms the novel male-specific locus on chromosome 17.

Primary candidate genes identified using eQTL

To identify primary candidate genes within our chromosome 2 and chromosome 17 cQTL, we performed linkage analysis using the microarray expression data from uninfected lung tissue of 54 mice from the AcB/BcA set originally published by Lee et al. in 2006 [195]. Compared with Lee et al.'s analysis, the current analysis had increased resolution and accuracy in detecting eQTL, because of increased density of genotypes in the RCS and the exclusion of known problem microarray probes. We hypothesized that genes showing highly significant *cis*-regulated expression differences between A/J and B6 genotypes within cQTL would be primary candidates for further validation. Benjamini–Hochberg correction of ANOVA-derived p values identified a set of 458 probes representing 473 genes from the microarray analysis that showed genome-wide significance (Fig. 4D).

Significant *cis*-eQTL were colocalized with the cQTL identified above (Fig. 6A). *Hc* (34.8 Mb, $p = 1.27 \times 10^{-3}$) was the most significant of five eQTL located within the female-specific cQTL on chromosome 2 (Fig. 6A, *left panel*). *Hc* was previously associated with influenza susceptibility, validating our candidate gene identification [159]. A/J mice are known to have a mutation that causes a stop

codon in *Hc*, resulting in A/J mice showing drastic reductions in the expression of transcripts for this gene compared with B6 mice. *Tnfrsf21* and *Pla2g7* were colocalized within the male-specific cQTL located on chromosome 17 (Fig. 6A, *right panel*). In both *cis*-eQTL on chromosome 17, the A/J allele had higher expression compared with the B6 allele. TNFRSF21 (43.2 Mb, $p = 1.18 \times 10^{-2}$) is a member of the TNFR superfamily and was shown to modulate Th cell activation [269]. *Pla2g7* (43.7 Mb, $p = 4.65 \times 10^{-5}$) was the third most significant eQTL identified in our microarray analysis and the most significant eQTL within the male-specific locus. PLA2G7 is a phospholipase that was shown to modulate proinflammatory responses [270], [271]. Fig. 6 has expanded views of the two loci, showing the colocalization of cQTL and eQTL (Fig. 6A), marker location (Fig. 6A), and SNP frequency in A/J and B6 mice (Fig. 6B). By identifying the most significant eQTL within our cQTL, we identified *Hc*, *Pla2g7*, and *Tnfrsf21* as strong candidate genes underlying susceptibility to influenza in our model.

Male-specific locus on chromosome 17 is associated with decreased survival and increased cytokine expression in RCS

To dissect male-specific susceptibility, we chose to focus on the BcA70 and BcA69 strains, which have the shortest survival time in males within the RCS panel. These two strains carry A/J alleles for *Pla2g7* and *Tnfrsf21* at the chromosome 17 cQTL but also contain A/J segments for *Hc*, which was previously reported to control influenza susceptibility [159], and for the *H2* cluster of immune genes, including *Tnf- α* , known to be involved in the pathology of acute influenza infection [272] (Fig. 7A). *H2* is located ~7 Mb proximal to the chromosome 17 cQTL mapped in this study. Therefore, to refine the localization of the chromosome 17 cQTL and its relative contribution to susceptibility, we examined the survival and expression profiles of a panel of cytokines and chemokines in a subset of informative RCS with ancestral recombination events segregating these three loci. Male mice from strain BcA74, which contains only the informative *H2* segment, and strain BcA76, which contains only the informative *Hc* segment, both showed increased survival time compared with

male mice from BcA70 and BcA69 strains (Fig. 7B). Thus, in this model, neither the *Hc* nor the *H2* region affected male mouse vulnerability to influenza infection. Further, at day 3 post-infection with influenza virus, we examined the panel cytokines previously shown to distinguish A and B6 progenitor strains (Fig. 2D) and found significant interstrain differences by ANOVA. *Pla2g7* expression segregated according to genotype and phenotype, with higher levels in susceptible mice. By contrast, *Tnfr21* expression was low in all RCS mice compared with the parental strain A/J. In addition, increased expression of *Ifn- β* , *Il-6*, *Tnf- α* , and *Kc* associated with susceptibility, as determined by the chromosome 17 cQTL (Fig. 7C).

DISCUSSION

We undertook this study to identify new candidate regions and genes involved in the host response to infection with a mouse-adapted H3N2 influenza strain. In our novel mouse model, influenza susceptibility was under complex genetic control, with sex significantly altering the severity of infection. We identified two sex-specific cQTL associated with influenza susceptibility. By comparing eQTL from expression profiles of uninfected lung tissue, we were able to identify primary candidate genes within these cQTL. A novel male-specific cQTL on chromosome 17 was confirmed by an F₂ cross and was associated with increased cytokine and chemokine expression and decreased survival in male RCS mice.

Functional dissection of susceptibility in our mouse model

After identifying differential susceptibility to influenza infection in terms of clinical signs of disease and survival between A/J and B6 mice, we tested additional aspects of the host response and immune system function to further characterize the differences. Both A/J and B6 mice had similarly high viral loads in lung tissue in early infection, and both strains were able to clear the virus at later time points. We noted a small, but significant, difference in the viral load of A/J and B6 mice only at day 3 postinfection. This similar pattern of viral load between resistant and susceptible strains of mice is not in line with previous research [273], [274], in which susceptible mouse strains had higher viral loads throughout infection.

In contrast to viral loads, A/J mice showed significantly increased expression of cytokines and chemokines, principally *Kc*, *Tnf- α* , and *Mcp-1*, compared with B6 mice. This differential inflammatory response was replicated in bone marrow-derived macrophages from A/J and B6 mice. Notably, compared with B6 mice, the increased inflammatory response in A/J mice was accompanied by greater infiltration of neutrophils and higher LDH levels in BALF. Because A/J mice

ultimately succumb to infection with HK/68, an exacerbated inflammatory response may be the primary determinant of susceptibility in this model.

Poor prognosis in highly pathogenic human influenza virus infection has been associated with high viral load and elevated cytokine levels [96]. Infecting primates with pathogenic influenza strains reproduces these features [274], [96], as does infection of mice with pathogenic H5N1 and 1918 influenza [95]. These results provide evidence for the conservation of host immune responses. Increased neutrophil chemoattractants, such as KC (the mouse homolog of human IL-8) [275], increased neutrophil populations [95], and increased apoptosis are all markers of severe disease in influenza [276]. The differential pathogenicity in A/J and B6 mice mirrors some of the hallmarks of influenza in humans, including the overactive cytokine response. As such, our model provides a strong platform for genetic dissection of the inflammatory response. Further, host susceptibility genes to H3N2 identified in this study may be relevant for infection outcome with many influenza strains.

Complex genetic control of influenza virus infection in mice

To investigate host genetic factors contributing to infection outcome, we screened a panel of closely related mouse strains created from A/J and B6 progenitors. We found a continuous distribution of susceptibility across these mouse strains, suggesting that infection outcome was under complex genetic control. This is in line with the complex nature of susceptibility usually seen in human infectious diseases [5]. Interestingly, we found a significant effect of sex in two of the RCS (male susceptibility in BcA70 and female susceptibility in BcA72) that was not present in progenitor strains. It is remarkable that the sexual dimorphism was only statistically significant in two strains, indicating a combined effect of background and sex in the susceptibility of the RCS mice to influenza. Sexual dimorphism with enhanced male susceptibility has been seen in model infections with many pathogens and is believed to reflect endocrine-immune interactions [277], [278]. For example, in the case of *Listeria monocytogenes*

infection, increased female resistance was linked to estrogen-mediated inhibition of proinflammatory caspase-12 expression [279]. On the contrary, in the case of both human and mice influenza infection, females can appear more susceptible than males [280]. Observational studies indicated that the incidence, severity, and fatality due to seasonal or pandemic influenza infections could be different between the sexes. However, they also often vary with age and geographic location, making the effect of sex alone difficult to interpret [281]. Further, it was shown that infection of 8-week-old mice with H1N1 virus induced low estradiol levels, which had proinflammatory effects and rendered females more susceptible than males [280], [282]. Notably, this sexual disparity was evident only when median infectious doses were used, whereas the effect of infections with low or high viral inoculi was sex independent [283]. Thus, comparisons are not straightforward with our experiments that have used older (12-week-old) mice and a different virus strain and dose. Thus, to draw more precise conclusions about the impact of sex on the outcome of infection with influenza virus, it will be important to disaggregate not only sex, but also age and dose effects, in future studies. Regardless, to our knowledge, our study is the first to identify sexually dimorphic genetic control of influenza infection, but the precise mechanisms determining sex differences in this model remain to be defined.

Identification of sexually dimorphic loci altering host response to influenza

We used the AcB/BcA panel of RCS to identify genetic regions linked to the host response to influenza infection. Overall, two regions, on chromosomes 2 and 17, reached genome-wide significance and were significantly linked to influenza outcome in the full AcB/BcA population. The chromosome 2 cQTL was the only locus to reach significance in the female population, while the cQTL on chromosomes 6 and 17 reached significance in the male population. It is notable that the chromosome 6 QTL, which was not highly evident in the full RCS screen, became very highly significant in the male population. This is most likely due to the loss of power associated with halving the number of individuals in the analysis. With the loss of information that came with the decrease in sample size,

it was impossible to differentiate between the chromosome 6 and chromosome 17 genetic regions in the male-only dataset. In this case, the chromosome 6 QTL is most likely a false positive that was inflated because it had a distribution of RCS genotypes similar to the chromosome 17 cQTL. The chromosome 17 cQTL was confirmed in an additional F₂ male population generated from susceptible BcA70 and resistant B6 mice. The F₂ cross also detected a weaker chromosome 4 locus with A strain alleles associated with increased susceptibility in male and female mice. Perhaps due to a smaller effect this locus was not detected in the RCS analysis. This result showcases the importance of secondary crosses to fully uncover the complexity of the genetic control of infectious traits.

Identifying primary candidate genes using eQTL

To identify primary candidates within the chromosome 2 and chromosome 17 loci, we looked for colocalization of our cQTL with that of lung eQTL identified in the same AcB/BcA mice. We hypothesized that highly significant *cis*-regulated expression differences between B6 and A/J mice may result from altered or absent gene function in mice containing the A/J or B6 genotype at a given locus. Integrating expression and clinical data previously identified candidate genes involved in complex disease [12]. Within the AcB/BcA mice, the most significant differentially regulated gene in the female-specific chromosome 2 locus was *Hc*. Otherwise immunocompetent mice [189], A/J animals share a common natural 2-bp deletion in a 5' exon of this gene, leading to a primary stop codon and complement deficiency [284]. Depletion of *Hc* was previously associated with increased susceptibility to influenza [159], [285] and infections by unrelated pathogens [286], [287], as well airway hyperresponsiveness [288]. Intact complement function was shown to have an adjunctive effect on host recovery from primary influenza [285]. The identification of this gene in our study supports our experimental approach, but it does not discount possible roles for other expression candidate genes in the chromosome 2 interval in influencing susceptibility in our model. Previous mouse work implicating *Hc* in influenza susceptibility was done only in female mice, which is in line with our mapping

results that suggest the effect is stronger in females than in males. Interestingly, complement C5 (encoded by *Hc*) activity may be altered by female sex hormones [280] providing a potential explanation for the sexual dimorphism that we observed.

Similarly, of the significant eQTL within the male-specific chromosome 17 locus, *Pla2g7* is the most compelling candidate. The protein encoded by *Pla2g7* is a multifunctional enzyme that is implicated in many disease processes, including cardiovascular disease, asthma, and macular degeneration [270], [271], [281], [289]. Two contrasting effects on inflammation have been attributed to PLA2G7: it can hydrolyze oxidized phospholipids to produce lysophosphatidylcholine and free oxidized fatty acids, leading to increased inflammation [281]; and can also catalyze the degradation of platelet-activating factor into inactive components, leading to reduced inflammation [284], [285], [286].

Recently, high levels of PLA2G7 activity were implicated in a swine model of coronary artery disease where it appeared to increase inflammation in the necrotic cores of atherosclerotic plaques [283]. In this animal model, pharmacological inhibition of PLA2G7 reduced inflammation and disease severity by reducing levels of proinflammatory lysophosphatidylcholine [287]. In our AcB/BcA model of influenza infection, the chromosome 17 locus and the A/J *Pla2g7* genotype were associated with high *Pla2g7* mRNA levels, increased inflammation (including increased *KC* expression), and decreased survival. Although our results provide an intriguing new candidate gene, more research is needed to dissect the exact genetic contributors of susceptibility within the sex-specific region of linkage on chromosome 17. As with *Hc*, previous work showed that female sex hormones have an effect on mRNA levels of *Pla2g7* [290], [291], providing a potential explanation for the sexual dimorphism observed at this locus.

Another interesting candidate within the chromosome 17 cQTL stemming from eQTL analysis is *Tnfrsf21* (also known as death receptor 6), a member of the TNF receptor superfamily. TNFRS21 activation is associated with NF- κ B and

JNK activation and apoptosis [292]. The gene was associated with increased lung inflammation in a murine model of asthma [269]. However, although we observed differential expression in the parental strains, we failed to observe differential regulation of the gene in RCS upon influenza infection, suggesting that, at least at the level of mRNA expression, TNFRS21 does not influence the course of infection.

We confirmed a previously identified cQTL and identified a novel sex-specific cQTL linked to aberrant host responses against influenza infection in mice. Expression mapping was used to identify primary candidates within these regions of linkage. *Hc*, located within the previously confirmed cQTL on chromosome 2, was depleted in susceptible mice, leading to reduced complement function. *Pla2g7* and *Tnfrsf21*, located in the male-specific cQTL on chromosome 17, were overexpressed in susceptible mice and associated with increased pathogenic levels of proinflammatory cytokines. By using RCS mice, we could detect genes with small effects while maintaining high genetic resolution, thus anchoring expression results in small intervals. By searching for genes with highly significant *cis*-regulatory differences within regions of linkage to influenza susceptibility, we were able to more efficiently identify compelling candidate genes involved in the host response to this globally relevant pathogen. This study may be used as a reference for future studies attempting to incorporate eQTL with traditional forward genetics for efficient gene discovery within the context of infection.

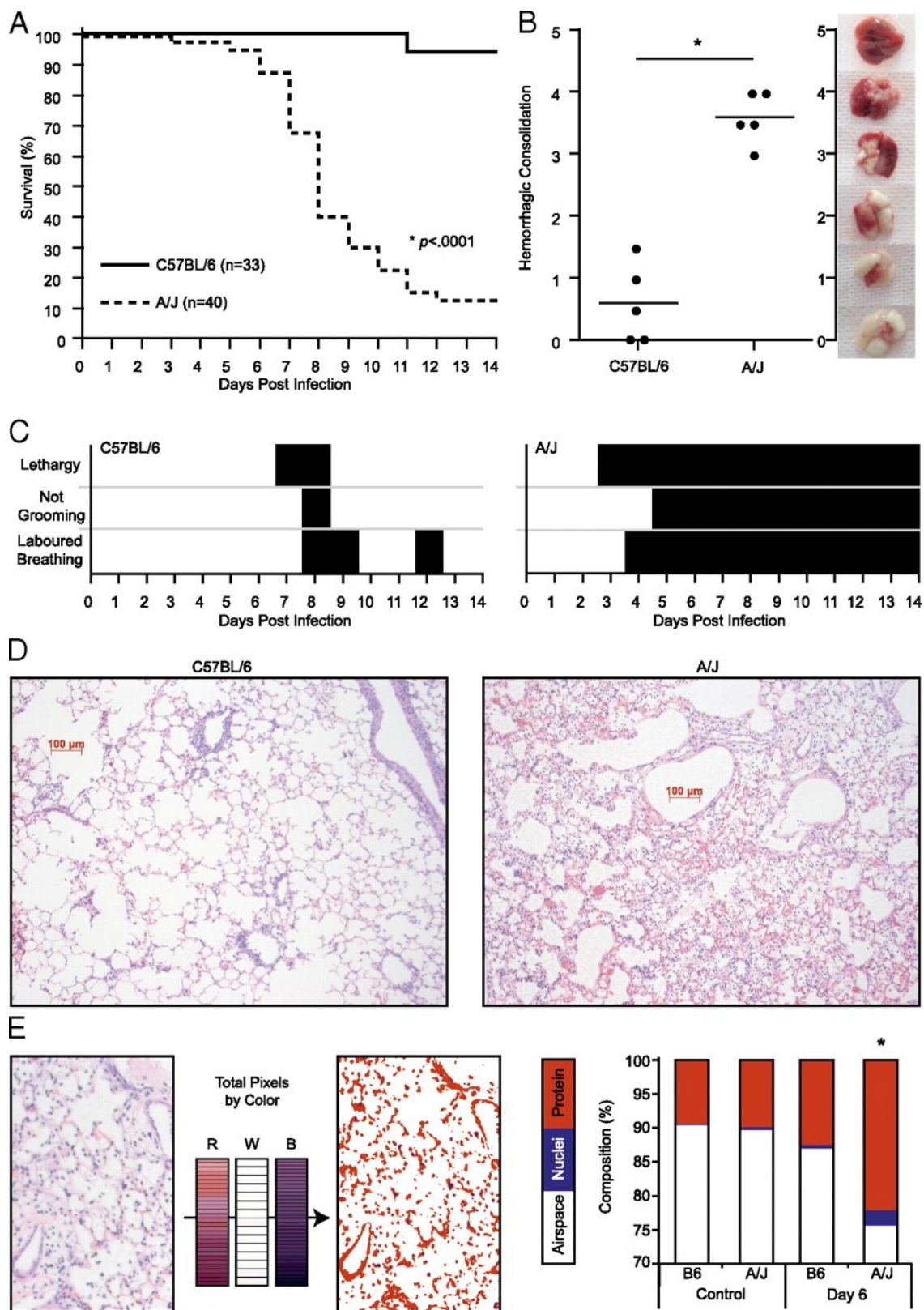


Figure 1 (legend appears on the following page)

Figure 1: Interstrain phenotypic differences in host response to mouse-adapted H3N2 clinical isolate HK/68. Male (six to ten per group) were intranasally infected with 10^4 PFU (**A, C–E**) or 10^3 PFU (**B**) and followed for 6 d (D, E) or 14 d (A–C). A/J mice had significantly reduced survival (A) ($*p < 0.0001$, χ^2 test) and increased clinical signs of infection (C) compared with B6 mice. Upon necropsy, A/J mice showed significantly increased levels of hemorrhagic consolidation compared with B6 mice ($*p < 0.0001$, ANOVA) (B). A/J mice showed more extensive disseminated interstitial inflammation and luminal protein infiltration in H&E-stained lung sections (D). Distribution into one of three preset binary conditions (white, blue, and red) of each pixel in digitized H&E slides identified a significant increase in eosin-associated collagen, muscle fiber, and proteinaceous infiltrates in infected A/J mice (E) ($*p < 0.001$, ANOVA). Original magnification $\times 10$ (D, E).

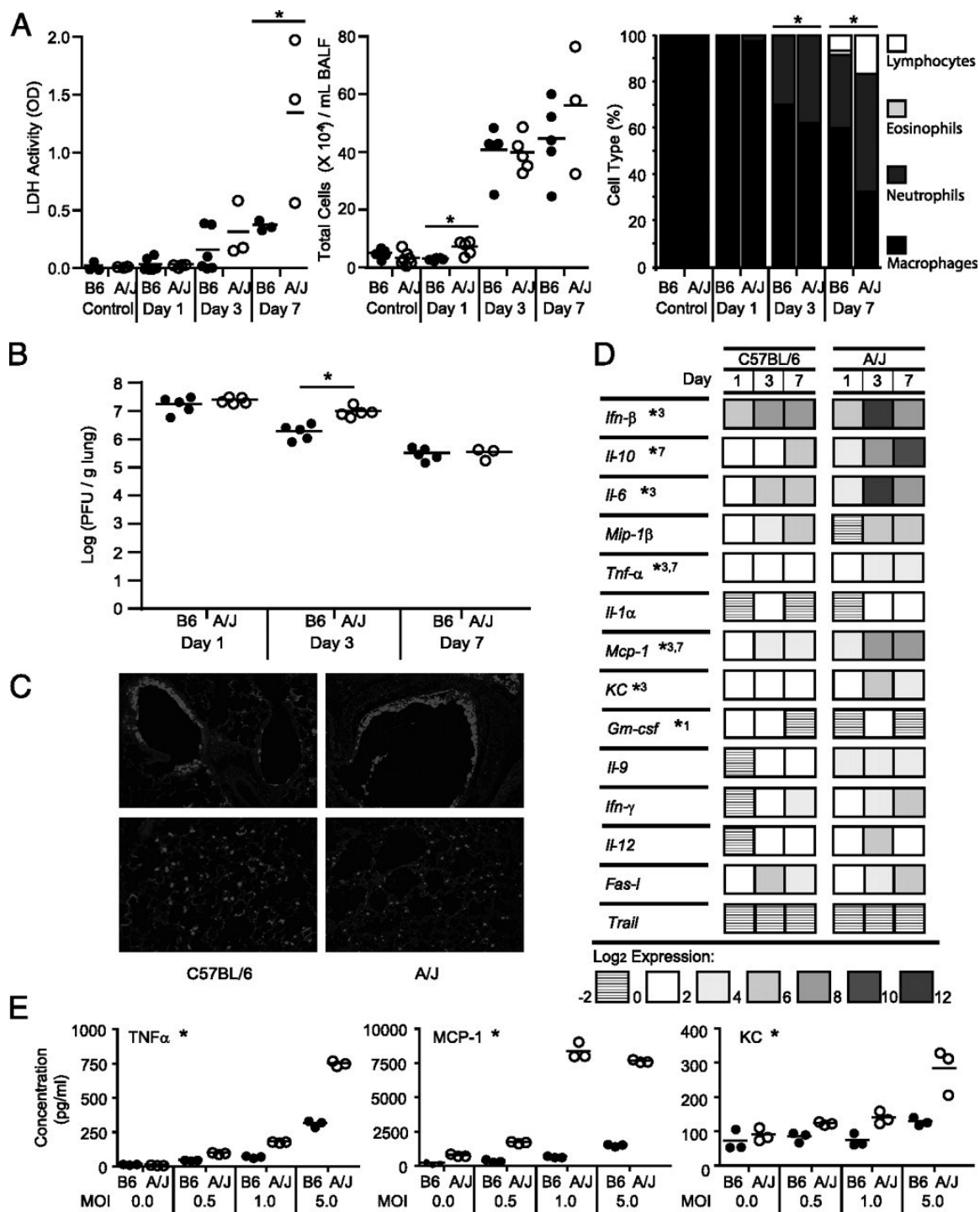


Figure 2 (legend appears on the following page)

Figure 2: Increased lung neutrophilia, cytotoxicity, and inflammation, but not viral load, correlate with susceptibility. Male mice (usually five per group) were infected with 10^4 PFU (A–C) or 10^3 PFU (D) of HK/68 virus. (A) A/J mice showed increased LDH levels (a measure of cell cytotoxicity) by day 7 postinfection (*left panel*, $*p < 0.001$, ANOVA) in BALF. After RBC lysis, the total number of cells in BALF was estimated using a hemocytometer. A/J and B6 mice showed a similar number of total cells in BALF across infection, with no significant differences occurring at day 0 (Control), day 3, or day 7 postinfection (*middle panel*). A small, but significant, increase in overall cell numbers was noted in A/J mice on day 1 (*middle panel*, $*p = 0.005$). The percentage of macrophages, neutrophils, eosinophils, and lymphocytes is shown (*right panel*). A/J mice showed significantly fewer macrophages compared with B6 mice on day 3 ($*p < 0.01$) and day 7 ($*p < 0.05$) postinfection (*right panel*). Similarly, A/J mice showed significantly increased neutrophils on day 7 postinfection (*right panel*, $*p < 0.001$). No significant differences were found within eosinophil or lymphocyte populations (*right panel*). A/J and B6 mice had similar viral loads throughout infection, as determined by plaque assay. (B) A/J mice had slightly higher levels on day 3 postinfection ($*p < 0.001$, ANOVA). (C) Similar localization of the virus was observed in A/J and B6 mice at day 6 postinfection when slides were stained with primary Ab against HK/68. Original magnification $\times 20$. (D) A/J mice had increased proinflammatory gene expression in whole-lung tissue compared with B6 mice when measured by qPCR; genes with significant differences by ANOVA are marked by asterisks and the day of significance (i.e., day 1, 3, or 7). *Il-9* was significant by ANOVA, but the day-specific post hoc tests were not significant. (E) Bone marrow-derived macrophages differentiated from A/J and B6 mice were incubated with various multiplicities of infection (MOI) of HK/68. TNF- α , MCP-1, and KC levels in cell supernatant were quantified by ELISA at 16 h postinfection and normalized to LDH activity; all differences in protein levels between A/J and B6 mice were significant by ANOVA ($*p < 0.001$).

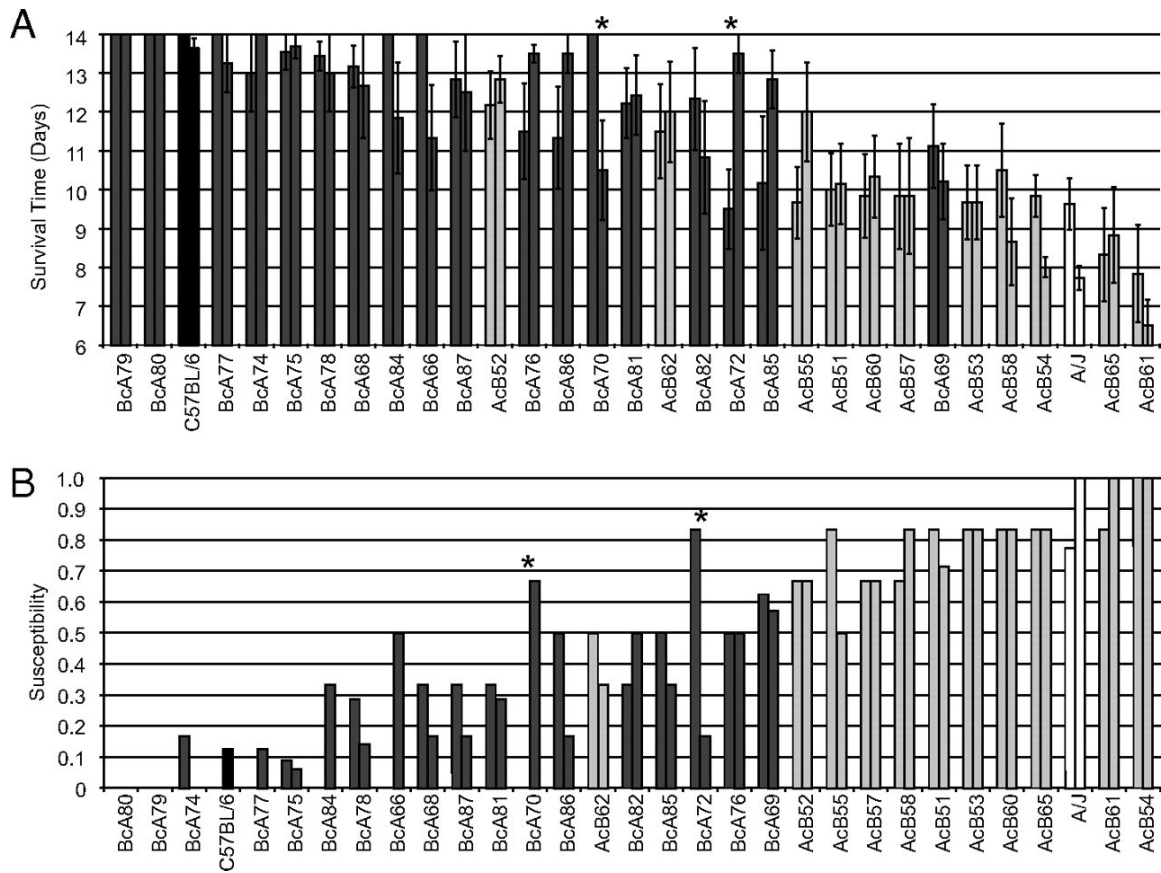


Figure 3: Complex sex-specific genetic control of host response. At least six male and six female mice from 29 RCS strains were infected with 10^4 PFU of HK/68 virus/22 g body weight and followed for 14 d ($n = 398$). Average survival time (**A**) and susceptibility (**B**) for BcA strains (dark gray), AcB strains (light gray), A/J (white), and B6 (black) are shown for females (*left bar*) and males (*right bar*) of each strain. Significant sex $\times$ strain interactions were identified for BcA70 and BcA72. $*p < 0.05$, ANOVA.

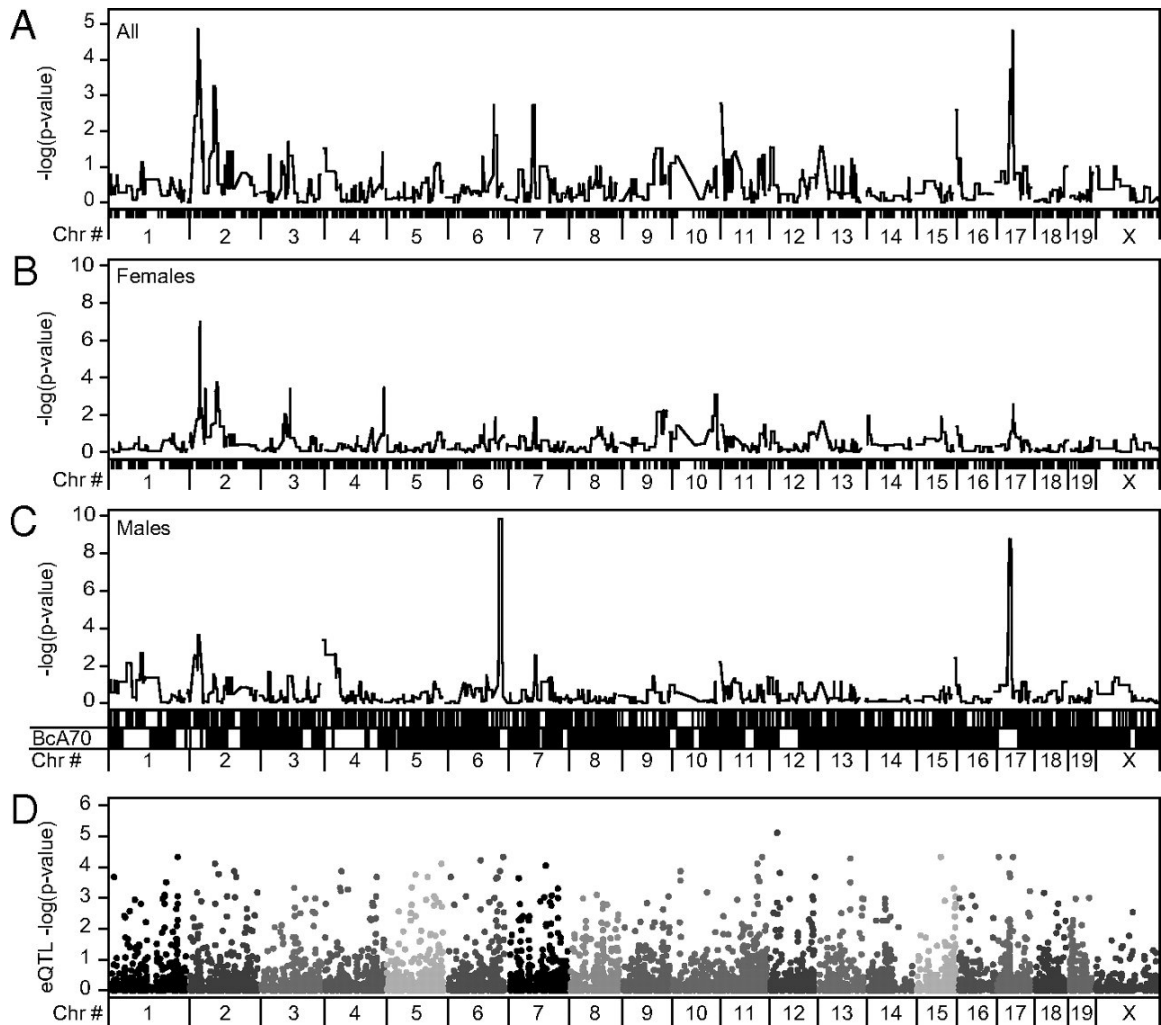


Figure 4: Increased susceptibility maps to sex-specific candidate cQTL. Genome-wide linkage analysis was done in the RCS population ($n = 398$) using a binary measure of susceptibility. The negative log genome-wide p values are shown for a pool of both sexes (A), females (B), and males (C), illustrating a female-specific locus on chromosome 2 and a male-specific locus on chromosome 17. Expression levels of genes were identified in whole-lung tissue of AcB/BcA mice using MOE430A microarrays. *Cis*-regulated expression differences in the AcB/BcA set of mice were identified by ANOVA. (D) The negative log of the Benjamini-Hochberg corrected p value is shown for each gene. Different shades of gray indicate different chromosomes.

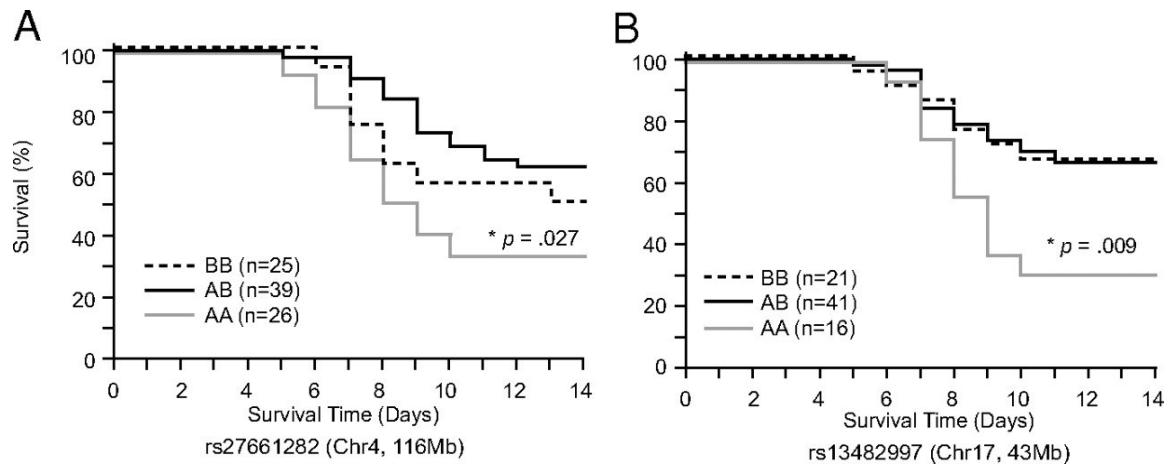


Figure 5: Male-specific cQTL confirmed through a secondary F₂ screen. F₂ mice (177 total) bred from susceptible BcA70 and resistant B6 mice were infected with 10⁴ PFU of HK/68 virus and followed for 14 d. Survival of the F₂ mice segregated by genotype at peak markers on chromosomes 4 (A) and 17 (B). Significant p values are indicated.

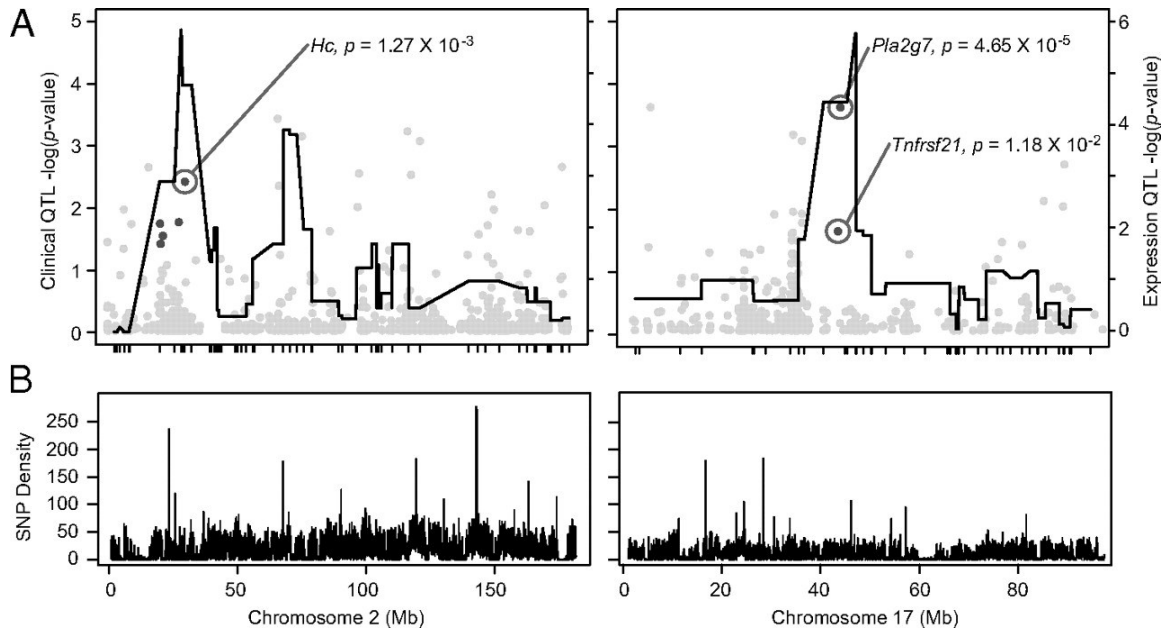


Figure 6: Primary candidate genes identified by combining cQTL and eQTL datasets. (A) Colocalization of the female-specific cQTL and lung tissue-derived eQTL on chromosome 2 (left panel, *Dpp7*, *Phpt1*, *Notch1*, *St6galnac4*, and *Hc*) and colocalization of the male-specific cQTL and lung tissue-derived eQTL on chromosome 17 (right panel, *Tnfrsf21* and *Pla2g7*) are shown. (B) SNP density per 25k bin is displayed for chromosomes 2 and 17 to indicate regions of genetic heterogeneity between A/J and B6 mice.

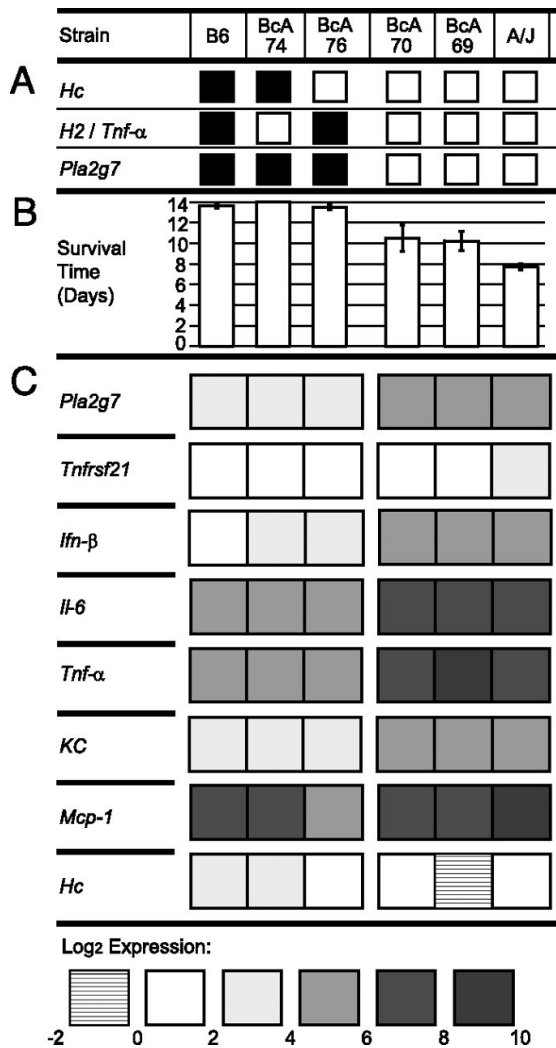


Figure 7: Chromosome 17 QTL associated with decreased survival and increased cytokine expression. Groups of at least three male mice were infected with 10^3 PFU of HK/68 virus. Genotypes at relevant loci (A) and male mice average survival time (B) for a subset of RCS. (C) Lung tissue was collected on day 3 postinfection and analyzed by qPCR for expression of a panel of cytokines and chemokines. ANOVA test revealed significant interstrain differences for all cytokines analyzed ($p < 0.05$). High expression of proinflammatory mediators was associated with the chromosome 17 cQTL and with the *Pla2g7* eQTL but not with *Hc* or the *H2* locus genotypes. This difference in expression was mirrored by the survival rates of male mice from the original RCS screen.

Table 1. PCR primers

Gene	Forward	Reverse
<i>Pla2g7</i>	5'-ATC GAC CTC ACC AAC AAA GC-3'	5'-TCA TCA TCT CCT TCC ACC AG-3'
<i>Ifnb</i>	5'-TGA CG GAGA AGA TGC AGA AG-3'	5'-ACC CAG TGC TGG AGA AAT TG-3'
<i>Mcp1</i>	5'-AGG TGT CCC AAA GAA GCT GTA-3'	5'-TCT GGA CCC ATT CCT TCT TG-3'
<i>Il6</i>	5'-CAT GTT CTC TGG GAA ATC GTG-3'	5'-TTC TGC AAG TGC ATC ATC G-3'
<i>KC</i>	5'-CAC CTC AAG AAC ATC CAG AGC-3'	5'-CTT GAG TGT GGC TAT GAC TTC G-3'
<i>Fas-I</i>	5'-CTG GTT GGA ATG GGA TTA GG-3'	5'-TAC TGG GGT TGG CTA TTT GC-3'
<i>Tnfa</i>	5'-CAT CTT CTC AAA ATT CGA GTG ACA A-3'	5'-TGG GAG TAG ACA AGG TAC AAC CC-3'
<i>Il9</i>	5'-GGC ATC AGA GAC ACC AAT TAC-3'	5'-TTG GGA CGG AGA GAC ACA AG-3'
<i>Ifng</i>	5'-CAC GGC ACA GTC ATT GAA AG-3'	5'-CAT CCT TTT GCC AGT TCC TC-3'
<i>Trail</i>	5'-TCA CCA AC GAGA TGA AGC AG-3'	5'-TCC GTC TTT GAG AAG CAA GC-3'
<i>Gmcsf</i>	5'-CAT CAA AGA AGC CCT GAA CC-3'	5'-GGT CTG CAC ACA TGT TAG CTT-3'
<i>Il1a</i>	5'-AGA GAG GGA GTC AAC TCA TTG G-3'	5'-GAG ATG GTC AAT GGC AGA AC-3'
<i>Mip1b</i>	5'-GAA GCT CTG CGT GTC TGG CCT CTC-3'	5'-CCA CAG CTG GCT TGG AGC AAA GAC-3'
<i>Il10</i>	5'-AGC CGG GAA GAC AAT AAC TG-3'	5'-TCC AGC TGG TCC TTT GTT TG-3'
<i>Gapdh</i>	5'-ACC ACA GTC CAT GCC ATC AC-3'	5'-TCC ACC ACC CTG TTG CTG TA-3'
<i>HPRT</i>	5'-CAG GCC AGA CTT TGT TGG AT-3'	5'-TTG CGC TCA TCT TAG GCT TT-3'

Table 2. Significant cQTL in the total, male, and female RCS populations

All	Male	Female	Position (Mb)
2	NS	2	24–38
NS	6	NS	127–135
17	17	NS	37–48

BRIDGING STATEMENT FROM CHAPTER 3 TO CHAPTER 4

In the previous chapter, we switched our attention from a model of ectromelia virus to influenza virus. In our preliminary influenza work, we identified 2 QTL associated with influenza susceptibility (on chromosomes 2 and 17). We confirmed the likelihood of the hemolytic complement (*C5*) gene on chromosome 2 in controlling this susceptibility locus, and identified *Pla2g7* on chromosome 17 as an interesting candidate gene for further study.

In the following chapter, we turn to chromosome 11 and the BcA80 strain of recombinant congenic mice. We identify several QTL on chromosome 11 associated with resistance to influenza infection. We further go on to identify a de-novo mutation in the *Nf1* gene as a candidate underlying resistance to influenza infection on chromosome 11. The impact and implications of these results are then discussed.

CHAPTER 4: MULTIPLE NOVEL QUANTITATIVE TRAIT LOCI ON MOUSE CHROMOSOME 11 CONTROL HOST SUSCEPTIBILITY TO INFLUENZA A/HK/1/68-MA20.

Gregory A. Boivin¹, Mathieu Mancini¹, Pablo E. Cingolani^{1,2}, Robert Sladek^{1,2,3},
and Silvia M. Vidal¹

¹ Department of Human Genetics, McGill University, Montreal, Quebec, Canada

² McGill University and Genome Quebec Innovation Centre, Montreal, Quebec,
Canada

³ Department of Medicine, Experimental Medicine Division, McGill University,
Montreal, Quebec, Canada

Running title: Multiple Novel QTL on Chromosome 11 Control Influenza
Susceptibility

ABSTRACT

Seasonal influenza outbreaks and recurrent influenza pandemics present major challenges to public health. By studying immunological responses to influenza in different host species, it may be possible to discover common mechanisms of susceptibility in response to various influenza strains. This could lead to novel therapeutic targets with wide clinical application. Here, we analyze the genetic components underlying the increased resistance of the BcA80 mouse strain using a mouse-adapted strain of influenza (A/HK/1/68-MA20 [H3N2]), which induces an overwhelming inflammatory response and death in susceptible mice. We perform linkage analysis with resistant BcA80 mice and susceptible C57BL/6 (B6) controls, and identify novel regions of linkage to influenza virus resistance on mouse chromosome 11. The contribution of proximal chromosome 11 to the survival time of H3N2-infected mice was further confirmed in congenic mice. We put forth candidate genes through the analysis of previously identified lung eQTL or novel mutations identified by whole genome sequencing.

INTRODUCTION

The influenza virus causes 250,000 deaths annually worldwide and is the only known pathogen to cause recurrent pandemics (e.g., 20 million deaths during the 1918 pandemic [244]). Predicting future strains of influenza that will significantly impact human health is difficult due to the broad host range of the virus and its ability to engage in reassortment events. The repeated emergence of pathogenic influenza strains and antiviral resistance illustrate the need to identify new mechanisms of host resistance against the virus. Doing so will allow new therapeutic strategies against this virus to be developed.

Host genetic factors are important in the onset, progression, and outcome of many viral infections [228], [229], [230], [231]. Many genetic risk factors have been identified for influenza virus infection, such as sex [247], blood type [248], [249] ethnicity [250], and familial history of susceptibility to influenza [251], [252]. Because more than one genetic factor often influence resistance to viral infections, they can be difficult to study in human populations [5], [253].

The host immune response is known to play a significant role in the pathogenesis of influenza virus infections. On one hand, a weak immune response such as the one in immunocompromised individuals and the elderly can lead to unimpaired viral replication and severe disease [92]. On the other hand, an exacerbated response, such as those associated with hypercytokinemia and enhanced infiltration of immune cells in the lung, can result in immunopathology and acute respiratory distress syndrome [94], [95], [96].

Genetic mapping with mouse models of infection has been used to successfully dissect both clinical and expression-based complex disease phenotypes [9], [10], [11], [12]. Mouse models have identified novel markers of influenza pathogenesis [255], [254], as well as genes involved in the host response to influenza [256],

[259], [260]. Additionally, several QTL have now been identified in the mouse with respect to influenza virus [157], [159], [160], [161].

Using the highly pathogenic H5N1 strain, QTL for influenza resistance were identified on chromosomes 2, 7, 11, 15, and 17 [159]. A natural premature stop codon in DBA/2 mice was found in the hemolytic complement gene on chromosome 2. Additional work also identified a QTL controlling cytokine levels (TNFa, IFNa, and CCL2) on chromosome 6 [160]. Another study using the collaborative cross recombinant inbred panel of mice and the PR8 strain of influenza identified QTL on chromosomes 1, 7, 15, and 16 [161].

Previously, our group identified QTL on chromosomes 2 and 17 by screening the AcB/BcA panel of recombinant congenic strains (RCS) of mice using a pathogenic mouse adapted strain of influenza [236]. The co-localization of expression QTL from lung tissue identified C5 as a potential gene controlling the susceptibility associated with the chromosome 2 QTL in our model, confirming previous studies above. We also validated the chromosome 17 QTL and identified *Pla2g7* and *Tnfrsf21* as good candidates controlling its linked susceptibility. Additionally, we identified the BcA80 strain of recombinant congenic mice as even more highly resistant to influenza infection than the resistant C57BL6 (B6) parent strain.

In the present study, we aim to break down the genetic components underlying the increased resistance of the BcA80 mouse strain. We utilize the mouse-adapted influenza strain A/HK/1/68-MA20 (H3N2)[262], which induces an overwhelming inflammatory response and death in susceptible mice. We perform linkage analysis with resistant BcA80 and susceptible B6 controls, and identify novel regions of linkage to influenza virus resistance on mouse chromosome 11. We put forth candidate genes through the analysis of previously identified lung eQTL or novel mutations identified by whole genome sequencing.

MATERIALS AND METHODS

Animals and Ethics

BcA80 ($N = 81$) is a recombinant congenic strain of the AcB/BcA panel derived from two successive backcrosses (N_3) of A/J onto C57BL/6 (B6) as previously reported [189]. BcA80 contains roughly one eighth of its genome from A/J, randomly interspersed through a B6 background. B6 mice ($N = 84$) were maintained in-house. Progeny ($N = 240$) from reciprocal (BcA80 $\times$ B6) F_2 crosses were generated in-house. Consomic Substitution Strain 11 (CSS11) mice ($N = 10$), which contain chromosome 11 from A/J and all other chromosomes from B6, were obtained from Jackson Laboratories (Bar Harbor, Maine). To develop congenic strains, CSS11 mice were back-crossed to B6 and the progeny were genotyped for several polymorphic microsatellite markers within the QTL region on chromosome 11. Males that carried heterozygous alleles were selected as breeders for the next generation. By the fifth generation, all breeders were heterozygous for the CSS11 alleles in the targeted region but homozygous for the B6 alleles in the rest of the genome. At this point, mice were intercrossed to obtain homozygous proximal chromosome 11 animals. Experimental protocols were in accordance with the institutional guidelines of the Canadian Council on Animal Care. Mice were maintained at McGill University animal facilities in compliance with the Canada Council on Animal Care as regulated by the McGill University Animal Care Committee.

Cell lines and virus strains

Madin–Darby canine kidney cells (MDCK, ATCC CCL-34) were maintained in Dulbecco's MEM medium supplemented with penicillin (100 U/ml), streptomycin (100 μ g/ml), and FBS (10%). A/HK/1/68-MA20 is a H3N2 mouse-adapted strain of influenza virus derived from a clinical isolate from Hong Kong during the 1968 pandemic, as previously reported [262]. The virus was adapted to mice by 20 serial lung-to-lung passages. The adaptation of the virus led to a $10^{3.5}$ change in the LD_{50} and the selection of 11 mutations, three of which are in common with

the virulent human H5N1 isolate A/HK/156/97 [262]. Titration of infectious virus was determined by a plaque assay as previously described [236].

Infection of mice, definition of phenotype, and tissue collection

The F₂ mice were infected in groups of 35-49 mice. All F₂ mice were aged 10-13 weeks at the time of infection. For infection, mice were anesthetized using a 50 µL intramuscular injection of a ketamine/xylazine cocktail and infected intranasally with 227 pfu/g A/HK/1/68-MA20, diluted in PBS to a final volume of 1 µL/g mouse.

For the susceptibility screen, mice were monitored daily for two weeks following inoculation. Clinical signs (weight loss, labored breathing, lack of grooming, and low motility) were recorded daily. Mice presenting respiratory distress or a body conditioning score lower than 2 were humanely sacrificed. For all analyses, mice were euthanized by CO₂. For qPCR and viral titer, lungs were perfused with 5 mL of cold sterile phosphate buffered solution (PBS), flash frozen, and stored at -80°C. For histology, the lung lobes were inflated under constant pressure of 25 cm-H₂O with 10% neutral buffered formalin (Sigma-Aldrich) as previously described [236].

Histology and image analysis

Fixed lung samples were sent to the Histology Department at the University of Montreal (Canada). Samples were embedded in paraffin blocks, excised in 5 µm thick sections and stained with hematoxylin and eosin (H&E) for light microscopic examination. Sections were also analyzed using custom colour recognition software designed to identify whitespace (open air spaces); hematoxylin-associated nuclei and eosin-associated collagen; muscle fiber; and proteinaceous infiltrates as previously described [236]. Briefly, each pixel in the histological image was filtered into one of three preset binary conditions (white, blue, and red) based on 8-bit color space.

FACS ANALYSIS

Mice were sacrificed on D7 post infection, and spleens were collected from mice on ice in 15ml falcon tubes containing 5 mL RPMI 10% FBS + 5 mL P/S (hereafter referred to as Complete RPMI). Spleens were gently crushed with the plunger of a 3 mL syringe over a cell strainer, with the cell suspension transferred to a new 15 mL tube. The suspension was centrifuged at 1400 rpm, at 5 min, 4⁰C (hereafter centrifuged). The supernatant was discarded. The pellets were resuspended in 2 mL RBC lysis buffer and homogenized with p1000 pipette to lyse red blood cells. The tubes were filled with cold complete RPMI after 3 minutes to stop the lysis reaction. The solution was again centrifuged and the pellet was resuspended in 10 mL cold Complete RPMI and passed through a cell strainer into a new 50 mL Falcon tube. Cells were counted, centrifuged and resuspended at 2×10^6 cells/ml in Complete RPMI.

100 μ L of cells were plated per well in a 96 well plate. The plate was centrifuged, and the supernatant was discarded. 100 μ L FACS buffer was added to each sample, and the plates were incubated for 20 minutes at 4⁰C for 20 minutes. The plates were again centrifuged, and the supernatant discarded. Cells were stained in 50 μ L FACS buffer for the following markers: CD69, Dx5 (CD49b), CD4, CD44, CD3, CD8, Gr-1 (Ly-6G), F4/80, B220 (CD45R), MHCII, CD11c, CD11b, TCRb. Cells were homogenized and incubated at 4⁰C for 20 min. Cells were centrifuged and washed twice with PBS. Viability dye eFluor506 (V500 was diluted in PBS and 50 μ L was added to each well. Cells were incubated at 4⁰C for 10 minutes. Cells were washed twice with PBS and centrifuged, and the supernatant was discarded.

Cells were fixed using 100 μ L PBS with 100 μ L 4% paraformaldehyde (PFA) per well, for a final concentration of 2% PFA per well. Fixation will stabilize the light scatter and inactivate most biohazardous agents. Cells were incubated at 4⁰C in the dark for 20 minutes. Cells were washed twice with PBS and centrifuged, and the supernatant was discarded. Cells were resuspended in 200 μ L FACS buffer,

and refrigerated at 4°C in the dark overnight. The following day, 1 million events per sample were acquired on the Canto II FACS machine and analyzed with FlowJo software (TreeStar).

Quantitative PCR

Total RNA was extracted by homogenization and Trizol (Invitrogen) extraction. cDNA was transcribed using M-MLV with random hexamers (Invitrogen) according to the manufacturer's instructions. Real-time quantitative PCR (qPCR) was performed using Platinum SYBR Green SuperMix-UDG (Invitrogen) together with experimental or control primers. Target transcripts were normalized to hypoxanthine phosphoribosyl-transferase 1 (*Hprt*) expression. Samples were run in duplicate with three to five mice per condition. Reactions were performed using the Step One Plus real time PCR system (Applied Biosystems), and expression was analyzed using StepOne v2.2.2. Relative mRNA expression was shown as $2^{-\Delta\Delta CT}$ for cytokines or $2^{-\Delta CT}$ for viral comparisons.

Genotyping and Sequencing

Genomic DNA for genotyping was prepared from tail biopsies using Proteinase K and serial phenol/chloroform extractions followed by ethanol precipitation as previously described [189]. The reciprocal (BcA80 × B6) F₂ crosses (N = 240) were genotyped with 34 markers spanning the informative A/J segments of BcA80 mice using Sequenom iPLEX Gold technology. BcA80 mice were sequenced using Illumina Hi-seq paired-ends whole genome sequencing, read length 100 bases. Reads were mapped to the reference genome (UCSC mm9 / ENSEMBL MM37.63) using BWA (version 0.5.9) [293]. Duplicates were removed using SamTools [294] (version 0.1.17) and variants called using BcfTools from the SamTools suite [294]. Variant calling with and without the BAQ model [295] was performed to assess differences between BcA80 and reference sequence (B6) in high variation regions. Variants were annotated using SnpEff [296] and further filtered using SnpSift [297] to obtain a list of non-synonymous, homozygous variants respect to the reference (B6) genome. Both the genotyping

and sequencing were performed at the McGill University and Genome Quebec Innovation Centre (Montreal, Canada).

Statistical and bioinformatic analyses

All of the statistical analyses were performed using the freely available package R [237]. F_2 linkage analyses were performed using the R “qtl” package. To identify candidate genes within regions of linkage, we examined previously reported cis-eQTL from the basal lung tissue of the AcB/BcA recombinant congenic panel of mice [236].

RESULTS

BcA80 mice are resistant to influenza virus infection as compared to B6

The BcA80 strain of mice was identified as being hyper-resistant in a previously reported preliminary screen that identified differential host responses to influenza virus infection in the AcB/BcA recombinant panel of mice [236]. To quantify this difference, we infected groups of BcA80 and B6 mice with two increasing doses of A/HK/1/68-MA20: 45 pfu/g and 227 pfu/g (equivalent to 1000 and 5000 pfu/22g mouse). BcA80 mice were significantly more resistant than B6 at both doses (Fig. 1A, $p < 0.05$ for 1000 pfu/22g, $p < 0.001$ for 5000 pfu/22g). BcA80 also showed differential clinical signs across infection. BcA80 showed less evidence of labored breathing, hunched posture and lethargy across the peak period of infection (roughly day 6-12 post infection), but showed worse (for the 1000 pfu/22g dose) or equivalent (for the 5000 pfu/22g dose) grooming habits during the first week of infection. Because previous results have shown that sex can significantly influence susceptibility in models of influenza infection, we looked for evidence of a sex effect in the current study. Male mice appear to have slightly increased levels of susceptibility as compared to female mice for both doses of influenza (data not shown), however these differences did not lead to significant results.

BcA80 mice show altered immune response to influenza infection compared to B6

To further characterize the differential susceptibility observed between BcA80 and B6 mice, we examined expression of viral RNA at days 1, 3, 5, and 7 and histological sections of lungs at day 6 post infection in male mice. BcA80 mice showed a significant reduction in lung virus expression by qPCR compared to B6 by day 7 post-infection (Fig. 2A, $p < 0.05$). This observation was confirmed using plaque forming assays to determine virus load on day 7 post infection. (Fig. 2B, $p < 0.01$). The kinetics of viral expression seemed otherwise comparable, with both strains reaching a maximum on day 1 and decreasing generally across infection.

BcA80 mice appeared to have a significantly increased inflammatory response as compared to B6 by day 6 post infection. Ten images representing H&E slides from each experimental condition (i.e., BcA80 uninfected (3 mice), B6 uninfected (3 mice), BcA80 infected (5 mice), and B6 infected (5 mice)) were analyzed using color recognition software. At day 6-post infection, we identified a significant increase in hematoxylin-associated nuclei and eosin-associated collagen, muscle fiber, and proteinaceous infiltrates in infected BcA80 mice (Fig. 2C, $p < 0.01$). Thus, in terms of histopathological phenotypes, BcA80 mice appear to elicit a stronger inflammatory response by day 6 post infection and clear virus from lung tissue faster than B6 mice. To account for this increase in inflammatory cells in the lung, we also examined spleens of BcA80 and B6 mice peripheral to the infection site. BcA80 mice showed increased spleen index on day 3 post infection, and increased spleen size on day 7 post infection compared to B6 controls (Fig. 2E-G). Flow cytometry analyses comparing BcA80 and B6 spleen cell populations on day 3 post infection revealed significantly increased frequencies of CD4⁺ T cells and cDCs, as well as significant decreases in activated CD8⁺ T and NKT cell populations (Fig. 2 H-O). These results indicate a significantly altered immune response in BcA80 as compared to B6 mice, comprising altered spleen growth after infection, altered composition of splenic cells post infection, and altered immune/inflammatory cell recruitment to the lung in response to influenza infection.

Resistance of BcA80 is mapped to multiple loci on chromosome 11

Because BcA80 and B6 mice show differential resistance to influenza infection, we created an F₂ cross from these two progenitor strains. F₂ mice ($N = 240$) were infected with 227 pfu/g as followed for 14 days post infection. A marker map was created that spanned all sections of A/J from the resistant BcA80 strain (i.e., sections of chromosomes 6, 9, 11, 12, 18 and 19). Two phenotypes, early sacrifice and bodyweight loss, were used to analyze the F₂. Survival was a binary measure indicating whether mice were sacrificed within the first week of infection. Bodyweight loss was recorded at day 5 post-infection. Using these measures of

survival and bodyweight loss, we performed genome-wide linkage and identified two clinical quantitative trait loci (cQTL) on chromosome 11 (0-28Mb and 40-57Mb, Fig. 3). The sex of mice did not have an impact on the mapping of the QTL. When including sex as a covariate in the analysis, the markers denoting the significant cQTL sizes were identical and negligible differences were seen in cQTL LOD scores (data not shown).

cQTL validated in CSS11 congenic mice

To identify whether the original cQTL derived from A/J regions of BcA80 were valid, we first tested CSS11 mice (N = 10) against B6 (N = 17) and BcA80 (N = 17, Fig. 4B). CSS11 mice are A/J in origin on chromosome 11, but are B6 in origin for all other locations. Contrary to expectations, the CSS11 mice show significantly reduced survival as compared to B6 or BcA80 mice ($p < .01$). Although the CSS11 mice have the proximal A/J segment of chromosome 11 from BcA80, which appears to confer resistance, they also contain the distal chromosome 11 from A/J as well. The increased susceptibility of CSS11 may be due to an overwhelming susceptibility from this distal A/J segment (i.e., 57M-122Mb). To test whether the proximal cQTL could confer resistance, they needed to be isolated from any A/J resistance loci in the distal half of chromosome 11.

We isolated three congenic strains of mice after five generations of successive backcrossing of CSS11 to B6 using marker-assisted introgression (Fig. 4, Table 01). Two heterozygotes were then crossed and homozygous offspring were bred to fix three chromosome 11 haplotypes. A proximal congenic was isolated with an A/J segment of 0Mb to 19.3-28.3Mb (Fig. 4A, the crossover resides somewhere between rs3723987 and rs3706694). A medial congenic was isolated with an A/J segment of 34.3-40.1Mb to 73.1-83.2Mb (Fig. 4A, the proximal crossover is between rs6239937 and rs6164170, the distal crossover is between rs3701609 and rs13481127). A full congenic haplotype was also isolated in an attempt to replicate the proximal chromosome 11 makeup of BcA80 from ancestral A/J and B6 genetic origin. The full congenic was isolated with an A/J segment of 0Mb to

54.7-62.8Mb (Fig. 4A, the crossover is between rs13481033 and rs13481061) into procured B6 mice. The proximal ($N = 11$), medial ($N = 11$), and full ($N = 45$) congenic mouse lines derived from CSS11 were then tested against BcA80 ($N = 5$) and B6 ($N = 7$) to see if, when isolated, they could confer any measure of resistance. Significant differences in survival were seen for all strains ($p < 0.01$, Fig. 4C).

Primary candidate genes identified using eQTL

To identify primary candidate genes within our proximal and medial cQTL on chromosome 11, we searched for genes within our intervals that show evidence of *cis*-regulation in uninfected lung tissue of RCS mice [236]. Significant *cis*-eQTL were colocalized with the cQTL identified above (Fig. 5). The eQTL that colocalize with the proximal cQTL and their corrected p-values are *Igfbp3* (2 probes, highest p-value = 3.91×10^{-3}), *Actr2* (1.52×10^{-2}) and *Slc1a4* (4.12×10^{-2}). *Igfbp3* is a member of the insulin-like growth factor binding protein family and may alter proliferative or apoptotic cellular effects. *Actr2* is an actin-related protein that impacts cell shape and motility. *Slc1a4* encodes a solute carrier that is known to transport alanine, serine, cysteine, and threonine.

The eQTL that colocalize with the medial cQTL and their corrected p-values are *Pttg1* (1.13×10^{-3}), *Tcf7* (1.85×10^{-2}), *Tnip1* (6.87×10^{-3}) and *Gria1* (1.58×10^{-2}). *Pttg1* plays a role in chromosome stability and DNA repair. *Tcf7* encodes a transcription factor that plays a role in lymphocyte differentiation and is expressed primarily in T cells. *Tnip1* is known to inhibit TNF-induced NFkB dependent gene expression. *Gria1* is an ionotropic glutamate receptor, a critical member of excitatory signaling in the central nervous system. Although none of these genes have been previously associated with influenza virus resistance, additional research may be warranted on their role in influenza infection.

Potential new De-novo in the *Nf1* gene on chromosome 11

Because the 40-57Mb cQTL was at its maximal point on the furthest distal marker in our cross, we hypothesized that there may be a background mutation in the B6 region of chromosome 11 in the BcA80 strain. BcA80 is A/J in origin from 0-57Mb, and is B6 in origin from 57-121 Mb. The RCS strains have been kept in house for over 30 generations. This is enough time for *de novo* mutations to become solidified in the RCS strains [192], [298]. Hi-seq paired-ends whole genome sequencing (Illumina) was done on two BcA80 samples. The samples were aligned to the reference B6 genome with an average depth of coverage of 15.35. A total of 169 homozygous mutations were identified in 111 genes in the B6 background of BcA80 (Supplemental Table 1). Of these, five are located on chromosome 11. Of these five, only an R262H mutation in the neurofibromin 1 gene (*Nf1*) at 79.22Mb of chromosome 11 was found to be present in BcA80 and absent in the internal B6 control. The remaining SNPs were identified in both BcA80 and the internal B6 control, reiterating the drift of the RCS colony from the reference strains from which they were originally created. *Nf1* was sequenced in the 240 mice from the original F₂. An intronic mutation in Breast carcinoma amplified sequence 3 (*Bcas3*), which is located 5Mb further distal to *Nf1*, was also sequenced in the F₂. After its inclusion, *Nf1* became the new peak marker in the F₂, controlling roughly 16% of the bodyweight phenotypic variance in the F₂ population (Fig. 6). Again, there was no significant sex effect within the F₂. There is the possibility that another mutation that does not reside on chromosome 11 in BcA80 is in linkage disequilibrium with *Nf1*, and is the true underlying cause of the increased resistance at the *Nf1* locus. More studies need to be done to isolate the *Nf1* mutation on a B6 background to validate these findings. Currently, however, *Nf1* must be viewed as a strong candidate as a factor underlying a portion of the resistance of BcA80 to influenza virus, in conjunction with the new QTL identified and validated on the proximal and medial sections of chromosome 11.

DISCUSSION

We undertook this study to identify new candidate regions and genes involved in BcA80 associated resistance to influenza virus infection. In our mouse model, influenza susceptibility was under complex genetic control, with multiple QTL influencing the overall resistance phenotype of the mouse strain. We identified three cQTL associated with influenza resistance on chromosome 11. Whole genome sequencing identified the likely candidate underlying one cQTL as *Nf1*, which encodes a potential activator of Ras activity [299]. The two remaining cQTL that were derived from ancestral A/J regions of the BcA80 genome were further validated with congenic mice made from CSS11 mice. By comparing eQTL from expression profiles of uninfected lung tissue, we were able to identify potential candidate genes within the two cQTL of ancestral A/J origin.

Functional dissection of susceptibility in our mouse model

After confirming the differential survival of BcA80 and B6 to influenza infection, we tested additional phenotypes in an attempt to further characterize BcA80 associated resistance. Both BcA80 and B6 mice had high levels of viral RNA in lung tissue in early infection, with the highest levels recorded at day 1 post infection. We found a significant difference in the amount of viral RNA between BcA80 and B6 at day 7 post infection. This difference was confirmed via plaque assay on day 7.

This finding is not in line with resistance to many strains of influenza virus, as resistant mice tend to have lower viral levels throughout infection [172], [300]. However, this finding is similar to our previous work with the A/HK/1/68-MA20 virus. We have previously reported a similar trend of high early viral levels for both susceptible A/J and resistant B6 mice [236]. As A/J and B6 are the progenitor strains for BcA80, it is not unreasonable that the hyper resistance of the strain is not associated with early differences in viral levels. We did find a significant difference of viral levels, both by qPCR and plaque assay at day 7

post infection. This effect seems to be indicative of the ability of BcA80 mice to clear infection better than B6 mice. In conjunction with the late difference in viral levels, we also found higher levels of infiltrating cells by histology in resistant BcA80 mice on day 6 post infection. Previous research generally associates increased inflammatory responding with susceptibility to pathogenic influenza [273], [274], [96], [95]. More research with congenic and knockout mouse strains will be needed to identify whether this increased inflammatory response in BcA80 mice is associated with resistance.

One potential explanation for the increased infiltration of cells in the lung after infection is a greater production of immune cells. We identified altered lung composition post infection, and altered percentages of immune cells in the spleen post infection in our BcA80 mice. Again, it is not clear whether this phenotype is aligned with the increased resistance seen in the BcA80 strain. More studies must be done to completely dissect the underlying causes of resistance in BcA80 mice and identify whether the increased spleen size and altered size of immune cell populations are associated with resistance in our model.

Complex genetic control of influenza virus infection in mice

Multiple BcA80 resistance cQTL were found on chromosome 11 in the present study. This is in accordance with the complex nature of resistance seen in many infectious diseases [5]. Nevertheless, it is rare to identify three closely linked QTL in one study, especially as the A/J progenitor from which two of the QTL were inherited is more susceptible than the B6 strain. Interestingly, while chromosome 11 has been identified as linked to influenza susceptibility in a previous screen [159], none of our loci appear to overlap the previously identified regions.

Identifying primary candidate genes using expression QTL

To identify primary candidates within the two cQTL on chromosome 11, we looked for colocalization of the cQTL with that of lung eQTL identified previously in AcB/BcA mice. Since the BcA80 mice are derived from A/J and B6 progenitors

and are a member of the AcB/BcA panel of recombinant congenic strains, we hypothesized that cis-regulated eQTL between A/J and B6 genotypes identified in the AcB/BcA panel may provide gene candidates with altered or absent gene function in our present study. Combining expression and clinical quantitative trait loci has previously identified candidate genes involved in complex disease [12], [236]. Although none of the candidate genes identified for each of the ancestral A/J cQTL have been previously associated with infectious disease, they do represent genes with altered functions within the intervals. One candidate within the medial QTL, *Tnfr1*, is compelling due to its relation to immune function. *Tnfr1* is highly expressed in stimulated bone marrow macrophages and has been shown to repress NFkB signaling and protect from apoptosis [301]. Further, genetic variations within *Tnfr1* have been shown to be associated with several autoimmune diseases [302], [303], [304], [305], [306]. *Tcf7* is another compelling candidate from the medial chromosome 11 QTL, as it is involved in specification and differentiation of T cells [307]. However, these genes mentioned do not discount possible roles for other candidate genes or polymorphisms in the chromosome 11 intervals. More research must be done to identify the underlying causative gene or polymorphisms present in these newly identified loci.

***Nf1* as a primary candidate gene for influenza resistance**

We sequenced the BcA80 genome in order to search for candidate genes in the distal B6 background of BcA80 chromosome 11. The peak marker in the BcA80 F₂ and the most likely candidate to confer resistance from this distal QTL is the de-novo non-synonymous mutation the *Nf1* gene. NF1 is a negative regulator of ras signaling [308]; and mutations in the human *NF1* gene have been previously associated with neurofibromatosis type 1 [309], lung adenocarcinomas [310], leukemia [311], and other genetic syndromes [312]. Mutations in *NF1* have also been shown to lead to proliferative disorders (e.g., cancers), including hypersensitivity to granulocyte macrophage colony stimulating factor [311]. This is the first study to our knowledge, however, to link mutations in *Nf1* to potential altered resistance to infectious disease. More research needs to be conducted to

see if the isolated BcA80 *Nf1* mutation retains resistance to influenza. Congenic mice are currently being created.

Summary

By breaking down the genetic contributions to influenza resistance of the BcA80 mouse strain, we identified three cQTL linked to resistance against influenza infection in mice (Table 02). Two of these cQTL were validated by congenics created from CSS11 mice. Previously reported eQTL were colocalized to the two validated ancestral A/J derived cQTL, providing functional candidates within these regions of linkage. Whole genome sequencing identified *Nf1* as a compelling candidate for underlying the resistance of the third cQTL. Ongoing studies will identify whether these candidate genes may be novel targets for therapy against this globally relevant pathogen.

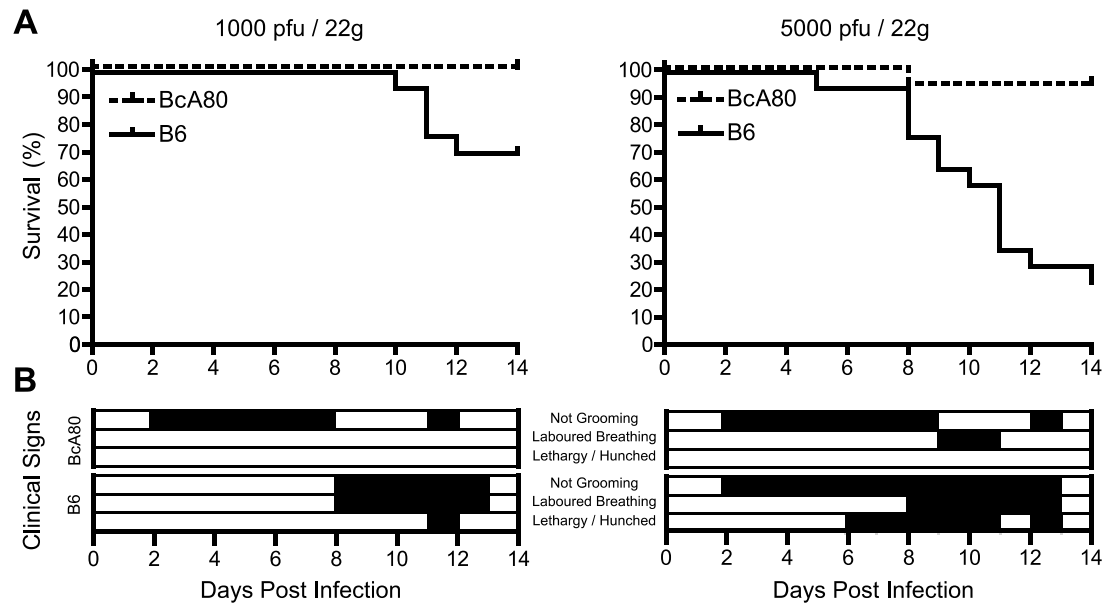


Figure 1: BcA80 mice are resistant to influenza infection as compared to B6 controls. BcA80 shows significantly greater survival of influenza infection at doses of 1000 pfu/22g (n = 16 BcA80, n = 17 B6, A left) and 5000 pfu/22g (n = 16 BcA80, n = 17 B6, A right, $p < 0.001$, Kaplan-Meier). This significant difference in survival was accompanied by differences in clinical signs of infection, including grooming behaviour, laboured breathing, and lethargy or hunched posture (B).

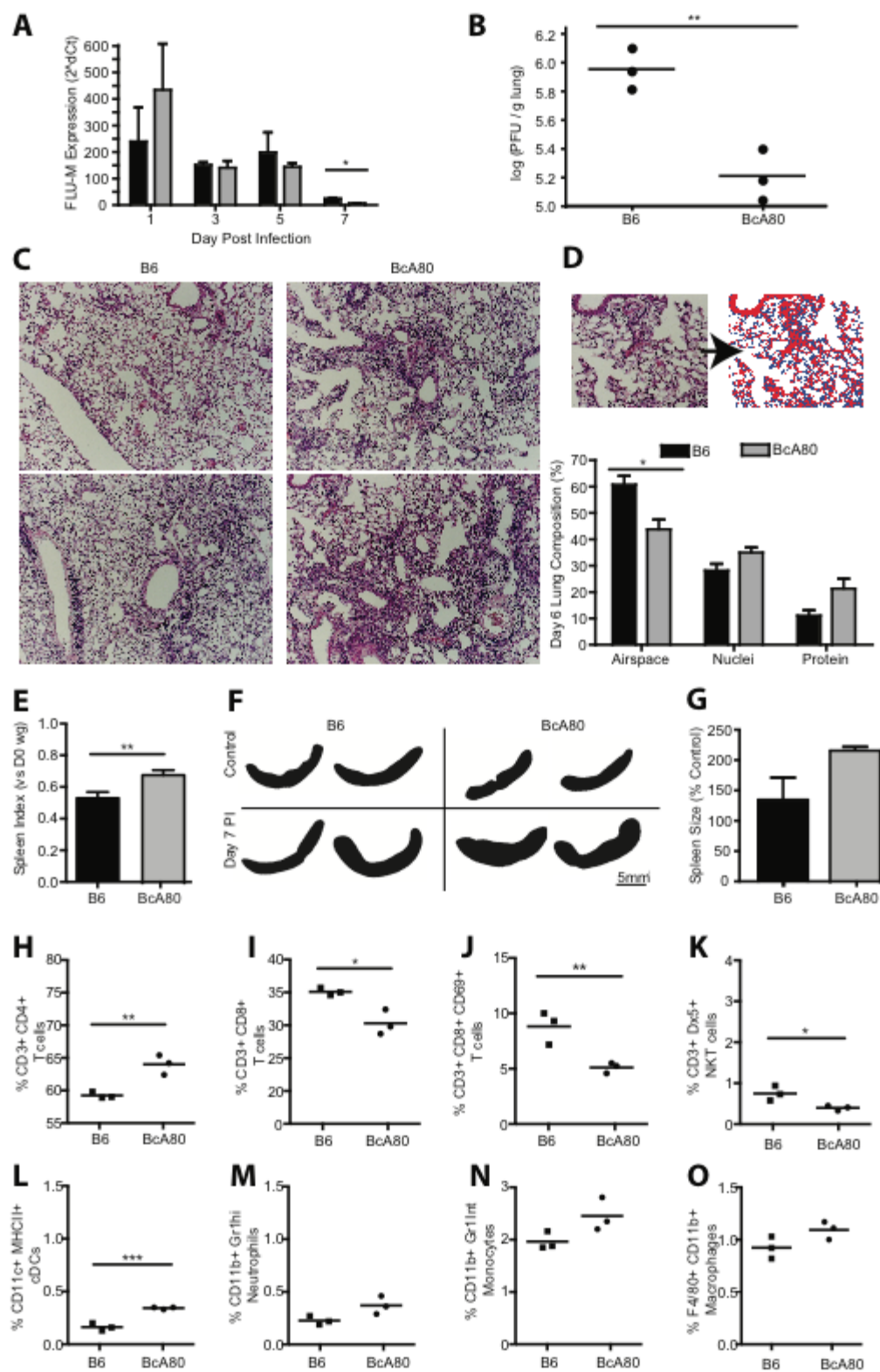


Figure 2 (legend appears on the following page)

Figure 2: Resistant BcA80 mice show altered viral titer, lung and spleen histology, and lung cell counts. Resistant BcA80 mice show significantly less virus by qPCR and viral titer by day 7 post infection (3-5 male mice per group, (A and B, t-test). Resistant BcA80 mice show increased cell infiltration as shown by H&E stained slides (C, representative images). Distribution into one of three preset binary conditions (white, blue, and red) of each pixel in digitized H&E slides identified a significant increase in eosin-associated collagen, muscle fiber, and proteinaceous infiltrates in infected A/J mice (3-5 male mice per group). Resistant BcA80 mice show increased spleen size by spleen index (E, t-test) and spleen size (F and G, representative, 2 male mice per group). Resistant BcA80 mice show altered cell profiles in infected lung tissue (H – O, 3-5 male mice per group, t-test).

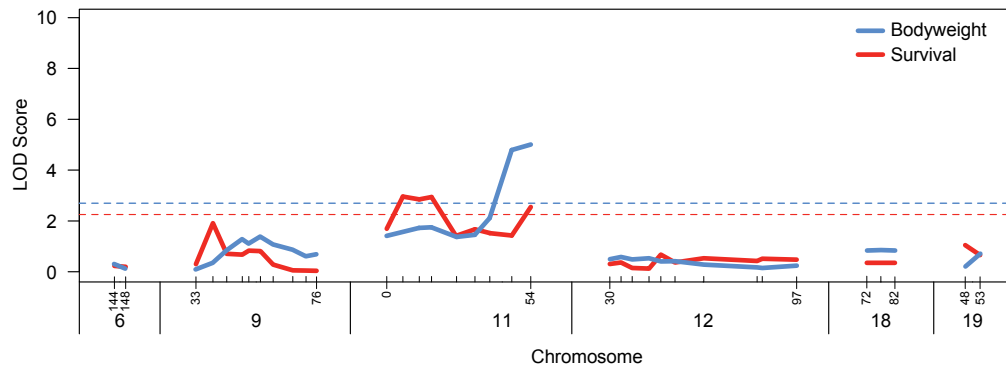


Figure 3: Genetic mapping of resistance to influenza virus infection. 240 BcA80 × B6 F₂ mice were tested for resistance to influenza virus through the analysis of D5 bodyweight or survival phenotypes. The cutoff for significance of the phenotypes is shown (blue for bodyweight, red for survival). There is evidence for a proximal QTL from ~0-28 Mb and a distal QTL from ~40-57 Mb on chromosome 11.

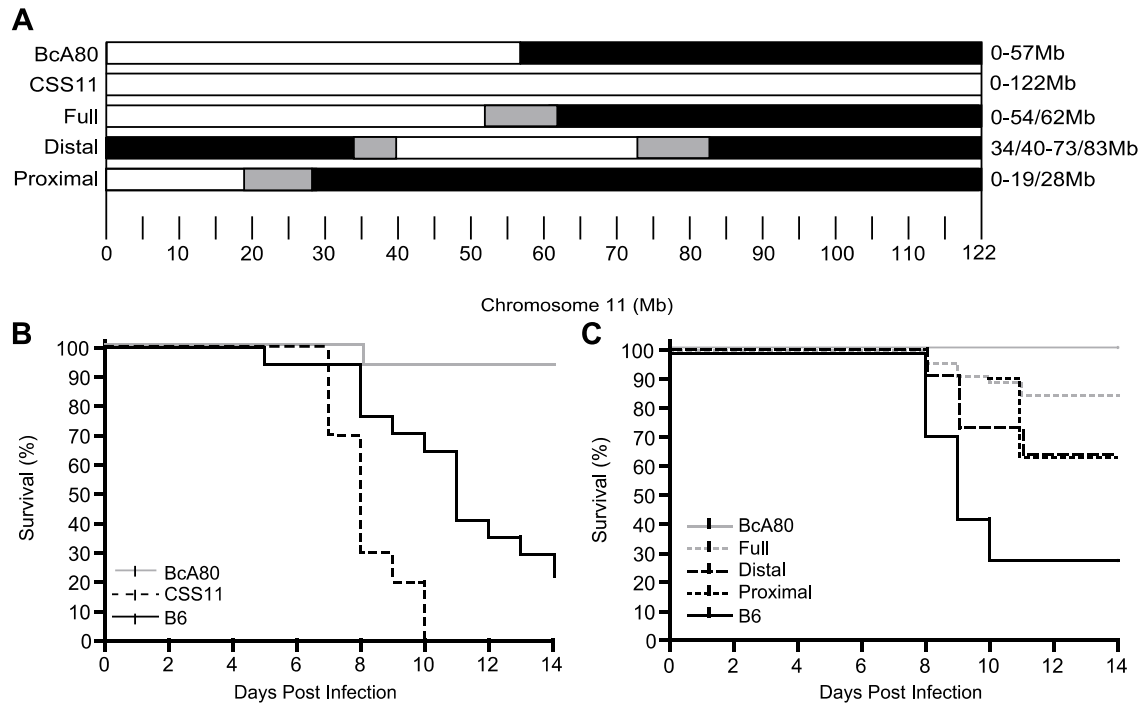


Figure 4: Additional resistance loci validated on chromosome 11. Congenic mouse strains were made from CSS11 and B6 mice to see if the chromosome 11 associated resistance of BcA80 mice stems from the proximal A/J segment (A). Although the CSS11 strain itself was more susceptible to infection (B, $n = 17$ male BcA80, $n = 10$ CSS11 male, $n = 17$ B6 male, $p < .01$ Kaplan-Meier), significantly increased resistance was associated with mice carrying the 0-19/28Mb ($n = 11$ male, $p < .01$ Kaplan-Meier), 34/40-73/83Mb ($n = 11$ male), and 0-54/62Mb ($n = 45$ male) A/J congenic segments as compared to B6 ($n = 7$ male) (C, Table 02 $n = 5$ male BcA80).

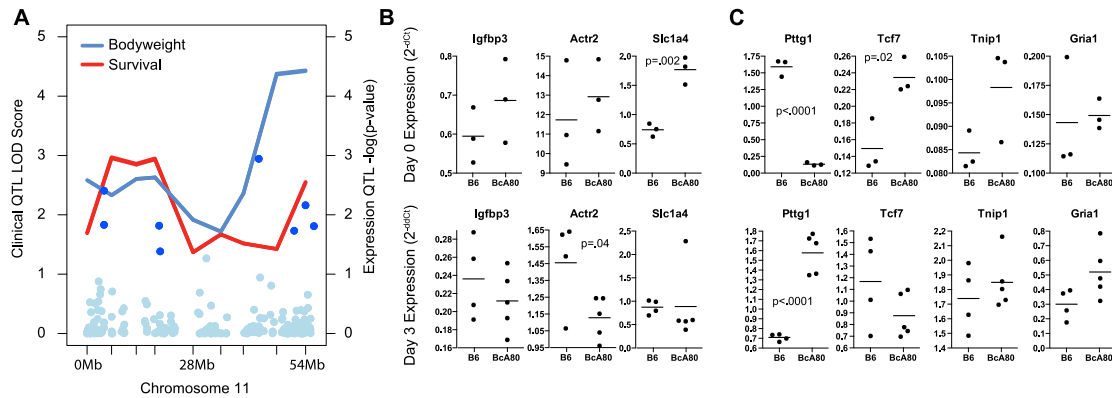


Figure 5: Additional candidate genes identified by combining clinical and expression QTL. The proximal and distal BcA80 resistance loci for bodyweight and survival are shown overlaid with lung tissue derived eQTL on chromosome 11. Significant eQTL under potential regions of linkage are shown in dark blue under the proximal peak (*Igf1bp3* (two probes), *Actr2*, and *Slc1a4*) and distal peak (*Pttg1*, *Tcf7*, *Tnfr1*, and *Gria1*) A. Expression of candidate genes is shown by qPCR in the proximal (B) and distal peak (C) for baseline (day 0) and day 3 post infection. Groups of 3-5 male mice (B and C, t-test).

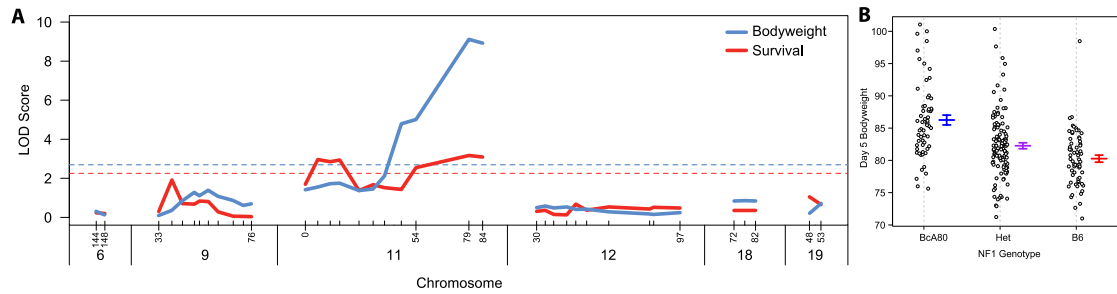


Figure 6: Genetic mapping of resistance to influenza virus implicates a *de novo* *Nf1* mutation. 240 BcA80 × B6 F₂ mice were tested for resistance to influenza virus through the analysis of D5 bodyweight or survival phenotypes. The largest peak in the cross was identified as an nsSNP in *Nf1* (79Mb) (A). The phenotype × genotype plot at the *Nf1* marker is shown (B).

Table 1. Markers used for congenic and F₂ mice.

Project	SNP	Chr	Pos (bp)
Congenic	rs13480836	11	3454200
Congenic	rs13480863	11	9193620
Congenic	rs3723987	11	19368928
Congenic	rs3706694	11	28314860
Congenic	rs6239937	11	34833394
Congenic	rs6164170	11	40117642
Congenic	rs13481014	11	47930884
Congenic	rs13481033	11	54712386
Congenic	rs13481061	11	62806119
Congenic	rs3701609	11	73077126
Congenic	rs13481127	11	83226031
Congenic	rs3688955	11	90397849
Congenic	rs6386362	11	106839739
Congenic	rs3699056	11	113865002
Congenic	rs26999278	11	118385117
F2	rs3658783	6	144.26375
F2	rs3711088	6	148.26047
F2	rs3655898	9	33.48717
F2	rs13480150	9	39.531947
F2	rs13462199	9	44.444901
F2	rs13480186	9	49.989194
F2	rs3723670	9	52.322903
F2	rs3666398	9	56.434767
F2	rs3714012	9	61.00434
F2	rs6174757	9	67.986295
F2	rs13480271	9	72.676666
F2	rs13480285	9	76.421401
F2	rs13480836	11	3.4542
F2	rs13480863	11	9.19362
F2	rs29468939	11	14.997358
F2	rs3723987	11	19.368928
F2	rs3706694	11	28.31486
F2	rs6239937	11	34.833394
F2	rs6164170	11	40.117642
F2	rs13481014	11	47.930884
F2	rs13481033	11	54.712386
F2	NF1	11	79.22295
F2	BCAS3	11	85.610219
F2	rs13481371	12	30.883465
F2	rs6223000	12	34.86761
F2	rs3658100	12	38.816648
F2	rs49445146	12	44.815196
F2	rs13481436	12	49.028484
F2	rs3670749	12	54.060261
F2	rs13481496	12	64.303588
F2	rs13481561	12	83.267516
F2	rs13481565	12	85.117957
F2	rs13481599	12	97.343222
F2	rs3706601	18	77.708702
F2	rs3023496	19	48.212356
F2	rs6304326	19	53.512609

Table 2: Location and Type of QTL.

QTL	Kind	Chromosome	Mb
Proximal	Ancestral	11	0-28
Distal	Ancestral	11	40-57
NF1	De novo	11	79

Supplemental Table 1: Mutations in BcA80 Identified in Sequencing

Chr	position	ref allele	alt alleles	gene name	Chr	position	ref allele	alt alleles	gene name	Chr	position	ref allele	alt alleles	gene name
chr1	7548331	G	A	Obsl1	chr7	1.47E+08	TCCCCC	TCCCCC	Kndc1	chr10	57690783	G	A	4933403O03Rik
chr1	84755604	C	T	Trp12	chr7	1.49E+08	GAAGATC	GAAGATC	Muc2	chr10	57690841	C	T	4933403O03Rik
chr1	87481514	T	C	Sp110	chr7	3415321	T	C	Caong8	chr10	57690853	T	C	4933403O03Rik
chr1	87481584	T	C	Sp110	chr7	3415332	G	C	Caong8	chr10	62480787	T	A	Hnmph3
chr1	87488130	A	G	Sp110	chr7	3415343	T	G	Caong8	chr10	76522670	G	T	Col18a1
chr1	90257246	C	A	Trpm8	chr7	1.16E+08	T	A	Olfr498	chr10	77454832	G	C	Pkl
chr1	90257248	C	A	Trpm8	chr7	1.38E+08	C	A	Dmbt1	chr10	94306815	G	C	Plknc1
chr1	90257261	C	G	Trpm8	chr7	1.49E+08	G	T	Muc5	chr10	99991288	C	T	Cop290
chr1	90257263	C	A	Trpm8	chr7	1.49E+08	A	G	Muc2	chr11	79222950	G	A	Nf1
chr1	90257274	G	A	Trpm8	chr7	1.49E+08	A	G	Muc2	chr11	87431022	G	C	Hsf5
chr1	90257278	C	A	Trpm8	chr7	1.49E+08	A	T	Muc2	chr11	87881937	A	T	Vezf1
chr1	90257289	T	G	Trpm8	chr8	73278766	AGGG	AGG	Rab3a	chr11	97213076	C	T	Gpr179
chr1	90257295	A	G	Trpm8	chr8	73288662	TGG	TG	Ifi30	chr11	1.01E+08	C	G	Naglu
chr1	90257303	G	T	Trpm8	chr8	1.07E+08	TCC	TC	Bean1	chr11	1.02E+08	C	G	Sost
chr1	1.74E+08	G	T	Cd84	chr8	40134541	A	C	Tusc3	chr12	1.06E+08	A	G	Serpina3i
chr1	1.78E+08	C	T	Olfr429	chr8	72589034	G	A	Supp1	chr12	1.14E+08	C	T	Ahnak2
chr2	90757866	GA	GAA	Ptpmt1	chr8	75016470	T	C	Slc35e1	chr12	1.14E+08	G	C	Ahnak2
chr2	53937899	C	T	Rprm	chr8	1.2E+08	T	C	Bcmo1	chr12	1.19E+08	T	C	Dnahc11
chr2	91425969	A	G	Ckap5	chr9	7465083	TG	TGG	Mmp1a	chr12	1.2E+08	G	C	Sp4
chr2	1.19E+08	C	T	Picb2	chr9	3001063	G	A	Gm10722	chr13	43466192	TCCCC	TCCCCC	Sirt5
chr2	1.54E+08	C	T	Dnm13b	chr9	3001079	A	T	Gm10722	chr13	21480526	A	G	Zxscan3
chr3	95492913	AGGGGG	AGGGGG	Mcl1	chr9	3004907	G	T	Gm11168	chr13	53212573	C	T	Ror2
chr3	95587821	A	G	Rprd2	chr9	3004915	T	C	Gm11168	chr13	83764759	A	G	Me2c
chr4	60152637	ACAGCAG	ACAGCAG	Mup2	chr9	3004921	C	G	Gm11168	chr13	1.01E+08	A	T	Naip2
chr4	1.24E+08	CGGGGG	CGGGGG	Pou3f1	chr9	3004924	C	G	Gm11168	chr13	1.2E+08	G	T	Paip1
chr4	53033092	G	T	Nipnap3b	chr9	3005149	T	G,C	Gm11168	chr14	4453772	A	G	2610042L04Rik
chr4	59007937	C	T	Dnajc25	chr9	3005175	T	A	Gm11168	chr14	33460327	C	T	Prrxl1
chr4	99468826	G	C	Itgb3bp	chr9	3016147	T	A	Gm10720	chr14	46149672	G	T	Fermt2
chr4	1.47E+08	T	A	Gm13152	chr9	3016521	T	C/C,G	Gm10720	chr15	1.02E+08	G	T	Krt74
chr4	1.56E+08	T	C	Kihl17	chr9	3016547	T	A	Gm10720	chr16	11162771	G	A	Zc3h7a
chr4	1.56E+08	T	C	Kihl17	chr9	3016725	G	C,A/C,T	Gm10720	chr16	19947220	A	G	Kihl6
chr5	38691528	TG	T	Otop1	chr9	3016928	C	A	Gm10720	chr16	32751753	C	G	Muc4
chr5	1.12E+08	GCC	GC	Ttc28	chr9	3016931	T	C/C,G	Gm10720	chr16	32754173	G	A/A,T	Muc4
chr5	1.32E+08	AG	AGG	Aut2	chr9	3016967	C	G	Gm10720	chr16	32754177	G	A	Muc4
chr5	13570208	A	G	Sema3a	chr9	3017092	T	C	Gm10720	chr16	32754419	T	A	Muc4
chr5	33080891	G	T	Psd	chr9	3018106	T	G	Gm10719	chr16	32755762	C	G	Muc4
chr5	36891324	G	C	Otop1	chr9	3018135	T	C	Gm10719	chr16	32755763	T	A	Muc4
chr5	50354957	G	A	Gpr125	chr9	3018898	G	T	Gm10719	chr16	32755770	T	A	Muc4
chr5	1.28E+08	C	G	Ubc	chr9	3020989	A	C/C,T	Gm11167	chr16	32755776	C	A	Muc4
chr5	1.3E+08	C	A	PspH	chr9	3021080	T	C	Gm11167	chr16	32755785	T	A	Muc4
chr5	1.38E+08	A	T	Muc3	chr9	3021253	A	G	Gm11167	chr16	32755789	C	G	Muc4
chr5	1.38E+08	A	G	Gm3054	chr9	3021533	T	C	Gm11167	chr16	32755791	G	A	Muc4
chr5	1.52E+08	T	A/A,C	D730045B01Rik	chr9	3021574	T	A	Gm11167	chr16	32755803	G	C	Muc4
chr6	38613345	GCC	GCCC	Clec2l	chr9	3032777	T	C/C,G	Gm10106	chr16	36772531	A	G	Slc15a2
chr6	52210629	CCG	CCGCG	Hoxa13	chr9	3032783	C	G,A,T/G	Gm10106	chr17	47490865	ACCC	ACC	Trerf1
chr6	72259796	TGGG	TGG	Sltpb	chr9	3032786	C	G,A,T	Gm10106	chr17	68123154	TGGATGC	TG	Lama1
chr6	52210841	C	G	Hoxa13	chr9	3032802	T	C	Gm10106	chr17	15098196	T	C	Phf10
chr6	52210842	G	C	Hoxa13	chr9	3037125	T	C	Gm10715	chr17	22826838	A	G	Gm9805
chr6	1.29E+08	A	G	Gm5684	chr9	3037462	T	A	Gm10715	chr17	25101780	C	T	Cramp1l
chr6	1.35E+08	G	C	Hbbp1	chr9	3037727	C	T	Gm16889	chr17	47537359	T	C	Guca1a
chr7	3412525	CAAAAA	CAAAA	Caong8	chr9	3037857	C	A	Gm10715	chr17	68102223	G	A	Lama1
chr7	3412569	GCC	GC	Caong8	chr9	3038271	T	C	Gm10715	chr18	58038742	TGCGGCG	TGCGGCG	Slc12a2
chr7	13846247	AGGG	AGG	6330408A02Rik	chr10	5992117	C	T	Akap12	chr18	53624431	A	G	Prdm6
chr7	13891273	GCCCCC	GCCCC	Lig1	chr10	57687131	C	T	Dux	chr18	53624602	T	G	Prdm6
chr7	38979802	GCCCCC	GCCCCC	1600014C10Rik	chr10	57687374	C	A	Dux	chr19	13912910	A	C	Olfr1501
chr7	1.05E+08	TGGCG	TGGCGAC	Rsf1	chr10	57687969	A	C	Dux					
chr7	1.08E+08	A	AG	C2cd3	chr10	57690761	A	G	4933403O03Rik					

BRIDGING STATEMENT FROM CHAPTER 4 TO CHAPTER 5

In the previous chapter, we paid particular focus to chromosome 11 and the BcA80 strain of recombinant congenic mice. We identified several QTL on chromosome 11 associated with resistance to influenza infection. We further went on to identify a *de novo* mutation in the *Nf1* gene as a candidate underlying resistance to influenza infection on chromosome 11. The impact and implications of these results were then discussed.

In the following chapter we return to the study of *Pla2g7*, one of the strong candidates identified through eQTL studies earlier in the thesis as potentially underlying the chromosome 17 associated altered susceptibility of our mice to influenza virus infection. We verify the importance of *Pla2g7* on influenza resistance through knockout and congenic studies, and the results and implications are discussed.

CHAPTER 5: LACK OF LIPOPROTEIN-ASSOCIATED PHOSPHOLIPASE ACTIVITY ENHANCES SURVIVAL AFTER INFECTION WITH MOUSE-ADAPTED INFLUENZA VIRUS

Gregory A. Boivin^{1,2}, Claudia Duerr^{2,3}, Martin Brandt⁴, Earl G. Brown⁵, Robert Sladek^{1,6,7}, Jörg H. Fritz^{2,3}, Colin MacPhee⁴, and Silvia M. Vida^{1,6,7}

¹ Department of Human Genetics, McGill University, Montreal, Quebec, Canada

² Complex Traits Group, McGill University, Montreal, Quebec, Canada

³ Department of Microbiology and Immunology, Montreal, Quebec, Canada

⁴ Biological Sciences, GlaxoSmithKline, King of Prussia, PA.

⁵ Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ontario, Canada

⁶ McGill University and Genome Quebec Innovation Centre, Montreal, Quebec, Canada

⁷ Department of Medicine, Experimental Medicine Division, McGill University, Montreal, Quebec, Canada

Running title: PLA2G7 Activity Controls Susceptibility to Influenza

ABSTRACT

Influenza virus presents a major challenge to public health. In the recent years, many genetic loci have been linked to influenza susceptibility in the mouse. Identifying the underlying genes or pathways in these loci may lead to novel therapeutic interventions against infection. In previous work, we identified *Pla2g7* (encoding lipoprotein-associated phospholipase A2 Group VII or PLA2G7) as a strong possible candidate controlling influenza virus resistance linked to mouse chromosome 17. Here, we show that PLA2G7 activity is inversely correlated with survival upon challenge with a pathogenic H3N2 influenza strain. Congenic mice bred with a natural high expressing *Pla2g7* allele show decreased survival to infectious challenge compared to wild-type controls (50% versus 70% survival, $p < 0.05$) and the level of PLA2G7 activity in plasma is an early indicator of fatal disease progression. By contrast, *Pla2g7*-deficient mice demonstrated increased survival compared to wild-type littermates (95% versus 70% survival, $p < 0.05$). Accordingly, the frequency of neutrophils and inflammatory monocytes was significantly decreased in the lung after infection compared to controls. Kinetics studies using qPCR and microarrays indicated a potential shift in the resistant *Pla2g7*-deficient mice away from exacerbated cytokine signaling later in infection, towards wound repair.

INTRODUCTION

There are upwards of 50,000 seasonal influenza virus-associated deaths per year in the United States alone [90]. This mortality rate can increase during pandemics (e.g., over 20 million in 1918, ~2 million in 1958, ~1 million in 1968, [91], [313]. Primary viral pneumonias, secondary bacterial infections, and cardiovascular complications account for a significant proportion of these deaths.

Host genetic factors are important in the onset, progression, and outcome of many viral infections [228], [229], [230], [231]. This holds true for influenza virus infections. Weakened immune responses (e.g., in the immunocompromised or the elderly) can lead to unimpaired viral replication and severe disease [92]. Alternatively, the exacerbated response associated with hypercytokinemia and enhanced infiltration of immune cells in the lung can result in immunopathology and acute respiratory distress syndrome [94], [95], [96].

Genetic mapping with mouse models of infection has been used to successfully identify novel markers of influenza pathogenesis [255], [254], genes involved in the host response to influenza (e.g., *Mx1* [256], [259], [260], [261], and vaccination strategies against influenza [257]. Additionally, many QTL linked to influenza susceptibility or resistance have now been identified through mouse studies [236], [157], [159], [160], [161].

Loci for influenza resistance have been identified on mouse chromosomes 1, 2, 7, 11, 15, 16, and 17 [159], [161], [236]. To date only two regions have been resolved at the gene level; with *Hc* and *Mx1* having been shown to underlie variation in host resistance or susceptibility to influenza in several replicated crosses. Clearly, additional studies need to be undertaken to identify novel resistance genes from these QTL. With these genes identified, it will be possible to develop new potential therapies for seasonal or pandemic influenza cases.

By co-localizing expression QTL (eQTL) and clinical QTL (cQTL) in an earlier study of influenza resistance, we identified *Pla2g7* as a strong candidate underlying the resistance/susceptibility phenotype associated with our chromosome 17 locus [236]. The protein encoded by *Pla2g7* (known as PLA2G7 or Lp-PLA2) is a circulating multifunctional enzyme that travels mainly with low-density lipoprotein (LDL) and is mainly produced by myeloid derived cells [314]. PLA2G7 may be involved in many disease processes, including cardiovascular disease and asthma [270], [271], [281]. Two contrasting effects on inflammation have been described. PLA2G7 can hydrolyze oxidized phospholipids (OxPL) to produce lysophosphatidylcholine (lyso-PC) and free oxidized fatty acids leading to oxidative stress and inflammation [283], [315]. It also catalyzes the degradation of platelet activating factor (PAF) into inactive components, leading to reduced inflammation [284], [285], [286].

High levels of PLA2G7 activity were implicated in a swine model of coronary artery disease, via increasing inflammation in the necrotic cores of atherosclerotic plaques [283]. In this animal model, pharmacological inhibition of *Pla2g7* reduced inflammation and disease severity by reducing levels of proinflammatory lyso-PC [287]. In our model of influenza infection, chromosome 17 imparted susceptibility to influenza virus infection was associated with high *Pla2g7* mRNA levels, increased inflammation (including increased Kc, Mcp1 and Il6 protein expression), and decreased survival in susceptible mice.

Here, we use *Pla2g7*-deficient mice (*Pla2g7*^{-/-}) and high expressing *Pla2g7* congenic mice to dissect the involvement of this gene in the host response against a pathogenic mouse adapted strain of influenza infection (A/HK/1/68-MA20 [262]). Targeted deletion of *Pla2g7* significantly increased survival compared to wild-type controls, which in turn survived longer than congenic mice expressing even higher levels of *Pla2g7*. Further, we document significantly increased circulating PLA2G7 activity in response to mouse adapted H3N2 influenza virus infection. The level of PLA2G7 activity positively correlated with

expression of inflammatory mediators and recruitment of myeloid cells in the lung but not with viral clearance. Finally, microarray analysis suggested a model whereby lack of PLA2G7 function contributes to the balance of pro- and anti-inflammatory pathways in wound repair.

MATERIALS AND METHODS

Animals and Ethics

C57BL/6 (B6) mice were purchased from the Jackson Laboratory (Bar Harbour, Maine). As previously described, *Pla2g7*-deficient mice were bred on a pure B6 background [316] and were kindly provided by GlaxoSmithKline Pharmaceuticals Ltd (King of Prussia, PA). B6 and *Pla2g7*^{-/-} progenitors were inter-crossed to obtain *Pla2g7*^{+/-} and *Pla2g7*^{-/-} littermates for use in the present study. High expressing *Pla2g7* congenic mice were developed by transferring a chromosome 17 *Pla2g7*-containing interval from the BcA70 strain into the B6 strain using marker assisted breeding and serial back-crosses. BcA70 is a recombinant congenic mouse strain that inherited 12.5% of genes from the strain A/J and the remaining genes from the strain B6 [189] but presents short survival time during influenza infection comparable to the phenotype of their minor donor strain [236]. *Pla2g7*-congenic mice (B6.A-*Pla2g7*) in which the critical interval between rs33584971 (38.5Mb) and rs13483002 (45.8Mb) was fixed were bred to homozygosity and maintained by sib-matings. Experimental protocols were in accordance with the institutional guidelines of the Canadian Council on Animal Care. Mice were maintained at McGill University animal facilities in compliance with the Canada Council on Animal Care as regulated by the McGill University Animal Care Committee.

Genotyping

Genomic DNA was prepared from tail biopsies using Proteinase K and serial phenol/chloroform extractions followed by ethanol precipitation as previously described (35). *Pla2g7* genotypes were identified in house through PCR amplification and gel electrophoresis using primers (5'-GAA ATC CCA AGC ATC TTG TCA GA-3', 5'-GAC CTT GAT CTT GAG CTG GGT AGT A-3') to distinguish *Pla2g7* wild-type (~1034 bp) and knock-out (~328 bp) alleles.

Cell lines and virus strains

Madin–Darby canine kidney (MDCK, ATCC CCL-34) and HeLa cells (ATCC: CCL-2) were grown and maintained in Dulbecco's MEM supplemented with penicillin (100 U/ml), streptomycin (100 µg/ml), and FBS (10%). Influenza virus A/HK/1/68-MA20 (H3N2) was grown in 10-day-old embryonated hen's eggs. A/HK/1/68-MA20 is a mouse-adapted influenza virus strain derived from a H3N2 strain that was clinically isolated in Hong Kong during the 1968 pandemic, as previously reported [262]. Titration of infectious virus was determined by plaque assay as described previously [263], [317].

Infection of mice, definition of phenotype, and tissue collection

Pla2g7 was initially identified as a candidate gene in a male specific QTL for influenza susceptibility [236]. Due to this fact, only male mice have been used for all influenza infections in the present study. Mice were aged 10 to 16 weeks at the time of infection. For all influenza experiments, mice were anesthetized with isoflurane and intranasally infected with 454 pfu of A/HK/1/68-MA20 per gram bodyweight, diluted in sterile PBS to 1ul/g bodyweight. This made for a dose of 104 pfu / 22g animal. For the survival screen, mice were monitored daily for two weeks following inoculation. Clinical signs (weight loss, labored breathing, lack of grooming, and low motility) were recorded daily. Mice presenting respiratory distress or a body-conditioning score lower than 2 were humanely sacrificed. For lung viral titer and gene expression, mice were sacrificed at day 0, 1, 2, 4, or 6 post infection by exposure to CO₂, lungs were perfused with PBS via the right ventricle, excised and snap frozen in liquid nitrogen.

PLA2G7 activity assay

The assay was completed as described previously [318]. Briefly, all assays were performed at 37°C in 50 mmol/L HEPES and 150 mmol/L NaCl, pH 7.4. PLA2G7 activity was measured using either 1-decanoyl-2-(4-nitrophenyl glutaryl) phosphate (DNGP) or [³H]PAF (Cascade Biochemicals) as a substrate. Serum samples (at different dilutions) were added to 50 µmol/L DNGP in buffer at 37°C. The absorbance increase was followed at 400 nm, using either a diode array

spectrophotometer (Hewlett-Packard) or a 96-well plate reader (Molecular Devices, T_{max}) running in kinetic mode. Product was quantified using the extinction coefficient, $\epsilon_{400}=15\,000\cdot\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$. For PAF-AH activity, [³H]PAF and sample were incubated in a final volume of 200 μL for 10 minutes at 37°C. The reaction was stopped by vortexing with 600 μL of $\text{CHCl}_3/\text{MeOH}$ (2:1), and the CHCl_3 and aqueous layers were separated by centrifugation. The aqueous layer was removed (250 μL) and vortexed with 250 μL of CHCl_3 . The aqueous layer was again removed and the [³H]acetate determined by scintillation counting. Protein was determined using the Pierce bicinchoninic acid assay kit, according to the manufacturer's instructions.

Histology and image analysis

Lung samples were sent to the McGill histology Facility. Tissues were embedded in paraffin blocks, excised in 5 μm thick sections, and stained with hematoxylin and eosin (H&E) for light microscopic examination. Sections were analyzed using custom color recognition software designed to identify whitespace (open airspaces); hematoxylin-associated nuclei and eosin-associated collagen; muscle fiber; and proteinaceous infiltrates. Briefly, each pixel in the histological image was filtered into one of three preset binary conditions (white, blue, and red) based on 8-bit color space.

Quantitative PCR

Total RNA was extracted using Trizol (Invitrogen) and transcribed into cDNA using M-MLV with random hexamers (Invitrogen) according to the manufacturer's instructions. Real-time quantitative PCR (qPCR) was performed using Platinum SYBR Green SuperMix-UDG (Invitrogen) together with experimental or control primers. Experimental primers targeted a panel of cytokines and chemokines and were designed to span exon junctions with the help of primer3 (Table S1). Target transcripts were normalized to hypoxanthine phosphoribosyltransferase 1 (*Hprt*). Samples were run in duplicate with three to five mice per condition. Reactions were performed using the Step One Plus real time PCR system (Applied

Biosystems), and expression was analyzed using StepOne v2.2.2. Relative mRNA expression was shown as $2^{-\Delta\Delta CT}$ for cytokine or $2^{-\Delta CT}$ for viral comparisons.

Fluorescence-activated cell sorting

To prepare cells for fluorescent sorting, lungs were perfused with 10 mL PBS, cut to small pieces and digested in RPMI/5%FBS/0.5 mg/mL Collagenase Type IV (Sigma) and 0.02 mg/mL DNase I (Roche) for 1 hr at 37°C. The lung pieces were then disrupted using a 10 mL syringe with an 18G1½ needle and filtered using a 100 µm cell strainer. The cell suspension was washed once with FACS buffer (PBS/2%FBS) and then stained for FACS analysis. Fc receptors were blocked using 2.4G2 supernatant. Then, the following antibodies were used for cell surface staining for 30 min on ice: CD45-eFluor 450 or PerCPCy5.5 (clone 104), Siglec-F-PE (clone E50-2440), Ly6C-FITC (clone AL-21), CD11b-APC (clone M1/70), Ly6G PerCP-Cy5.5 (clone 1A8), F4/80-PE-Cy7 (clone BM8), CD11c-APC-eFluor780 (clone N418), NK1.1-PE (clone PK136), CD8a-FITC (clone 53-6.7), CD19-APC (clone eBio1D3 (1D3)), CD3e-eFluor450 (clone 145-2C11), CD4-PE-Cy7 (clone GK1.5), B220-APC-eFluor780 (clone RA3-6B2), CD103-PE (clone 2E7), MHCII-FITC (clone M5/114.15.2). Dead cells were identified by staining with the LIVE/DEAD Fixable Aqua Dead Cell Stain Kit (Invitrogen) accordingly to the manufacturer's instructions. Cells were acquired on a Canto II and analyzed with FlowJo software (TreeStar).

Microarray Analysis

Total RNA was extracted using Trizol (Invitrogen) and mRNA was purified from these extractions using the RNeasy MinElute Cleanup Kit (Qiagen). Quality control with Agilent Bioanalyzer (Agilent Technologies), cDNA synthesis, and microarray hybridization was performed at the McGill University and Genome Quebec Innovation Centre. RNA from three *Pla2g7*^{-/-} and *Pla2g7*^{+/+} male mice on day 0, 1 and 6 were analyzed on Illumina Mouse-WG6 version 2.0 Expression Beadchips (Illumina). The raw data was background corrected via negative controls, quantile normalized, and log₂ transformed. Significantly expressed

probes were fit with a linear model using contrasts to compare the relevant groups of interest. Significance levels were adjusted for multiple testing. The statistical analyses were performed using the LIMMA package [319] in Bioconductor, an R-based open-source software development project in statistical genomics [320]. Pathways that were potentially dysregulated between the microarray conditions were identified using Ingenuity Pathway Analysis.

RESULTS

Increased PLA2G7 activity is associated with decreased survival during influenza virus infection.

In a previous screen, our group used the AcB/BcA recombinant congenic strains (RCS) of mice to identify a sex-dependent QTL linked to influenza susceptibility on chromosome 17. We have since created congenic mice from one of the informative RCS mice, in which the critical ~38.5-45.8Mb interval on chromosome 17 from the A/J origin was fixed onto the C57BL/6 (B6) genetic background (Fig. 1). The congenic B6.APla2g7 strain was initially selected with markers spanning the proximal region of chromosome 17 to exclude the major histocompatibility complex (Fig. 1, Supplemental Table 1). We also obtained *Pla2g7*^{-/-} mice from GlaxoSmithKline, which have been described previously [316].

Baseline PLA2G7 activity levels were recorded for each of the mouse lines. Taking C57BL/6 levels as a 100% control, the *Pla2g7*^{-/-} mice had next to non-existent levels of activity, while the *Pla2g7* congenic mice had ~140% control of PLA2G7 activity at baseline, confirming their knockout and overproducing status (Fig. 2A). *Pla2g7*^{-/-}, control B6, and high expression *Pla2g7* congenic mice were tested for survival with ~10⁴ pfu of the pathogenic A/HK/1/68-MA20 mouse adapted strain of influenza virus. Knockout mice showed reduced mortality, while the congenic over-expressors showed increased mortality as compared to wild type controls (Fig. 2B, *p* < 0.0001).

To test whether increased PLA2G7 levels could be predictive of morbidity in a population of mice, we infected 60 B6 control mice with either 10⁴ pfu / 22g animal (50 mice) or PBS control (10 mice) and followed them for 14 days. Blood was collected from all mice at day 1, day 3, and day 6 post-infection. Moribund mice that succumbed to infection showed elevated levels of PLA2G7 activity by day 3 as compared to mice that survived the infection and mice infected with

PBS, illustrating the potential for PLA2G7 activity as an early biomarker of severe influenza infection. By day 6 post infection, PLA2G7 activity levels were elevated in both moribund mice and mice that survived infection as compared to PBS controls (Fig. 2C, ***p < 0.001, **p < 0.01).

Histological Analysis shows increased nuclei in lung tissue of *Pla2g7*^{-/-} mice at baseline.

In an attempt to identify any gross morphological differences between the more resistant *Pla2g7*^{-/-} and control mice, we analyzed histological sections of lung tissues from mice at baseline and day 7 post infection. While no major differences were identified by day 7 post infection, an increase in eosin associated red color space was identified at baseline (Fig. 3A, B). This finding could be indicative of increased proteinaceous material, collagen, or muscle fibers in the slides of resistant mouse lungs at the baseline.

Resistant *Pla2g7*^{-/-} mice show altered cytokine kinetics

Pla2g7^{-/-} mice show increased levels of pro-inflammatory mediators by day 1 post infection (e.g., IL-6, A, C, D, E, F). Interestingly, by day 6 post infection this trend had reversed, with the resistant *Pla2g7*^{-/-} mice having lower levels of KC (p < 0.05, Fig. 4A) and a trend towards lower *Mcp-1* and *Tnfa* (Fig. 4D, E). These lower cytokine levels at day 6 post infection were mirrored by lower expression levels of the influenza viral matrix gene M1, implying that the resistant *Pla2g7*^{-/-} mice may be better able to mount a productive immune response and resolve influenza infection quicker than wild-type controls.

To identify if these changes were also seen in immune related cell populations, we used flow cytometric analyses from lung tissue on day 1 and 6 post infection. While there was a trend towards increased numbers of monocytes in the resistant *Pla2g7*^{-/-} mice on day 1, the result was not significant (data not shown). We did, however, identify significantly decreased populations of neutrophils and inflammatory monocytes at day 6 post infection (Fig. 4G). While there was a

trend towards fewer cells overall in the resistant *Pla2g7^{-/-}* mice, this was not significant (Fig. 4H).

Microarray analysis Identifies differential expression of genes at Day 1 and Day 6 post infection.

Homogenized whole lung tissue from resistant *Pla2g7^{-/-}* and wild-type control mice from days 0, 1, and 6 post infection were collected and used for gene expression profiling with microarrays. Principal component analysis (PCA) showed reproducibility of biological replicates, and an effect of infection and time-course on global gene expression changes (Supplemental Fig. 1). After background correction and normalization, ~43% of the probes were dropped from future analyses due to below threshold expression levels, leaving 26,034 probesets to be analyzed for differential expression.

Expression differences were identified by subtracting expression levels from either day 1 or day 6 from the baseline for each mouse strain and comparing those differences. A linear model was built and a contrast matrix was defined to determine significantly differentially expressed genes. Remarkably, on day 1 post infection, only two probesets were identified as differentially expressed, *Cyr61* and *Trim72*. On day 6 post infection, 271 genes were identified as significantly differentially expressed (Fig. 5, Supplemental Table 2). The findings were analyzed with Ingenuity Pathway Analysis software. Nineteen different pathways were significantly altered on day 6 post-infection (Table 1). Many interesting related pathways were found through this analysis, including Rho signaling (i.e., RhoGDI signaling, RhoA signaling, and signaling by Rho family GTPases) and tight junction signaling (i.e., Tight junction signaling, Agrin signaling at Neuromuscular Junction, Epithelial Adherens Junction Signaling). Interestingly, the genes implicated in these pathways overlap to a high degree. *Acta2* (actin, alpha 2, smooth muscle), *Actc1* (actin, alpha, cardiac muscle), *Myl7* (myosin, light chain 7), and *Itgb1* (integrin, beta 1) comprise differentially regulated genes in each of the pathways, implicating a potential role of cytoskeletal or integrin

related components in the differential response of resistant *Pla2g7^{-/-}* mice later in infection.

DISCUSSION

Pla2g7 was identified as a potential candidate underlying susceptibility to influenza virus linked to a chromosome 17 locus [236]. To test whether the gene has a direct effect on resistance to pathogenic influenza infection, we tested *Pla2g7*^{-/-} mice, control mice, and congenic mice that carry a naturally high expression allele of the *Pla2g7* gene.

Pla2g7 encodes a phospholipase that is known to hydrolyze phospholipids into fatty acids and other lipophilic molecules. The activity of PLA2G7 is multifunctional, with contradicting proposed effects on inflammation. PLA2G7 can hydrolyze oxidized phospholipids (OxPL) producing lysophosphatidylcholine (lyso-PC) and free oxidized fatty acids, leading to increased inflammation [283]. Alternatively, it can catalyze the degradation of platelet activating factor (PAF) into inactive components, leading to reduced inflammation [284], [285]. Pharmacological inhibition of PLA2G7 has been shown to reduce inflammation and disease severity in a model of coronary atherosclerosis in diabetic and hypercholesterolemic swine [287]. These results were consistent with our original hypothesis, which stated that reducing PLA2G7 activity would reduce the negative damaging effects of hyper-inflammation across influenza infection, and result in increased resistance to the virus.

We found that mice lacking the *Pla2g7* gene showed effectively no PLA2G7 protein activity, and significantly increased survival on challenge with influenza virus infection. In contrast, mice with a natural variant that increased *Pla2g7* expression showed increased PLA2G7 activity levels and increased susceptibility to influenza virus challenge. Interestingly, *Pla2g7*^{-/-} mice showed a trend towards higher levels of inflammation early in infection (e.g., *Il-6*, day 1 post infection), and more subdued levels of inflammation later on during infection (e.g., *Kc*, day 6 post infection). This coincided with fewer neutrophils and inflammatory monocytes being recruited to the lung later in infection (e.g., day 6 post infection).

Lower levels of cytokines and reduced number of inflammatory cells have been seen in *Pla2g7*^{-/-} mice after myocardial infarction [321], resulting in an altered balance between inflammation and repair, and ultimately more efficient healing. This similar finding could indicate that the mechanism involved in increased resistance may potentially be generalized to more than pathogenic influenza infection. There was further evidence for altered lung baseline lung histology in the resistant *Pla2g7*^{-/-} mice. More analyses needs to be done to understand exactly what is underlying this difference and whether it can explain the altered inflammatory signaling seen across infection.

We performed microarray analyses on resistant *Pla2g7*^{-/-} mice and controls and found that remarkably fewer differences were apparent between resistant *Pla2g7*^{-/-} and susceptible control mice on day 1 post infection. Only *Cyr61* and *Trim72* were identified as differentially regulated. *Cyr61* encodes a secreted protein that is involved in integrin signaling and adhesion of endothelial cells. It has also been shown to play a role in cell proliferation, angiogenesis, differentiation, apoptosis, and extracellular matrix formation [322], [323], [324], [325]. Interestingly, it also appears to play a role in wound healing through the up-regulation of *Il-6* [326]. TRIM72 is a muscle specific protein that also has a main role in repair and has specifically been identified as promoting repair in alveolar epithelial cells after in vivo injury [327].

By day 6 post infection, 271 genes were differentially expressed between resistant *Pla2g7*^{-/-} mice and susceptible controls. Pathway analysis identified Rho related signaling and signaling at junctions among others. Rho has been identified as a key regulator of neutrophil hyper-responsiveness and influenza disease severity after infection [328]. Interestingly, several integrins (*Itga3*, *Itgb1*) and actin / myosin related genes (*Acta2*, *Actc*, *Myl7*) were consistently identified as members of the differentially regulated pathways. Integrin signaling has been associated with altered resistance to influenza infection through its action on

TGF β [329], [330] and on transepithelial migration of monocytes after influenza infection [104].

The microarray data overall seem to point to a shifting of the negative inflammatory response so often associated with increased susceptibility in pathogenic influenza infection to a response more closely associated with tissue repair through the initial upregulation of *Cyr61* and *Trim72*, and later through the action of integrins and other junction associated signaling pathways. More research needs to be done to identify the exact mechanism of action that underlies the resistance associated with reduced PLA2G7 activity.

This study represents more intriguing evidence pointing to parallels in healthy immune response to both infectious disease and to cardiovascular disease states. PLA2G7 is known for its effects on phospholipids, and has been studied for its effects in both myocardial infarctions and atherosclerosis. In these conditions, immune responses participate in progression and outcome of disease [331]. In either disease state, the evidence appears to implicate a healthy response with reduced inflammation and increased tissue repair. More research on the effects of reduced PLA2G7 activity in pathogenic disease states should be investigated.

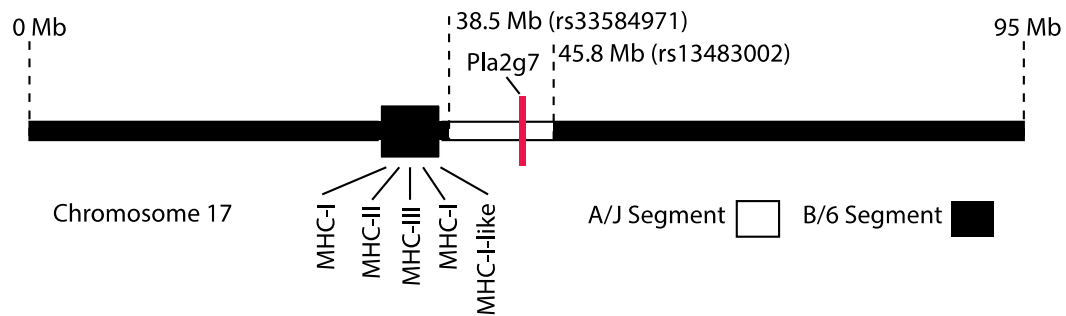


Figure 1: Haplotype structure of the congenic segment of chromosome 17 fixed in B6.APla2g7 mice. The position of the A/J segment inherited from susceptible BcA17 mice bearing a natural high expression allele of *Pla2g7*. The localization of the major histocompatibility complex (MHC) containing MHC class I, II, III and MHC-like molecules is indicated. The congenic region encompasses the critical chromosome 17 interval controlling male-survival during influenza virus infection as defined by Boivin et al (2012). Physical positions of markers used to select congenic B6.APla2g7 are shown in Mb (Build 38).

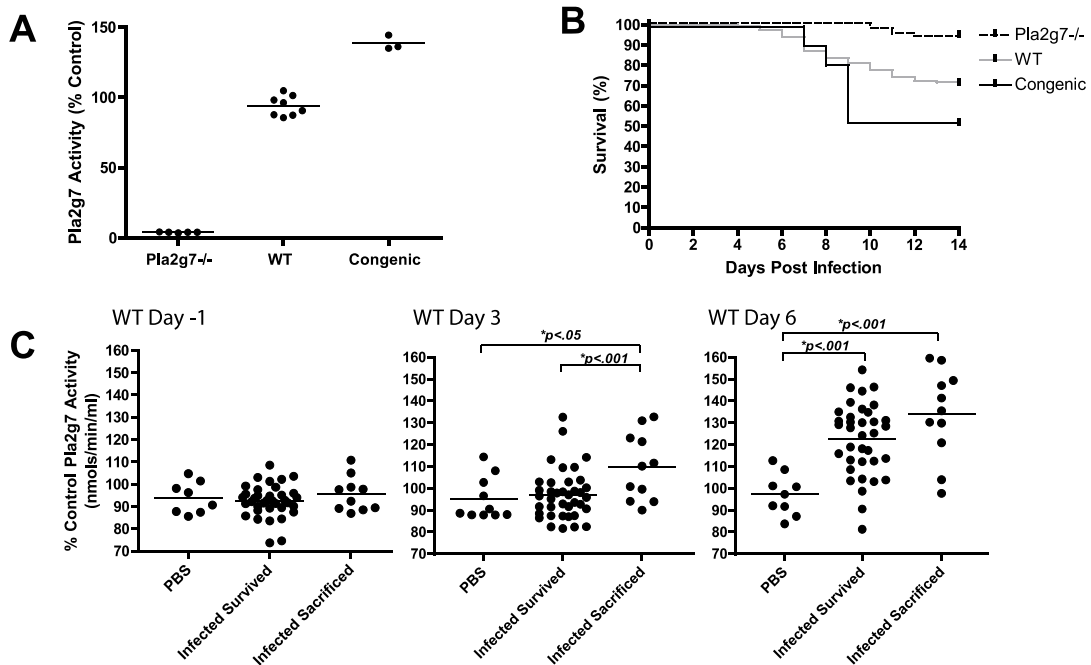


Figure 2: Increased PLA2G7 Activity is Associated with Susceptibility to Influenza Virus. Altered PLA2G7 Activity of Pla2g7^{-/-} mice, B6 mice, and overexpressor *Pla2g7* congenic mice was confirmed (A, $p < 0.001$, ANOVA). Results from 76 Pla2g7^{-/-}, 116 control B6 mice, and 21 high expressor *Pla2g7* congenic mice were tested for survival across 14 days (B, $p < 0.0001$, Kaplan-Meier). 60 B6 mice were infected with either 10^4 pfu / 22g animal (50 mice) or PBS control (10 mice) and followed for 14 days. Blood was collected from all mice at day -1, day 3, and day 6 post infection. Moribund mice that succumbed to infection showed elevated levels of PLA2G7 activity by day 3 as compared to mice that survived the infection and mice infected with PBS, illustrating the potential for PLA2G7 activity as an early biomarker of severe influenza infection. By day 6 post infection, PLA2G7 activity levels were elevated in both moribund mice and mice that survived infection as compared to PBS controls (C, *** $p < 0.001$, ** $p < 0.01$, t-test).

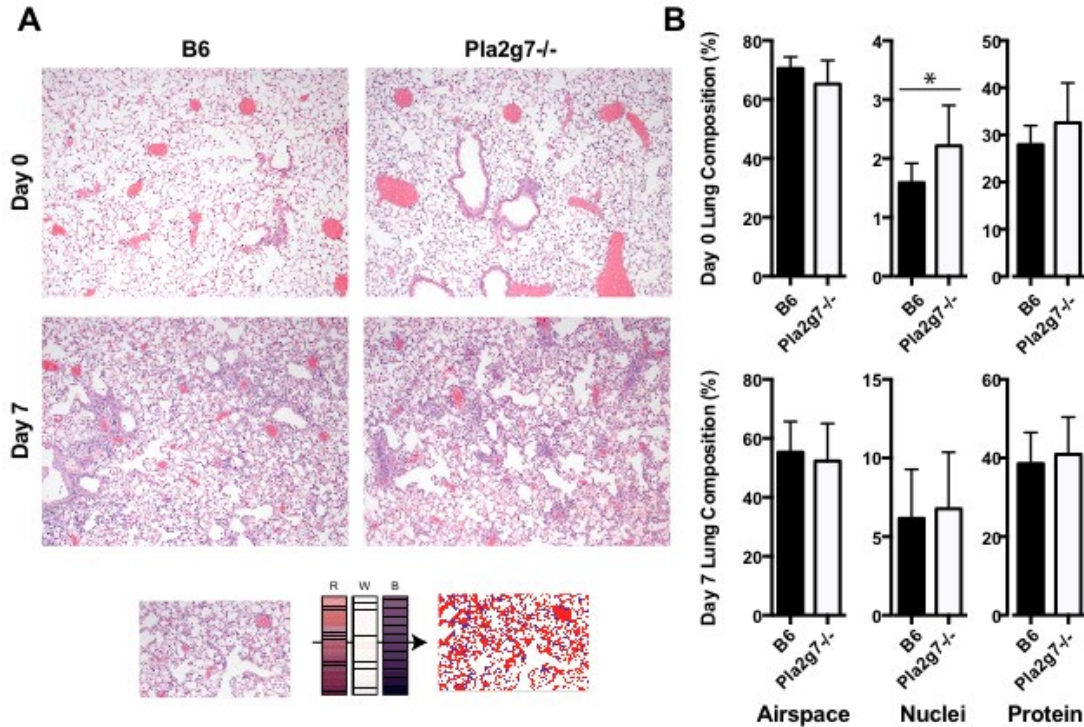


Figure 3: Histological analysis of resistant *Pla2g7^{-/-}* and wild-type controls.

H&E-stained lung sections are shown for *Pla2g7^{-/-}* and wild-type mice (A). Distribution into one of three preset binary conditions (white, blue, and red) of each pixel in digitized H&E slides identified a significant increase in eosin-associated collagen, muscle fiber, and proteinaceous infiltrates in *Pla2g7^{-/-}* mice (B, groups of 10-16 male mice, $p < .05$, t-test).

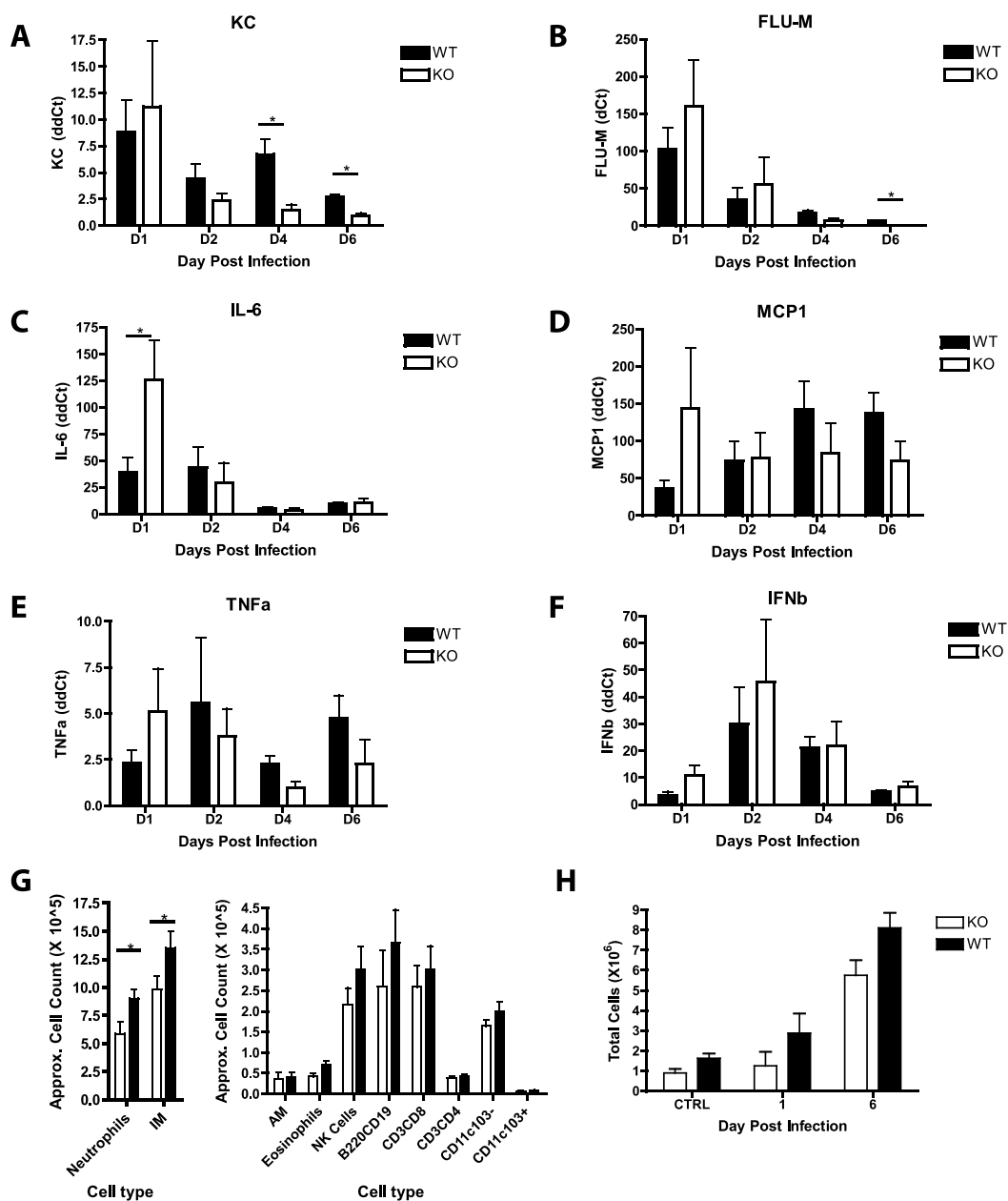


Figure 4 (legend appears on the following page)

Figure 4: Resistant *Pla2g7^{-/-}* mice show elevated cytokine signaling on day 1 post infection, but reduced cytokine levels by day 6 post infection as compared to B6 controls. *Pla2g7^{-/-}* mice show increased cytokine levels day 1 post infection (A,C,D,E,F) but reduced cytokine levels and virus by day 6 post infection (A-E, t-test). *Pla2g7^{-/-}* mice show reduced levels of neutrophils and inflammatory monocytes by day 6 post infection by flow cytometry (G, $p < .05$, t-test). No differences in overall cell numbers were seen in control mice or mice at day 1 or day 6 post infection (H). (* $p < 0.05$, t-test). 3-5 male mice per group.

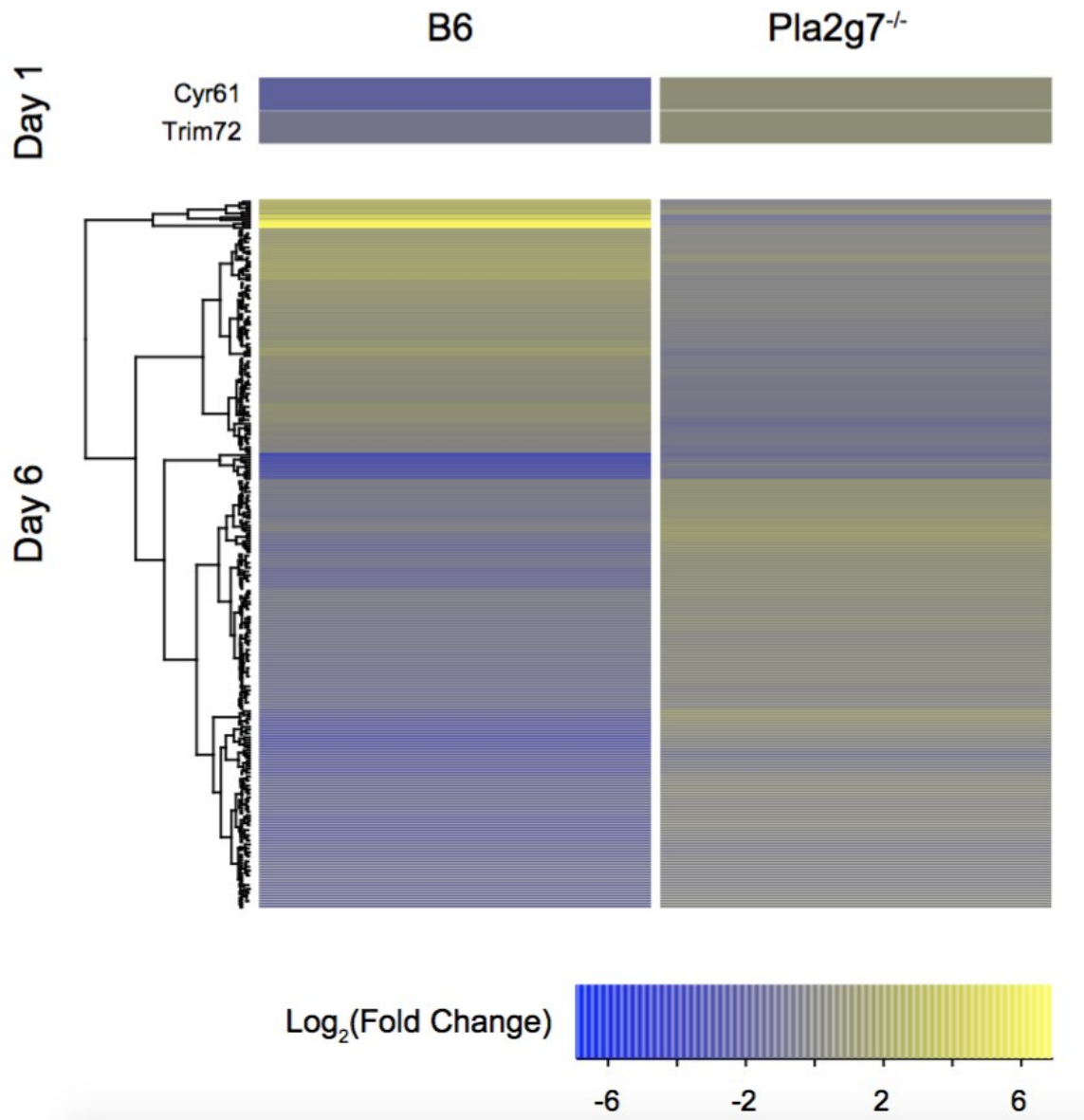
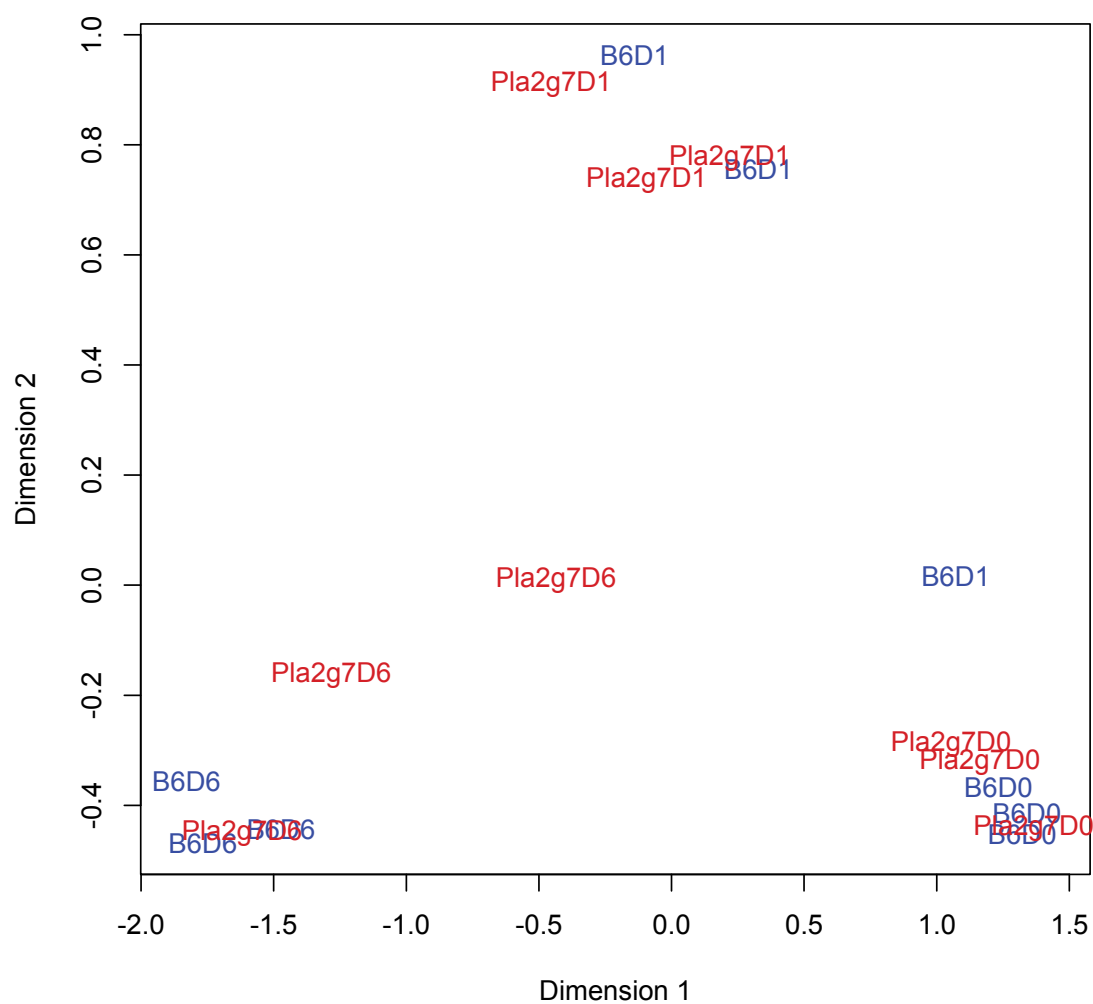


Figure 5: Dysregulated gene expression between *Pla2g7*^{-/-} and control mice. Microarrays were run on resistant *Pla2g7*^{-/-} and wild-type B6 mice at baseline, day 1, and day 6 post infection. Genes that are differentially upregulated on days 1 and 6 are shown. Only 2 genes were differentially upregulated on day 1 post infection, *Cyr61* and *Trim72*. 271 genes were differentially upregulated on day 6 post infection. The left side of the table shows a dendrogram representing the similarity in expression signatures between genes for day 6.



Supplemental Figure 1: A graph illustrating the multidimensional scaling of microarray samples.

Table 1: Significantly Altered Pathways by IPA on day 6 post-infection.

Pathway	# of significant genes in pathway	total # of genes in pathway	pvalue
ILK Signaling	9	169	4.96E-04
RhoGDI Signaling	9	164	3.99E-04
Cell Cycle: G1/S Checkpoint Regulation	5	57	1.02E-03
Tight Junction Signaling	8	155	1.24E-03
Agrin Interactions at Neuromuscular Junction	5	65	1.85E-03
Germ Cell-Sertoli Cell Junction Signaling	7	147	3.89E-03
Epithelial Adherens Junction Signaling	6	135	1.02E-02
Signaling by Rho Family GTPases	8	224	1.14E-02
Leukocyte Extravasation Signaling	7	185	1.31E-02
Tec Kinase Signaling	6	147	1.50E-02
Cyclins and Cell Cycle Regulation	4	67	1.79E-02
RhoA Signaling	5	114	1.94E-02
NRF2-mediated Oxidative Stress Response	6	163	2.36E-02
Virus Entry via Endocytic Pathways	4	82	2.51E-02
Phospholipase C Signaling	7	212	2.55E-02
Phosphatidylcholine Biosynthesis I	2	7	3.66E-03
Phospholipases	3	51	3.16E-02
Paxillin Signaling	4	94	3.87E-02
Integrin Signaling	6	188	4.28E-02

Supplemental Table 1: Chromosome 17 Markers

SNP	Chr	Position (Mb)
rs13482845	17	4.6
rs3696835	17	22.8
rs6397584	17	27.6
rs3693494	17	29.9
rs3672987	17	33.2
rs6298471	17	36.7
rs33584971	17	38.5
rs13482997	17	43.9
rs13483002	17	45.8
rs13483055	17	60.4
rs6288047	17	91.9

Over the course of the thesis, we created and presented a new marker map for the RCS mice and introduced a more effective statistical method to find novel regions of linkage in our panel of inbred mouse strains (EMMA). While studying ectromelia virus, we verified a strong QTL. Further, we identified an extremely strong candidate gene (*H2-T23*) through eQTL studies within our region. We went on to verify the genes involvement in the susceptibility of mice to ectromelia virus through complementation studies.

While studying influenza virus, we identified 2 QTL associated with susceptibility (on chromosomes 2 and 17), and several QTL on chromosome 11 associated with resistance to infection. We confirmed the likelihood that the hemolytic complement (*C5*) gene on chromosome 2 is the key molecular player at this susceptibility locus, although we agree that future analysis of the chromosome 2 region for potential modifier loci is warranted. We further identified a novel susceptibility gene, *Pla2g7* on chromosome 17, through eQTL analyses. By studying knockout and congenic strains of mice, we verified the importance of *Pla2g7* on influenza resistance. While concentrating on the BcA80 mouse strain, we further identified a *de novo* mutation in the *Nf1* gene as a candidate underlying resistance to influenza infection on chromosome 11. The impact and implications of these results will now be brought up in a general discussion.

GENERAL DISCUSSION

Overview of findings

Over the course of this thesis, we identified novel genes and QTL underlying resistance to multiple infectious diseases using the AcB/BcA panel of recombinant congenic mice. We chose to study the host response against Influenza virus and Ectromelia virus. Influenza virus causes 250,000 deaths annually worldwide and is the only known pathogen to cause recurrent pandemics (e.g., 20 million deaths in 1918; 2 million deaths in 1957; 1 million deaths in 1968). Ectromelia virus is a mouse model of smallpox, a disease that has killed more individuals than all other infectious diseases combined, and whose viral mechanisms have yet to be fully understood.

We proposed that the use of inbred mice will serve to clarify disease mechanisms and identify target genes that impact host resistance against the viruses we studied. We further proposed that since the use of recombinant congenic strains offer better QTL localization and increased availability of expression data for target tissues compared to standard mouse crosses, this would be an excellent platform to dissect complex trait divergence in A/J and C57BL/6 inbred mouse strains, provided that population structure and other confounding variables are taken into account.

In the study of the ectromelia virus, the mouse model of smallpox, our statistical analysis of the RCS identified regions of chromosome 7 and 17 as potential susceptibility loci. The chromosome 17 cQTL (~29-41 Mb) was confirmed in a secondary F₂ cross. Using microarray expression data from livers of AcB/BcA mice, we identified *H2-T23*, the gene encoding Qa1^b, as an eQTL within our chromosome 17 cQTL, making it a strong candidate for further study. Qa1^b is a non-classical major histocompatibility complex (MHC) molecule that is recognized by receptors present in both CD8 T-cells and NK cells. Additional infection studies of *H2-T23*-deficient mice confirmed the likely involvement of *H2-T23* in resistance to ectromelia virus in our model. This work provides a new

mechanism of the differential susceptibility of hosts to poxvirus infections. Our work also resulted in the creation of multiple congenic and F₁ mouse lines that isolated *H2-T23* from susceptible A/J mice. These new mouse lines may provide the groundwork for future studies outlining the interaction of Qa1^b with NK cells and CD8⁺ T cells. Future work must be done to understand how polymorphisms between A/J and B6 in *H2-T23* alter the interactions of infected cells to both NK and T-cell populations.

In the Influenza virus study, we identified cQTL on chromosomes 2 and 17. The chromosome 17 cQTL (~37-48 Mb) was confirmed in a secondary F₂ cross. Using microarrays from lungs of AcB/BcA mice, we identified *Pla2g7* as an eQTL within our chromosome 17 cQTL and *Hc* (encoding hemolytic complement C5) as a candidate within our chromosome 2 cQTL. *Hc* had previously been identified as significant in susceptibility to influenza infection. *Pla2g7* had previously been identified as involved in other inflammatory diseases (e.g., atherosclerosis). Our work associated increased *Pla2g7* expression with increased levels of pro-inflammatory cytokines. We tested congenic mice overexpressing *Pla2g7* and knockout mice lacking functional *Pla2g7* alongside controls and concluded that higher *Pla2g7* gene expression levels were significantly associated with influenza susceptibility. Time-course expression studies, with microarrays and qPCR, associated reduced levels of *Pla2g7* with lower levels of pro-inflammatory cytokines later in infection, and an increased expression of genes associated with tissue repair.

We also identified the BcA80 mouse strain as hyper-resistant to influenza infection in our study of the AcB/BcA panel. The study of the BcA80 mouse strain lead to three new cQTL on chromosome 11 that are associated with resistance to influenza virus. Congenic mouse lines created from CSS11 mice validated two of these cQTL. Co-localized eQTL with the two validated ancestral A/J derived cQTL were identified, providing strong functional candidates within these regions of linkage. Additionally, whole genome sequencing identified *Nf1* as a candidate for underlying the resistance of the third chromosome 11 cQTL. More research is

needed to fully dissect the complex resistance to influenza virus associated with mouse chromosome 11 in our model. However, the congenic mouse lines that we created combined with the eQTL co-localization studies and whole genome sequencing studies we undertook provide excellent foundations for further study.

Together, these studies show that the AcB/BcA mouse panel can be a powerful tool to dissect complex genetic phenotypes. This mouse panel provided smaller cQTL intervals with baseline eQTL information from tissues of interest, leading to the identification of potential candidate genes involved in susceptibility to globally relevant diseases and pathogens.

Advantages of the RCS

The mouse panel that we chose to use in our studies has several distinct advantages over traditional mouse F_2 populations. The panel is created from only two progenitor strains, A/J and C57BL/6. While this restriction does reduce genetic diversity in our panel as compared to some others, it also allows us to more easily identify the genetic loci underlying susceptibility to viral infection in our setting. The simplicity of the system can aid in breaking down a complex trait like resistance to viral disease. Another advantage of this panel is that each line, bred to near homozygosity, could be tested multiple times for multiple phenotypes. The increased breadth of information available through this multiple testing can increase the mechanistic understanding of a QTL by more easily identifying disease-related endophenotypes associated with a particular QTL, and by providing a head start in the creation of congenic mice for further study. We took advantage of multiple phenotypes tested on the RCS, including previously conducted expression studies and cis-eQTL analysis, to more efficiently identify candidate genes in our QTL intervals. A similar additional advantage of the RCS is that more mice can be tested per genotype and per phenotype, reducing the overall variability in the panel while tightening the relationship between genotype and phenotype of interest.

Many of the genetic regions of linkage for influenza resistance from mouse

studies have come from the use of the B×D (i.e., C57BL/6DBA/2) recombinant inbred panel. For example, regions on chromosomes 5 and 19, as well as on chromosomes 2, 16, and 17 were linked to weight loss and survival after infection with the laboratory strain A/PR/8/34 [157]. Regions on chromosomes 2, 7, 11, 15, and 17 were found when screening the panel with a highly pathogenic H5N1 influenza strain [159]. Additional work with the BXD panel also identified a QTL controlling cytokine levels (e.g., TNFa, IFNa, and CCL2) on chromosome 6 [160].

One advantage that the AcB/BcA panel of mice has over the BXD panel or a traditional F₂ population is the reduced size of genetic linkage. This reduced interval size can simplify the identification of causal genes or polymorphisms, especially in the context of complex traits. The AcB/BcA panel has been used to study the genetic correlates of many infectious, inflammatory, metabolic, and neoplastic phenotypes [191], [192], [193], [194], and has also been used to generate expression datasets for both lung [195] and liver (our ECTV study) available for analysis.

Analysis of the RCS

Over the course of our influenza and ectromelia virus studies, we were able to infect 29 strains of RCS for each infection model. For ectromelia virus, only 3 of 29 (BcA69, BcA70, and BcA74) showed highly deviant phenotypes (i.e., these 3 strains are ~87.5% B6 genetically, but show survival times that are closer to A/J than B6). Upon challenge with influenza virus, there was a gradual increase in susceptibility scores (or decrease in survival times) across the entire panel of 29 strains. Published research has shown that most viral infections are complex in nature. That is, multiple genes control susceptibility [5]. Our influenza studies confirm this finding, with multiple QTL being identified both originally and within our secondary studies of the BcA80 RCS strain. Interestingly, in our ECTV studies, the strain distribution suggests that the susceptibility in the AcB/BcA panel is controlled by a single genetic locus. This does not imply, however, that the nature of susceptibility to ECTV is not complex.

It may be that the strength of our main effect is much higher than others in the dataset, effectively masking true genetic effects underlying the phenotypic distribution of viral susceptibility in the RCS panel. To attempt to uncover any potential masked effects, the strain distributions could be examined using the genotype at our major locus as a covariate. Any significant secondary peaks could then be further examined, potentially through additional secondary F_2 crosses. Alternatively, it may be that the constrained diversity inherent in the RCS panel may be working to its advantage in simplifying the complexity identified in the AcB/BcA model. The A/J and C57BL/6 strains are both inbred lines of mice, significantly limiting the potential diversity in the RCS cross as compared to crosses containing many mouse lines (e.g., heterogeneous stock or collaborative cross mice). It is possible that the diversity of our model did not allow for the representation of the genetic differences underlying some previously identified QTL. Lastly, methodological differences between our study and previous ECTV studies may have led us to identify only chromosome 17 QTL, with no evidence of linkage on chromosomes 1, 2, or 6.

Strain × Sex interactions

While studying the influenza virus, we found two significant sex × strain interactions in the dataset. Females were significantly more susceptible than males to the identified BcA72 strain, and males were significantly more susceptible than females to the identified BcA70 strain. Previously documented sex-specific genetic and hormonal differences may contribute to the severity of influenza and the clearance of viral infection [247]. Interestingly, around the time that my work began, most research studies of immune responses did not stratify by sex [332]. Generally, androgens, including dihydrotestosterone and testosterone, suppress the activity of immune cells [333]. The effect of estrogens is more complex: low doses have been found to enhance proinflammatory cytokine production, while high doses reduce it [334].

Past studies have linked female sex hormones to both the expression of *Hc* [280] and *Pla2g7* [335], [336], [337], potentially explaining the results of our initial RCS

influenza screen. We found that our chromosome 2 peak was stronger in females than in males. The susceptibility linked to the locus can likely be attributed to the effect of the premature stop codon in C5. There may, however, be other mechanisms involved in the susceptibility linked to this locus. More research is warranted to isolate C5 from the other potential nearby modifier loci, and to identify the exact nature of the sexual dimorphism of the locus. The chromosome 17 QTL, which we can likely attribute to the effect of elevated *Pla2g7* levels, was stronger in males than in females. In both cases, the published effects of sex hormones on levels of both C5 and *Pla2g7* are in line with our results. Estrogens have been shown to generally decrease the effect of both C5 (high levels beneficial) and PLA2G7 (high levels detrimental), potentially explaining the increased effect of our chromosome 2 QTL and the decreased effect of our chromosome 17 QTL in females.

Evolution in Statistical analysis to mixed model analyses

In our primary QTL analyses, we used a mixed model to identify 2 main peaks in each of our studies. In our ectromelia study, we identified susceptibility loci on chromosomes 7 and 17. In our influenza study, we identified chromosomes 2 and 17. Both of the loci identified in our influenza study showed evidence of being sex specific. The chromosome 2 locus was stronger in females than in males in the RCS screen. The chromosome 17 locus was stronger in males than in females in the RCS screen. Secondary F₂ crosses were performed to validate these QTL. In both cases, the chromosome 17 QTL was validated, while chromosome 7 for ectromelia, and chromosome 2 for influenza were not. There could be many reasons that could account for this lack of validation. In the case of the influenza chromosome 2 locus, it may have been a sex related effect that did not show up as well in our male only population. In the case of chromosome 7 for ectromelia virus, the genetic distribution of the RCS mice at the chromosome 7 and 17 loci was identical. The chromosome 7 peak was only identified because we did not have the power available in the RCS to resolve a difference between the two peaks (i.e., there were no RCS mouse strains that contained an informative

genetic segment for either the chromosome 7 or 17 region, they either had both, or none). It is for this reason that it is incredibly important to test as many RCS strains as possible during a study to increase the power of detecting increasingly smaller genetic regions of linkage.

In the case of the Influenza virus study, there could have been a sex effect involved in the lack of validation for our initial chromosome 2 QTL. Initially, due to a major sex effect present in our study, we analyzed our initial RCS population separately, in females, and in males. Our chromosome 2 QTL for influenza susceptibility was strongest in females, whereas our chromosome 17 QTL was strongest in males. Due to the fact that a previous study had already associated our prime candidate underlying our chromosome 2 QTL (i.e., C5) with susceptibility to influenza, we chose to focus our validation on the chromosome 17 QTL, and use only male mice for the validation. Interestingly, the study that originally identified C5 as linked to influenza susceptibility was performed using only female mice [159].

Originally, when I arrived in the lab, the accepted QTL analysis for RCS studies was not as complex. ANOVA were run on for single markers, using the strain background as a fixed covariate (i.e., either AcB or BcA). This analysis often resulted in large numbers of long genomic segments being linked to a disease trait. For example, 16 regions of linkage were identified in a screen of airway responsiveness in response to methacholine as assessed by whole-body plethysmography [338]. It is entirely possible that these QTL were valid, but it is equally possible that some of these QTL were spurious results brought on by the complex nature of the dichotomous backgrounds (i.e. AcB and BcA) in the RCS.

Over the course of my PhD we experimented with many mixed model analyses in order to reduce the number of spurious QTL and speed the process from initial phenotype to causative gene. A big step forward in our statistical analysis was the move away from a simple covariate used to control our population background to a covariate matrix. The covariate matrix looks at the co-relatedness of every possible pair of RCS strains from available genotype

information (e.g., 2 BcA strains may have 85% similarity in terms of their genetic makeup as judged by the available marker map, whereas another pair may have 90% similarity). These numbers are used together in a covariate matrix in order to fine tune the co-relatedness of the mouse strains in our population and adjust more accurately to the population structure. Another large step forward came through the optimization of running a mixed model in a reduced amount of time, without the need for permutation testing [212].

Another improvement that we made over the course of my thesis was an increase in the density of our marker map for the RCS. We genotyped each RCS line using the Mouse Diversity Genotyping Array from Affymetrix. This panel had over 100,000 polymorphic markers between the C57BL/6 and A/J mouse strains. Combining this with our microsatellite marker panel and a medium density panel from Illumina, we were able to create an extremely accurate map of haplotype blocks for the RCS population. This improvement, along with the greatly improved statistical control of our population background, drastically increased the results that we were able to obtain from RCS studies. Simply reanalyzing the results from previous RCS studies with our newer method could identify many novel QTL. This, combined with our novel method of candidate gene identification, could result in the identification of many compelling candidate genes based on function for previously tested infectious diseases and other phenotypes.

Using eQTL to identify candidate genes underlying cQTL

Originally, identifying a causative gene underlying a QTL within an RCS study was an incredibly difficult task to accomplish over the course of a PhD. Along with the large candidate intervals and spurious QTL, there was the arduous task of identifying candidate genes for further study. There were many online resources available to aid in the attempt. For example, all genes with non-synonymous polymorphisms could be identified through a database on the Mouse Genome Informatics website. Then rudimentary *in silico* analyses could be conducted to identify potentially damaging SNPs within the linked interval (e.g.,

SIFT from the Venter Institute). Alternatively, expression analyses of various tissues and cell types through published datasets (e.g., BioGPS.org) could be used to identify whether the candidate genes were expressed in relevant tissues, or upregulated in response to an immune trigger (e.g. LPS).

Our method made the identification of candidate genes less subjective by identifying functional baseline differences in genes within tissues of interest (e.g., lung tissue for influenza and liver tissue for ECTV). We reanalyzed a set of microarrays that had been previously run on the RCS. From this analysis, we obtained a list of genes that were differentially cis-regulated within the RCS population. That is, we determined whether the RCS strains with an A/J genotype at a particular gene were significantly differentially expressed than those RCS strains with C57BL/6 genotypes at that gene. This information proved incredibly valuable in identifying valid candidate genes. Both of the genes that we went on to validate in this thesis, *H2-T23* for ectromelia and *Pla2g7* for influenza, were identified in this manner. Further, the genes were both identified with high levels of statistical significance (10^{-14} for *H2-T23* and 10^{-5} for *Pla2g7* – the 3rd strongest eQTL in the dataset). This strategy again gave an objective, accurate, and validated method to obtain candidate genes in RCS studies based on function, and again streamlined the process of going from a phenotype to the identification of a causative gene underlying its control.

In addition to providing the tools to reanalyze previous RCS results with greater marker density, control for population structure, and functional candidate identification, this thesis also provides the tools to identify compelling candidates from eQTL studies. The improved marker map and statistical analyses resulting from this thesis, combined with the expression information from the multiple microarrays for multiple tissues for each RCS strain that were previously studied, provides the basis to identify potential genetic regions controlling the expression levels of multiple genes both in *cis* and in *trans*. This information could be highly useful for future research.

H2-T23

H2-T23 encodes the highly conserved MHC class 1 molecule QA1^b in mice. QA1^b provides resistance to B6 mice against ECTV when interacting with CD94-NKG2E on NK cells that are simultaneously stimulated by NKG2D [87]. There is also evidence for QA1^b interacting with CD94-NKG2A molecules on CD-8 + T cells regulating cytotoxic T cell function [240].

Interestingly, although QA1^b deficient mice have previously been shown to succumb to mousepox, they had an overall mortality rate of ~20% when infected with the virus [87]. Because mice deficient in CD94 (the receptor for QA1^b) are substantially more susceptible than QA1^b deficient mice, it is possible that other mechanisms, such as the CD-8 T-cell regulation mentioned above, may contribute to the susceptibility of QA1^b deficient mice [339], [340].

In our F₂ and congenic work, we saw a susceptibility level of ~60-80%, which is much higher than that seen for QA1^b deficient mice. It is possible that the A/J polymorphism may disrupt NK cell killing, while not interrupting the CD8 regulatory T-cell function, resulting in decreased immune responding and increased mortality as compared to the *H2-T23* knockout. It is also entirely possible that more polymorphisms may be present in the *H2* region of the congenic mice that we created, and may be influencing the susceptibility phenotype as well. More research is needed to completely dissect the contribution of each of these potential options in our ECTV model. The congenic mouse lines that we created during this PhD would be useful in starting this dissection.

Pla2g7

The protein encoded by *Pla2g7* is a secreted phospholipase with high specificity for oxidatively modified sn-2 fatty acid residues in phospholipids that has been implicated in many disease processes. Specifically, two contrasting effects on inflammation have been described. PLA2G7 can hydrolyze oxidized phospholipids to produce lysophosphatidylcholine and free oxidized fatty acids,

leading to increased inflammation [281]. It also catalyzes the degradation of platelet-activating factor into inactive components, leading to reduced inflammation [284].

High levels of PLA2G7 activity were implicated in a swine model of coronary artery disease by increasing inflammation in the necrotic cores of atherosclerotic plaques [283]. In this animal model, pharmacological inhibition of PLA2G7 reduced inflammation and disease severity by reducing levels of proinflammatory lysophosphatidylcholine [287].

Recently, PLA2G7 has been shown to alter the inflammatory response after myocardial infarction. Further, it has been shown to disrupt healing by altering the balance between repair and inflammation [321]. Our microarray analyses also indicated a trend towards mechanisms of tissue repair in our resistant *Pla2g7*-deficient mice (e.g., through the early upregulation of *Cyr61* and *Trim72*, and through the later up-regulation of integrin related pathways).

Inflammation is a complex process, often with multiple mechanisms present to balance pro-inflammatory and anti-inflammatory responses. The complexity of these networks has made it challenging to identify those specific pathways or key enzymes that contribute directly to disease states.

The development of Darapladib, a novel therapeutic agent that directly inhibits the Lp-PLA₂ enzyme, allows the possibility of directly testing whether the inhibition of *pla2g7* may translate into improved clinical outcomes in humans [341].

Darapladib, an inhibitor of PLA2G7, was initially developed as a treatment for atherosclerosis. In several phase 3 clinical trials, Darapladib failed to reduce the risk of coronary heart disease, death, myocardial infarction, and urgent coronary revascularization compared with placebo in patients with acute coronary syndrome treated with standard medical care [342]. The drug, however, has been shown to reduce inflammatory cytokine production [343], [287].

Unfortunately, the efficacy of the drug is extremely poor in mice, making it ineffective for use in our studies, but many other human-related studies could be completed to further understand its role in inflammation and its potential link to resistance to infectious disease.

Additionally, a common loss-of-function mutation in *PLA2G7* (the V279F allele) has been identified in individuals of Japanese, Chinese, and Korean ancestry [344]. This allele leads to a natural deficiency (heterozygous ~50%) or absence (homozygous ~100%) of LP-PLa2 activity. We have recently reached out to collaborators with human cohorts for other pathogenic diseases (e.g., Respiratory Syncytial Virus and Dengue Virus) to test cohorts of individuals with Asian ancestry for an association between *PLA2G7* alleles and resistance to infection. Our results with Dengue Virus have so far been promising. Over the course of my PhD, studies have been published using human cohorts with increased susceptibility to influenza virus [6]. It would be my hope that with time, more of these cohorts will be available for study of candidate genes identified over the course of my work.

CONCLUSION

Despite the useful results from the work presented in this thesis, and all of the useful research that has come from the AcB/BcA panel of RCS mice, testing further phenotypes with this model has been put on hold. The mice have been removed from our animal facilities, and have been cryogenically preserved. The month to month cost of maintaining such an expansive mouse colony is high. This, along with the continual accumulation of *de novo* mutations in each of the strains has made these mice increasingly tricky for continued study. Other mouse genetics projects with increased potential for identifying underlying causative polymorphisms (e.g., ENU mutagenesis) have replaced the RCS in our forward genetics programs.

Even these newer mouse studies, however, suffer from their own inherent difficulties (e.g. with ENU, compounding heterozygous mutations and the equally prohibitive costs of maintaining animal colonies). The cost of human expression analyses and sequencing efforts have continued to decrease over the course of this thesis, to a point where one must question the overall ethical nature of using the mouse as a genetic model to study causative mutations or identify eQTL.

Despite the vast array of genetic and phenotypic information we have accrued over the decades of mouse research, there are many results that do not translate directly to humans. It may be increasingly useful to restrict future forward genetic studies, wherever possible, to the species in which we wish to make the biggest impact, humans. This will, of course, bring with it its own set of methodological and ethical issues, but could stop the excess breeding and culling of many future living, breathing, feeling individuals: laboratory mice.

I sacrificed hundreds if not thousands of mice to obtain the results in this thesis. It is my hope that the work presented here can be used to increase our understanding of basic immunology, and potentially provide a new target for therapies against pathogenic infectious disease. I am confident that when I started this thesis, my work with animals was warranted, and that there was not

another viable option to effectively study my chosen infectious diseases in living populations. If I were to start this thesis now, with the decreases in cost and availability of human sequencing, and with the presence of human cohorts in infectious disease slowly becoming available, I am not completely sure that this statement would still hold true.

Of course, the mouse is still useful in modeling certain highly dangerous diseases with no current human cohorts (or where the creation of a human cohort would be impossible), and in unraveling the mechanisms of disease associated with causative mutations or eQTL. It is not my intention to soften the future utility of mouse genetics. However, I urge anyone reading this who plans to use animal models in the future to evaluate their research with their peers and mentors, and decide whether the questions they are asking could be answered in a more ethically sound manner with other genetic tools.

Overall, this thesis streamlined the methodology of studies using recombinant congenic mice by increasing the density of the marker map, increasing the control of population structure during statistical analysis, creating an objective and valid means of identifying candidate genes in QTL by function, and validating two genes underlying susceptibility to two infectious diseases. The discovery of either or both of these genes could lead to valuable future work, either immunological in nature (*H2-T23*), or with the potential to directly improve treatment of pathogenic human infection (*Pla2g7*).

GENERAL REFERENCES

1. Kaufmann, S.H., *Robert Koch, the Nobel Prize, and the ongoing threat of tuberculosis*. N Engl J Med, 2005. **353**(23): p. 2423-6.
2. Comstock, G.W., *Tuberculosis in twins: a re-analysis of the Proffit survey*. Am Rev Respir Dis, 1978. **117**(4): p. 621-4.
3. Herndon, C.N. and R.G. Jennings, *A twin-family study of susceptibility to poliomyelitis*. Am J Hum Genet, 1951. **3**(1): p. 17-46.
4. Lin, T.M., et al., *Hepatitis B virus markers in Chinese twins*. Anticancer Res, 1989. **9**(3): p. 737-41.
5. Hill, A.V., *Aspects of genetic susceptibility to human infectious diseases*. Annu Rev Genet, 2006. **40**: p. 469-86.
6. Everitt, A.R., et al., *IFITM3 restricts the morbidity and mortality associated with influenza*. Nature, 2012. **484**(7395): p. 519-23.
7. Watanabe, T., S. Watanabe, and Y. Kawaoka, *Cellular networks involved in the influenza virus life cycle*. Cell Host Microbe, 2010. **7**(6): p. 427-39.
8. Karlas, A., et al., *Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication*. Nature, 2010. **463**(7282): p. 818-22.
9. Karp, C.L., et al., *Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma*. Nat Immunol, 2000. **1**(3): p. 221-6.
10. Min-Oo, G., et al., *Pyruvate kinase deficiency in mice protects against malaria*. Nat Genet, 2003. **35**(4): p. 357-62.
11. Desrosiers, M.P., et al., *Epistasis between mouse Klra and major histocompatibility complex class I loci is associated with a new mechanism of natural killer cell-mediated innate resistance to cytomegalovirus infection*. Nat Genet, 2005. **37**(6): p. 593-9.
12. Mehrabian, M., et al., *Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits*. Nat Genet, 2005. **37**(11): p. 1224-33.
13. Vidal, S.M., et al., *Forward genetic dissection of immunity to infection in the mouse*. Annu Rev Immunol, 2008. **26**: p. 81-132.
14. Eppig, J.T., et al., *The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease*. Nucleic Acids Res, 2015. **43**(Database issue): p. D726-36.
15. Drinkwater, N.R. and M.N. Gould, *The long path from QTL to gene*. PLoS Genet, 2012. **8**(9): p. e1002975.
16. Flint, J., et al., *Strategies for mapping and cloning quantitative trait genes in rodents*. Nat Rev Genet, 2005. **6**(4): p. 271-86.
17. Riedel, S., *Edward Jenner and the history of smallpox and vaccination*. Proc (Bayl Univ Med Cent), 2005. **18**(1): p. 21-5.
18. Littman, R.J. and M.L. Littman, *Galen and the Antonine plague*. Am J Philol, 1973. **94**: p. 243-55.
19. Barquet, N. and P. Domingo, *Smallpox: the triumph over the most terrible of the ministers of death*. Ann Intern Med, 1997. **127**(8 Pt 1): p. 635-42.
20. Fenner F, et al., *Smallpox and its Eradication*. 1988, WHO: WHO.

21. Lyons AS, P.R., *Medicine - An Illustrated History*. 1987, New York: Abradale Press, Harry N Abrams Inc.
22. Solomon, M., *Mozart: A Life*. 1995, New York: Harper Collins.
23. Bumgarner, J., *The Health of the Predisents: The 41 United States Presidents Through 1993 from a Physician's Point of View*. 1993, NC: McFarland and Company.
24. Goldman, A.S. and F.C. Schmalstieg, Jr., *Abraham Lincoln's Gettysburg illness*. J Med Biogr, 2007. **15**(2): p. 104-10.
25. Montefiore, *Young Stalin*. 2007, New York: Random House.
26. Reed, K.D., et al., *The detection of monkeypox in humans in the Western Hemisphere*. N Engl J Med, 2004. **350**(4): p. 342-50.
27. Rao, A.R., *Smallpox*. 1972, Bombay: Kothari Book Depot.
28. Fenner F, H.D., Arita I, Jezek D, Ladnyi D, *Smallpox and its Eradication*. 1988, WHO: WHO.
29. Downie, A.W., et al., *Haemorrhagic smallpox*. J Hyg (Lond), 1969. **67**(4): p. 619-29.
30. Downie, A.W., et al., *Virus and virus antigen in the blood of smallpox patients; their significance in early diagnosis and prognosis*. Lancet, 1953. **265**(6778): p. 164-6.
31. Bras, G., *The morbid anatomy of smallpox*. Doc Med Geogr Trop, 1952. **4**(4): p. 303-51.
32. Jahrling, P.B., et al., *Exploring the potential of variola virus infection of cynomolgus macaques as a model for human smallpox*. Proc Natl Acad Sci U S A, 2004. **101**(42): p. 15196-200.
33. Rosengard, A.M., et al., *Variola virus immune evasion design: expression of a highly efficient inhibitor of human complement*. Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8808-13.
34. Smith, C.A., et al., *Poxvirus genomes encode a secreted, soluble protein that preferentially inhibits beta chemokine activity yet lacks sequence homology to known chemokine receptors*. Virology, 1997. **236**(2): p. 316-27.
35. Upton, C., et al., *Poxvirus orthologous clusters: toward defining the minimum essential poxvirus genome*. J Virol, 2003. **77**(13): p. 7590-600.
36. Lefkowitz, E.J., C. Wang, and C. Upton, *Poxviruses: past, present and future*. Virus Res, 2006. **117**(1): p. 105-18.
37. Dunlop, L.R., et al., *Variola virus immune evasion proteins*. Microbes Infect, 2003. **5**(11): p. 1049-56.
38. Laliberte, J.P., A.S. Weisberg, and B. Moss, *The membrane fusion step of vaccinia virus entry is cooperatively mediated by multiple viral proteins and host cell components*. PLoS Pathog, 2011. **7**(12): p. e1002446.
39. Senkevich, T.G., et al., *Poxvirus multiprotein entry-fusion complex*. Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18572-7.
40. Mercer, J. and A. Helenius, *Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells*. Science, 2008. **320**(5875): p. 531-5.
41. Schramm, B. and J.K. Locker, *Cytoplasmic organization of POXvirus DNA replication*. Traffic, 2005. **6**(10): p. 839-46.

42. Moss, B., *Poxvirus DNA replication*. Cold Spring Harb Perspect Biol, 2013. **5**(9).
43. Smith, G.L., A. Vanderplasschen, and M. Law, *The formation and function of extracellular enveloped vaccinia virus*. J Gen Virol, 2002. **83**(Pt 12): p. 2915-31.
44. Khodakevich, L., et al., *The role of squirrels in sustaining monkeypox virus transmission*. Trop Geogr Med, 1987. **39**(2): p. 115-22.
45. Arita, I., et al., *Outbreaks of monkeypox and serological surveys in nonhuman primates*. Bull World Health Organ, 1972. **46**(5): p. 625-31.
46. Stanford, M.M., et al., *Immunopathogenesis of poxvirus infections: forecasting the impending storm*. Immunol Cell Biol, 2007. **85**(2): p. 93-102.
47. Zaucha, G.M., et al., *The pathology of experimental aerosolized monkeypox virus infection in cynomolgus monkeys (Macaca fascicularis)*. Lab Invest, 2001. **81**(12): p. 1581-600.
48. McConnell, S., et al., *Protection of Rhesus Monkeys against Monkeypox by Vaccinia Virus Immunization*. Am J Vet Res, 1964. **25**: p. 192-5.
49. Saijo, M., et al., *LC16m8, a highly attenuated vaccinia virus vaccine lacking expression of the membrane protein B5R, protects monkeys from monkeypox*. J Virol, 2006. **80**(11): p. 5179-88.
50. Earl, P.L., et al., *Immunogenicity of a highly attenuated MVA smallpox vaccine and protection against monkeypox*. Nature, 2004. **428**(6979): p. 182-5.
51. Fenner, F., *Adventures with poxviruses of vertebrates*. FEMS Microbiol Rev, 2000. **24**(2): p. 123-33.
52. Smith, S.A. and G.J. Kotwal, *Immune response to poxvirus infections in various animals*. Crit Rev Microbiol, 2002. **28**(3): p. 149-85.
53. Esteban, D.J. and R.M. Buller, *Ectromelia virus: the causative agent of mousepox*. J Gen Virol, 2005. **86**(Pt 10): p. 2645-59.
54. Smith, G.L., *Vaccinia virus immune evasion*. Immunol Lett, 1999. **65**(1-2): p. 55-62.
55. Schriewer, J., R.M. Buller, and G. Owens, *Mouse models for studying orthopoxvirus respiratory infections*. Methods Mol Biol, 2004. **269**: p. 289-308.
56. Smee, D.F. and R.W. Sidwell, *A review of compounds exhibiting anti-orthopoxvirus activity in animal models*. Antiviral Res, 2003. **57**(1-2): p. 41-52.
57. De Clercq, E., *Cidofovir in the treatment of poxvirus infections*. Antiviral Res, 2002. **55**(1): p. 1-13.
58. Chaudhri, G., et al., *Polarized type 1 cytokine response and cell-mediated immunity determine genetic resistance to mousepox*. Proc Natl Acad Sci U S A, 2004. **101**(24): p. 9057-62.
59. Karupiah, G., et al., *Importance of interferons in recovery from mousepox*. J Virol, 1993. **67**(7): p. 4214-26.
60. Karupiah, G., et al., *Immunobiology of infection with recombinant vaccinia virus encoding murine IL-2. Mechanisms of rapid viral clearance in immunocompetent mice*. J Immunol, 1991. **147**(12): p. 4327-32.

61. Karupiah, G., et al., *Different roles for CD4+ and CD8+ T lymphocytes and macrophage subsets in the control of a generalized virus infection*. J Virol, 1996. **70**(12): p. 8301-9.
62. Jacoby, R.O., P.N. Bhatt, and D.G. Brownstein, *Evidence that NK cells and interferon are required for genetic resistance to lethal infection with ectromelia virus*. Arch Virol, 1989. **108**(1-2): p. 49-58.
63. Delano, M.L. and D.G. Brownstein, *Innate resistance to lethal mousepox is genetically linked to the NK gene complex on chromosome 6 and correlates with early restriction of virus replication by cells with an NK phenotype*. J Virol, 1995. **69**(9): p. 5875-7.
64. Mullbacher, A., et al., *Perforin is essential for control of ectromelia virus but not related poxviruses in mice*. J Virol, 1999. **73**(2): p. 1665-7.
65. Melo-Silva, C.R., et al., *The ectromelia virus SPI-2 protein causes lethal mousepox by preventing NK cell responses*. J Virol, 2011. **85**(21): p. 11170-82.
66. Smith, V.P. and A. Alcamí, *Inhibition of interferons by ectromelia virus*. J Virol, 2002. **76**(3): p. 1124-34.
67. Mossman, K., et al., *Species specificity of ectromelia virus and vaccinia virus interferon-gamma binding proteins*. Virology, 1995. **208**(2): p. 762-9.
68. Alcamí, A., et al., *Vaccinia virus strains Lister, USSR and Evans express soluble and cell-surface tumour necrosis factor receptors*. J Gen Virol, 1999. **80** (Pt 4): p. 949-59.
69. Saraiva, M. and A. Alcamí, *Crme, a novel soluble tumor necrosis factor receptor encoded by poxviruses*. J Virol, 2001. **75**(1): p. 226-33.
70. Esteban, D.J., A.A. Nuara, and R.M. Buller, *Interleukin-18 and glycosaminoglycan binding by a protein encoded by Variola virus*. J Gen Virol, 2004. **85**(Pt 5): p. 1291-9.
71. Smith, V.P. and A. Alcamí, *Expression of secreted cytokine and chemokine inhibitors by ectromelia virus*. J Virol, 2000. **74**(18): p. 8460-71.
72. Kluczyk, A., et al., *The immunosuppressive activity of peptide fragments of vaccinia virus C10L protein and a hypothesis on the role of this protein in the viral invasion*. Peptides, 2002. **23**(5): p. 823-34.
73. DiPerna, G., et al., *Poxvirus protein N1L targets the I-kappaB kinase complex, inhibits signaling to NF-kappaB by the tumor necrosis factor superfamily of receptors, and inhibits NF-kappaB and IRF3 signaling by toll-like receptors*. J Biol Chem, 2004. **279**(35): p. 36570-8.
74. Kettle, S., et al., *Vaccinia virus serpin B13R (SPI-2) inhibits interleukin-1beta-converting enzyme and protects virus-infected cells from TNF- and Fas-mediated apoptosis, but does not prevent IL-1beta-induced fever*. J Gen Virol, 1997. **78** (Pt 3): p. 677-85.
75. Komiyama, T., et al., *Inhibition of interleukin-1 beta converting enzyme by the cowpox virus serpin CrmA. An example of cross-class inhibition*. J Biol Chem, 1994. **269**(30): p. 19331-7.
76. Ray, C.A., et al., *Viral inhibition of inflammation: cowpox virus encodes an inhibitor of the interleukin-1 beta converting enzyme*. Cell, 1992. **69**(4): p. 597-604.

77. Dobbelstein, M. and T. Shenk, *Protection against apoptosis by the vaccinia virus SPI-2 (B13R) gene product*. J Virol, 1996. **70**(9): p. 6479-85.
78. Quan, L.T., et al., *Granzyme B is inhibited by the cowpox virus serpin cytokine response modifier A*. J Biol Chem, 1995. **270**(18): p. 10377-9.
79. Stewart, T.L., S.T. Wasilenko, and M. Barry, *Vaccinia virus F1L protein is a tail-anchored protein that functions at the mitochondria to inhibit apoptosis*. J Virol, 2005. **79**(2): p. 1084-98.
80. Langland, J.O. and B.L. Jacobs, *The role of the PKR-inhibitory genes, E3L and K3L, in determining vaccinia virus host range*. Virology, 2002. **299**(1): p. 133-41.
81. Cameron, C.M., et al., *Myxoma virus M128L is expressed as a cell surface CD47-like virulence factor that contributes to the downregulation of macrophage activation in vivo*. Virology, 2005. **337**(1): p. 55-67.
82. Ng, A., et al., *The vaccinia virus A41L protein is a soluble 30 kDa glycoprotein that affects virus virulence*. J Gen Virol, 2001. **82**(Pt 9): p. 2095-105.
83. Reading, P.C., J.B. Moore, and G.L. Smith, *Steroid hormone synthesis by vaccinia virus suppresses the inflammatory response to infection*. J Exp Med, 2003. **197**(10): p. 1269-78.
84. Mullbacher, A., *Cell-mediated cytotoxicity in recovery from poxvirus infections*. Rev Med Virol, 2003. **13**(4): p. 223-32.
85. Brownstein, D.G. and L. Gras, *Differential pathogenesis of lethal mousepox in congenic DBA/2 mice implicates natural killer cell receptor NKR-P1 in necrotizing hepatitis and the fifth component of complement in recruitment of circulating leukocytes to spleen*. Am J Pathol, 1997. **150**(4): p. 1407-20.
86. Brownstein, D.G., et al., *Serial backcross analysis of genetic resistance to mousepox, using marker loci for Rmp-2 and Rmp-3*. J Virol, 1992. **66**(12): p. 7073-9.
87. Fang, M., et al., *CD94 is essential for NK cell-mediated resistance to a lethal viral disease*. Immunity, 2011. **34**(4): p. 579-89.
88. Wallensten, A., *Influenza virus in wild birds and mammals other than man*. Microbial Ecology in Health and Disease, 2007. **19**(2): p. 122 - 139.
89. Nicholson, K.G.W., R.G.; Hay, A.J. , *Textbook of influenza*. 1998, London: Blackwell Science.
90. Thompson, W.W., et al., *Mortality associated with influenza and respiratory syncytial virus in the United States*. JAMA, 2003. **289**(2): p. 179-86.
91. Palese, P., C.F. Basler, and A. Garcia-Sastre, *The makings of a killer*. Nat Med, 2002. **8**(9): p. 927-8.
92. Henzler, T., et al., *Image findings of patients with H1N1 virus pneumonia and acute respiratory failure*. Acad Radiol, 2010. **17**(6): p. 681-5.
93. Ramsey, C. and A. Kumar, *H1N1: viral pneumonia as a cause of acute respiratory distress syndrome*. Curr Opin Crit Care, 2011. **17**(1): p. 64-71.

94. Baskin, C.R., et al., *Early and sustained innate immune response defines pathology and death in nonhuman primates infected by highly pathogenic influenza virus*. Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3455-60.
95. Kash, J.C., et al., *Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus*. Nature, 2006. **443**(7111): p. 578-81.
96. Kobasa, D., et al., *Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus*. Nature, 2007. **445**(7125): p. 319-23.
97. Webster, R.G., et al., *Evolution and ecology of influenza A viruses*. Microbiol Rev, 1992. **56**(1): p. 152-79.
98. Ito, T. and Y. Kawaoka, *Host-range barrier of influenza A viruses*. Vet Microbiol, 2000. **74**(1-2): p. 71-5.
99. CDC. *Types of Influenza Viruses*. 2014; Available from: <http://www.cdc.gov/flu/about/viruses/types.htm>.
100. Carrat, F., et al., *Time lines of infection and disease in human influenza: a review of volunteer challenge studies*. Am J Epidemiol, 2008. **167**(7): p. 775-85.
101. Beigel, J.H., et al., *Avian influenza A (H5N1) infection in humans*. N Engl J Med, 2005. **353**(13): p. 1374-85.
102. Lin, K.L., et al., *CCR2+ monocyte-derived dendritic cells and exudate macrophages produce influenza-induced pulmonary immune pathology and mortality*. J Immunol, 2008. **180**(4): p. 2562-72.
103. McGill, J., J.W. Heusel, and K.L. Legge, *Innate immune control and regulation of influenza virus infections*. J Leukoc Biol, 2009. **86**(4): p. 803-12.
104. Herold, S., et al., *Alveolar epithelial cells direct monocyte transepithelial migration upon influenza virus infection: impact of chemokines and adhesion molecules*. J Immunol, 2006. **177**(3): p. 1817-24.
105. Ison, M.G., *Antivirals and resistance: influenza virus*. Curr Opin Virol, 2011. **1**(6): p. 563-73.
106. Marjuki, H., et al., *An investigational antiviral drug, DAS181, effectively inhibits replication of zoonotic influenza A virus subtype H7N9 and protects mice from lethality*. J Infect Dis, 2014. **210**(3): p. 435-40.
107. Belser, J.A., et al., *DAS181, a novel sialidase fusion protein, protects mice from lethal avian influenza H5N1 virus infection*. J Infect Dis, 2007. **196**(10): p. 1493-9.
108. Furuta, Y., et al., *T-705 (favipiravir) and related compounds: Novel broad-spectrum inhibitors of RNA viral infections*. Antiviral Res, 2009. **82**(3): p. 95-102.
109. Fiore, A.E., et al., *Antiviral agents for the treatment and chemoprophylaxis of influenza --- recommendations of the Advisory Committee on Immunization Practices (ACIP)*. MMWR Recomm Rep, 2011. **60**(1): p. 1-24.

110. Hurt, A.C., et al., *Oseltamivir resistance and the H274Y neuraminidase mutation in seasonal, pandemic and highly pathogenic influenza viruses*. *Drugs*, 2009. **69**(18): p. 2523-31.
111. Lackenby, A., et al., *Emergence of resistance to oseltamivir among influenza A(H1N1) viruses in Europe*. *Euro Surveill*, 2008. **13**(5).
112. Triana-Baltzer, G.B., et al., *Phenotypic and genotypic characterization of influenza virus mutants selected with the sialidase fusion protein DAS181*. *J Antimicrob Chemother*, 2011. **66**(1): p. 15-28.
113. Palache, A.M., *Influenza vaccines. A reappraisal of their use*. *Drugs*, 1997. **54**(6): p. 841-56.
114. Ershler, W.B., *Influenza vaccination in the elderly: can efficacy be enhanced?* *Geriatrics*, 1988. **43**(9): p. 79-83.
115. CDC. 2007 - 2008 Influenza (Flu) Season. 2008; Available from: <http://www.cdc.gov/flu/pastseasons/0708season.htm>.
116. Ahmed, F., J.A. Singleton, and A.L. Franks, *Clinical practice. Influenza vaccination for healthy young adults*. *N Engl J Med*, 2001. **345**(21): p. 1543-7.
117. Noda, T., *Native morphology of influenza virions*. *Front Microbiol*, 2011. **2**: p. 269.
118. Neumann, G., et al., *Orthomyxovirus replication, transcription, and polyadenylation*. *Curr Top Microbiol Immunol*, 2004. **283**: p. 121-43.
119. Garcia-Sastre, A., et al., *Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems*. *Virology*, 1998. **252**(2): p. 324-30.
120. Haller, O., P. Staeheli, and G. Kochs, *Protective role of interferon-induced Mx GTPases against influenza viruses*. *Rev Sci Tech*, 2009. **28**(1): p. 219-31.
121. Cox, N.J. and K. Subbarao, *Global epidemiology of influenza: past and present*. *Annu Rev Med*, 2000. **51**: p. 407-21.
122. Ibricevic, A., et al., *Influenza virus receptor specificity and cell tropism in mouse and human airway epithelial cells*. *J Virol*, 2006. **80**(15): p. 7469-80.
123. Harrison, S.C., *Viral membrane fusion*. *Nat Struct Mol Biol*, 2008. **15**(7): p. 690-8.
124. Martin-Benito, J., et al., *Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle*. *EMBO Rep*, 2001. **2**(4): p. 313-7.
125. Cros, J.F. and P. Palese, *Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses*. *Virus Res*, 2003. **95**(1-2): p. 3-12.
126. Boulo, S., et al., *Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes*. *Virus Res*, 2007. **124**(1-2): p. 12-21.
127. Cilloniz, C., et al., *Lethal influenza virus infection in macaques is associated with early dysregulation of inflammatory related genes*. *PLoS Pathog*, 2009. **5**(10): p. e1000604.
128. Maher, J.A. and J. DeStefano, *The ferret: an animal model to study influenza virus*. *Lab Anim (NY)*, 2004. **33**(9): p. 50-3.

129. van Riel, D., et al., *Human and avian influenza viruses target different cells in the lower respiratory tract of humans and other mammals*. Am J Pathol, 2007. **171**(4): p. 1215-23.
130. Cameron, C.M., et al., *Gene expression analysis of host innate immune responses during Lethal H5N1 infection in ferrets*. J Virol, 2008. **82**(22): p. 11308-17.
131. Svitek, N., et al., *Severe seasonal influenza in ferrets correlates with reduced interferon and increased IL-6 induction*. Virology, 2008. **376**(1): p. 53-9.
132. Houser, K.V., et al., *Impact of prior seasonal H3N2 influenza vaccination or infection on protection and transmission of emerging variants of influenza A(H3N2)v virus in ferrets*. J Virol, 2013. **87**(24): p. 13480-9.
133. Krammer, F., et al., *Assessment of influenza virus hemagglutinin stalk-based immunity in ferrets*. J Virol, 2014. **88**(6): p. 3432-42.
134. Lund, J.M., et al., *Recognition of single-stranded RNA viruses by Toll-like receptor 7*. Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5598-603.
135. Hornung, V., et al., *5'-Triphosphate RNA is the ligand for RIG-I*. Science, 2006. **314**(5801): p. 994-7.
136. Pichlmair, A., et al., *RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates*. Science, 2006. **314**(5801): p. 997-1001.
137. Gack, M.U., et al., *TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity*. Nature, 2007. **446**(7138): p. 916-920.
138. Der, S.D., et al., *Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays*. Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15623-8.
139. Haller, O., M. Frese, and G. Kochs, *Mx proteins: mediators of innate resistance to RNA viruses*. Rev Sci Tech, 1998. **17**(1): p. 220-30.
140. Fritz, J.H., et al., *Nod-like proteins in immunity, inflammation and disease*. Nat Immunol, 2006. **7**(12): p. 1250-7.
141. Allen, I.C., et al., *The NLRP3 inflammasome mediates in vivo innate immunity to influenza A virus through recognition of viral RNA*. Immunity, 2009. **30**(4): p. 556-65.
142. Allen, I.C., et al., *NLRX1 protein attenuates inflammatory responses to infection by interfering with the RIG-I-MAVS and TRAF6-NF-kappaB signaling pathways*. Immunity, 2011. **34**(6): p. 854-65.
143. Ichinohe, T., et al., *Inflammasome recognition of influenza virus is essential for adaptive immune responses*. J Exp Med, 2009. **206**(1): p. 79-87.
144. Nayak, B., et al., *Contributions of the avian influenza virus HA, NA, and M2 surface proteins to the induction of neutralizing antibodies and protective immunity*. J Virol, 2010. **84**(5): p. 2408-20.
145. Liu, J., et al., *Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein*. Nat Struct Biol, 1997. **4**(11): p. 896-9.
146. Nemeroff, M.E., et al., *Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs*. Mol Cell, 1998. **1**(7): p. 991-1000.

147. Fortes, P., A. Beloso, and J. Ortin, *Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport*. EMBO J, 1994. **13**(3): p. 704-12.
148. Satterly, N., et al., *Influenza virus targets the mRNA export machinery and the nuclear pore complex*. Proc Natl Acad Sci U S A, 2007. **104**(6): p. 1853-8.
149. Gack, M.U., et al., *Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I*. Cell Host Microbe, 2009. **5**(5): p. 439-49.
150. Hale, B.G., et al., *The multifunctional NS1 protein of influenza A viruses*. J Gen Virol, 2008. **89**(Pt 10): p. 2359-76.
151. Fernandez-Sesma, A., et al., *Influenza virus evades innate and adaptive immunity via the NS1 protein*. J Virol, 2006. **80**(13): p. 6295-304.
152. Haye, K., et al., *The NS1 protein of a human influenza virus inhibits type I interferon production and the induction of antiviral responses in primary human dendritic and respiratory epithelial cells*. J Virol, 2009. **83**(13): p. 6849-62.
153. Moltedo, B., et al., *Cutting edge: stealth influenza virus replication precedes the initiation of adaptive immunity*. J Immunol, 2009. **183**(6): p. 3569-73.
154. Schmolke, M. and A. Garcia-Sastre, *Evasion of innate and adaptive immune responses by influenza A virus*. Cell Microbiol, 2010. **12**(7): p. 873-80.
155. Arnheiter, H. and O. Haller, *Antiviral state against influenza virus neutralized by microinjection of antibodies to interferon-induced Mx proteins*. EMBO J, 1988. **7**(5): p. 1315-20.
156. Haller, O., et al., *Mx GTPases: dynamin-like antiviral machines of innate immunity*. Trends Microbiol, 2015. **23**(3): p. 154-63.
157. Nedelko, T., et al., *Distinct gene loci control the host response to influenza H1N1 virus infection in a time-dependent manner*. BMC Genomics, 2012. **13**: p. 411.
158. Alberts, R., et al., *Genome-wide analysis of the mouse lung transcriptome reveals novel molecular gene interaction networks and cell-specific expression signatures*. Respir Res, 2011. **12**: p. 61.
159. Boon, A.C., et al., *Host genetic variation affects resistance to infection with a highly pathogenic H5N1 influenza A virus in mice*. J Virol, 2009. **83**(20): p. 10417-26.
160. Boon, A.C., et al., *A novel genetic locus linked to pro-inflammatory cytokines after virulent H5N1 virus infection in mice*. BMC Genomics, 2014. **15**: p. 1017.
161. Ferris, M.T., et al., *Modeling host genetic regulation of influenza pathogenesis in the collaborative cross*. PLoS Pathog, 2013. **9**(2): p. e1003196.
162. Valdar, W., et al., *Mapping in structured populations by resample model averaging*. Genetics, 2009. **182**(4): p. 1263-77.

163. Auffray, J.V., F; and Britton-Davidian J, *The house mouse progression in Eurasia: a palaeontological and archaeozoological approach*. Biol. J. Linnean Soc. , 1990(41): p. 13-25.
164. Horiuchi, Y., et al., *Polymorphisms distinguishing different mouse species and t haplotypes*. Genet Res, 1992. **60**(1): p. 43-52.
165. Boursot, P., Auffray, J. -C., Britton-Davidian, J. and Bonhomme, F. , *The evolution of house mice*. Ann Rev Ecol Syst, 1993(24): p. 119-152.
166. Silver, L.M., *Mouse Genetics, Concepts and Applications*. 1995: Oxford University Press.
167. Yonekawa, H., et al., *Evolutionary relationships among five subspecies of Mus musculus based on restriction enzyme cleavage patterns of mitochondrial DNA*. Genetics, 1981. **98**(4): p. 801-16.
168. Yang, H., et al., *Subspecific origin and haplotype diversity in the laboratory mouse*. Nat Genet, 2011. **43**(7): p. 648-55.
169. Flint, J. and E. Eskin, *Genome-wide association studies in mice*. Nat Rev Genet, 2012. **13**(11): p. 807-17.
170. Trammell, R.A. and L.A. Toth, *Genetic susceptibility and resistance to influenza infection and disease in humans and mice*. Expert Rev Mol Diagn, 2008. **8**(4): p. 515-29.
171. Ding, M., L. Lu, and L.A. Toth, *Gene expression in lung and basal forebrain during influenza infection in mice*. Genes Brain Behav, 2008. **7**(2): p. 173-83.
172. Srivastava, B., et al., *Host genetic background strongly influences the response to influenza a virus infections*. PLoS One, 2009. **4**(3): p. e4857.
173. Schell, K., *Studies on the innate resistance of mice to infection with mousepox. II. Route of inoculation and resistance; and some observations on the inheritance of resistance*. Aust J Exp Biol Med Sci, 1960. **38**: p. 289-99.
174. Wallace, G.D. and R.M. Buller, *Kinetics of ectromelia virus (mousepox) transmission and clinical response in C57BL/6j, BALB/cByj and AKR/J inbred mice*. Lab Anim Sci, 1985. **35**(1): p. 41-6.
175. Panchanathan, V., G. Chaudhri, and G. Karupiah, *Correlates of protective immunity in poxvirus infection: where does antibody stand?* Immunol Cell Biol, 2008. **86**(1): p. 80-6.
176. Niemialtowski, M.G., et al., *The inflammatory and immune response to mousepox (infectious ectromelia) virus*. Acta Virol, 1994. **38**(5): p. 299-307.
177. Bhatnagar, S., et al., *Positional cloning of a type 2 diabetes quantitative trait locus; tomosyn-2, a negative regulator of insulin secretion*. PLoS Genet, 2011. **7**(10): p. e1002323.
178. Joe, B., et al., *Positional identification of variants of Adamts16 linked to inherited hypertension*. Hum Mol Genet, 2009. **18**(15): p. 2825-38.
179. Shirley, R.L., et al., *Mpdz is a quantitative trait gene for drug withdrawal seizures*. Nat Neurosci, 2004. **7**(7): p. 699-700.
180. Sebastiani, G., et al., *Cloning and characterization of the murine toll-like receptor 5 (Tlr5) gene: sequence and mRNA expression studies in Salmonella-susceptible MOLF/Ei mice*. Genomics, 2000. **64**(3): p. 230-40.

181. Nadeau, J.H., et al., *Analysing complex genetic traits with chromosome substitution strains*. Nat Genet, 2000. **24**(3): p. 221-5.
182. Singer, J.B., et al., *Genetic dissection of complex traits with chromosome substitution strains of mice*. Science, 2004. **304**(5669): p. 445-8.
183. Iakoubova, O.A., et al., *Genome-tagged mice (GTM): two sets of genome-wide congenic strains*. Genomics, 2001. **74**(1): p. 89-104.
184. Darvasi, A. and M. Soller, *Advanced intercross lines, an experimental population for fine genetic mapping*. Genetics, 1995. **141**(3): p. 1199-207.
185. Darvasi, A. and M. Soller, *A simple method to calculate resolving power and confidence interval of QTL map location*. Behav Genet, 1997. **27**(2): p. 125-32.
186. Valdar, W., et al., *Genome-wide genetic association of complex traits in heterogeneous stock mice*. Nat Genet, 2006. **38**(8): p. 879-87.
187. Solberg Woods, L.C., *QTL mapping in outbred populations: successes and challenges*. Physiol Genomics, 2014. **46**(3): p. 81-90.
188. Demant, P. and A.A. Hart, *Recombinant congenic strains--a new tool for analyzing genetic traits determined by more than one gene*. Immunogenetics, 1986. **24**(6): p. 416-22.
189. Fortin, A., et al., *Recombinant congenic strains derived from A/J and C57BL/6J: a tool for genetic dissection of complex traits*. Genomics, 2001. **74**(1): p. 21-35.
190. Nadeau, J.H., L.D. Arbuckle, and E. Skamene, *Genetic dissection of inflammatory responses*. J Inflamm, 1995. **45**(1): p. 27-48.
191. Fortin, A., et al., *The AcB/BcA recombinant congenic strains of mice: strategies for phenotype dissection, mapping and cloning of quantitative trait genes*. Novartis Found Symp, 2007. **281**: p. 141-53; discussion 153-5, 208-9.
192. Min-Oo, G., et al., *Phenotypic expression of pyruvate kinase deficiency and protection against malaria in a mouse model*. Genes Immun, 2004. **5**(3): p. 168-75.
193. Min-Oo, G., et al., *Complex genetic control of susceptibility to malaria: positional cloning of the Char9 locus*. J Exp Med, 2007. **204**(3): p. 511-24.
194. Fortin, A., et al., *Identification of a new malaria susceptibility locus (Char4) in recombinant congenic strains of mice*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10793-8.
195. Lee, P.D., et al., *Mapping cis-acting regulatory variation in recombinant congenic strains*. Physiol Genomics, 2006. **25**(2): p. 294-302.
196. Yamamoto, F., et al., *Molecular genetic basis of the histo-blood group ABO system*. Nature, 1990. **345**(6272): p. 229-33.
197. Lander, E.S. and D. Botstein, *Mapping mendelian factors underlying quantitative traits using RFLP linkage maps*. Genetics, 1989. **121**(1): p. 185-99.
198. Haley, C.S. and S.A. Knott, *A simple regression method for mapping quantitative trait loci in line crosses using flanking markers*. Heredity (Edinb), 1992. **69**(4): p. 315-24.

199. Lander, E.S. and N.J. Schork, *Genetic dissection of complex traits*. Science, 1994. **265**(5181): p. 2037-48.
200. Ewens, W.J. and R.S. Spielman, *The transmission/disequilibrium test: history, subdivision, and admixture*. Am J Hum Genet, 1995. **57**(2): p. 455-64.
201. Pritchard, J.K. and N.A. Rosenberg, *Use of unlinked genetic markers to detect population stratification in association studies*. Am J Hum Genet, 1999. **65**(1): p. 220-8.
202. Patterson, N., A.L. Price, and D. Reich, *Population structure and eigenanalysis*. PLoS Genet, 2006. **2**(12): p. e190.
203. Jayakar, S.D., *On the detection and estimation of linkage between a locus influencing a quantitative character and a marker locus*. Biometrics, 1970. **26**(3): p. 451-64.
204. Haseman, J.K. and R.C. Elston, *The investigation of linkage between a quantitative trait and a marker locus*. Behav Genet, 1972. **2**(1): p. 3-19.
205. Goldgar, D.E., *Multipoint analysis of human quantitative genetic variation*. Am J Hum Genet, 1990. **47**(6): p. 957-67.
206. Schork, N.J., *Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations*. Am J Hum Genet, 1993. **53**(6): p. 1306-19.
207. Amos, C.I., *Robust variance-components approach for assessing genetic linkage in pedigrees*. Am J Hum Genet, 1994. **54**(3): p. 535-43.
208. Blangero, J. and L. Almasy, *Multipoint oligogenic linkage analysis of quantitative traits*. Genet Epidemiol, 1997. **14**(6): p. 959-64.
209. Comuzzie, A.G., et al., *A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2*. Nat Genet, 1997. **15**(3): p. 273-6.
210. Almasy, L. and J. Blangero, *Multipoint quantitative-trait linkage analysis in general pedigrees*. Am J Hum Genet, 1998. **62**(5): p. 1198-211.
211. Yu, J., et al., *A unified mixed-model method for association mapping that accounts for multiple levels of relatedness*. Nat Genet, 2006. **38**(2): p. 203-8.
212. Kang, H.M., et al., *Efficient control of population structure in model organism association mapping*. Genetics, 2008. **178**(3): p. 1709-23.
213. Darvasi, A., *Genomics: Gene expression meets genetics*. Nature, 2003. **422**(6929): p. 269-70.
214. Broman, K.W., *Mapping expression in randomized rodent genomes*. Nat Genet, 2005. **37**(3): p. 209-10.
215. Jansen, R.C. and J.P. Nap, *Genetical genomics: the added value from segregation*. Trends Genet, 2001. **17**(7): p. 388-91.
216. Brem, R.B., et al., *Genetic dissection of transcriptional regulation in budding yeast*. Science, 2002. **296**(5568): p. 752-5.
217. Cheung, V.G., et al., *Natural variation in human gene expression assessed in lymphoblastoid cells*. Nat Genet, 2003. **33**(3): p. 422-5.
218. Monks, S.A., et al., *Genetic inheritance of gene expression in human cell lines*. Am J Hum Genet, 2004. **75**(6): p. 1094-105.

219. Morley, M., et al., *Genetic analysis of genome-wide variation in human gene expression*. Nature, 2004. **430**(7001): p. 743-7.
220. Schadt, E.E., et al., *Genetics of gene expression surveyed in maize, mouse and man*. Nature, 2003. **422**(6929): p. 297-302.
221. Buckland, P.R., *Allele-specific gene expression differences in humans*. Hum Mol Genet, 2004. **13 Spec No 2**: p. R255-60.
222. Knight, J.C., *Regulatory polymorphisms underlying complex disease traits*. J Mol Med (Berl), 2005. **83**(2): p. 97-109.
223. de Koning, D.J. and C.S. Haley, *Genetical genomics in humans and model organisms*. Trends Genet, 2005. **21**(7): p. 377-81.
224. Gibson, G. and B. Weir, *The quantitative genetics of transcription*. Trends Genet, 2005. **21**(11): p. 616-23.
225. Sladek, R. and T.J. Hudson, *Elucidating cis- and trans-regulatory variation using genetical genomics*. Trends Genet, 2006. **22**(5): p. 245-50.
226. WHO. *Smallpox Disease Facts*.
227. Breman, J.G. and D.A. Henderson, *Poxvirus dilemmas--monkeypox, smallpox, and biologic terrorism*. N Engl J Med, 1998. **339**(8): p. 556-9.
228. Dupuis, S., et al., *Impaired response to interferon-alpha/beta and lethal viral disease in human STAT1 deficiency*. Nat Genet, 2003. **33**(3): p. 388-91.
229. Lee, S.H., et al., *Maneuvering for advantage: the genetics of mouse susceptibility to virus infection*. Trends Genet, 2003. **19**(8): p. 447-57.
230. Tabeta, K., et al., *The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9*. Nat Immunol, 2006. **7**(2): p. 156-64.
231. Casrouge, A., et al., *Herpes simplex virus encephalitis in human UNC-93B deficiency*. Science, 2006. **314**(5797): p. 308-12.
232. Buller, R.M. and G.J. Palumbo, *Poxvirus pathogenesis*. Microbiol Rev, 1991. **55**(1): p. 80-122.
233. Brownstein, D.G., *Comparative genetics of resistance to viruses*. Am J Hum Genet, 1998. **62**(2): p. 211-4.
234. Brownstein, D.G., P.N. Bhatt, and L. Gras, *Ectromelia virus replication in major target organs of innately resistant and susceptible mice after intravenous infection*. Arch Virol, 1993. **129**(1-4): p. 65-75.
235. Brownstein, D.G. and L. Gras, *Chromosome mapping of Rmp-4, a gonad-dependent gene encoding host resistance to mousepox*. J Virol, 1995. **69**(11): p. 6958-64.
236. Boivin, G.A., et al., *Mapping of clinical and expression quantitative trait loci in a sex-dependent effect of host susceptibility to mouse-adapted influenza H3N2/HK/1/68*. J Immunol, 2012. **188**(8): p. 3949-60.
237. Team, R.C., *R: A language and environment for statistical computing*. 2014, R Foundation for Statistical Computing: Vienna, Austria.
238. Benjamini, Y.H., Y. , *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society Series B, 1995(57): p. 289-300.

239. Wallace, G.D., R.M. Buller, and H.C. Morse, 3rd, *Genetic determinants of resistance to ectromelia (mousepox) virus-induced mortality*. J Virol, 1985. **55**(3): p. 890-1.
240. Lohwasser, S., et al., *The non-classical MHC class I molecule Qa-1(b) inhibits classical MHC class I-restricted cytotoxicity of cytotoxic T lymphocytes*. Int Immunol, 2001. **13**(3): p. 321-7.
241. Organization, W.H. *The top 10 causes of death*. Fact sheet No. 310. 2008.
242. Organization, W.H. *Influenza (seasonal)*. Fact sheet No. 21. 2009.
243. Brown, E.G., *Influenza virus genetics*. Biomed Pharmacother, 2000. **54**(4): p. 196-209.
244. Johnson, N.P. and J. Mueller, *Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic*. Bull Hist Med, 2002. **76**(1): p. 105-15.
245. Taubenberger, J.K., *The origin and virulence of the 1918 "Spanish" influenza virus*. Proc Am Philos Soc, 2006. **150**(1): p. 86-112.
246. Klein, S.L. and S. Huber, *Sex Differences in Susceptibility to Viral Infection*, in *Sex Hormones and Immunity to Infection*, L.S. Klein and C. Roberts, Editors. 2010, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 93-122.
247. Klein, S.L., et al., *The impact of sex, gender and pregnancy on 2009 H1N1 disease*. Biol Sex Differ, 2010. **1**(1): p. 5.
248. Aho, K., R. Pyhala, and R. Visakorpi, *ABO associated genetic determinant in H1N1 influenza*. Tissue Antigens, 1980. **16**(4): p. 310-3.
249. Lebiush, M., L. Rannon, and J.D. Kark, *The relationship between epidemic influenza (A(H1N1) and ABO blood group*. J Hyg (Lond), 1981. **87**(1): p. 139-46.
250. Togashi, T., et al., *Influenza-associated acute encephalopathy in Japanese children in 1994-2002*. Virus Res, 2004. **103**(1-2): p. 75-8.
251. Albright, F.S., et al., *Evidence for a heritable predisposition to death due to influenza*. J Infect Dis, 2008. **197**(1): p. 18-24.
252. Horby, P., et al., *What is the evidence of a role for host genetics in susceptibility to influenza A/H5N1?* Epidemiol Infect, 2010. **138**(11): p. 1550-8.
253. Weiss, L.A., et al., *The sex-specific genetic architecture of quantitative traits in humans*. Nat Genet, 2006. **38**(2): p. 218-22.
254. Alberts, R., et al., *Gene expression changes in the host response between resistant and susceptible inbred mouse strains after influenza A infection*. Microbes Infect, 2010. **12**(4): p. 309-18.
255. Tumpey, T.M., et al., *Pathogenicity of influenza viruses with genes from the 1918 pandemic virus: functional roles of alveolar macrophages and neutrophils in limiting virus replication and mortality in mice*. J Virol, 2005. **79**(23): p. 14933-44.
256. Lenschow, D.J., et al., *IFN-stimulated gene 15 functions as a critical antiviral molecule against influenza, herpes, and Sindbis viruses*. Proc Natl Acad Sci U S A, 2007. **104**(4): p. 1371-6.

257. Suguitan, A.L., Jr., et al., *Live, attenuated influenza A H5N1 candidate vaccines provide broad cross-protection in mice and ferrets*. PLoS Med, 2006. **3**(9): p. e360.
258. Szretter, K.J., et al., *Role of host cytokine responses in the pathogenesis of avian H5N1 influenza viruses in mice*. J Virol, 2007. **81**(6): p. 2736-44.
259. Lindenmann, J., *Inheritance of Resistance to Influenza Virus in Mice*. Proc Soc Exp Biol Med, 1964. **116**: p. 506-9.
260. Tumpey, T.M., et al., *The Mx1 gene protects mice against the pandemic 1918 and highly lethal human H5N1 influenza viruses*. J Virol, 2007. **81**(19): p. 10818-21.
261. Salomon, R., et al., *Mx1 gene protects mice against the highly lethal human H5N1 influenza virus*. Cell Cycle, 2007. **6**(19): p. 2417-21.
262. Brown, E.G., et al., *Pattern of mutation in the genome of influenza A virus on adaptation to increased virulence in the mouse lung: identification of functional themes*. Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6883-8.
263. Brown, E.G., *Increased virulence of a mouse-adapted variant of influenza A/FM/1/47 virus is controlled by mutations in genome segments 4, 5, 7, and 8*. J Virol, 1990. **64**(9): p. 4523-33.
264. Guillot, L., et al., *Enhanced innate immune responsiveness to pulmonary Cryptococcus neoformans infection is associated with resistance to progressive infection*. Infect Immun, 2008. **76**(10): p. 4745-56.
265. Keleta, L., et al., *Experimental evolution of human influenza virus H3 hemagglutinin in the mouse lung identifies adaptive regions in HA1 and HA2*. J Virol, 2008. **82**(23): p. 11599-608.
266. Richer, E., et al., *N-ethyl-N-nitrosourea-induced mutation in ubiquitin-specific peptidase 18 causes hyperactivation of IFN- α signaling and suppresses STAT4-induced IFN- γ production, resulting in increased susceptibility to Salmonella typhimurium*. J Immunol, 2010. **185**(6): p. 3593-601.
267. Cobben, N.A., et al., *Diagnostic value of BAL fluid cellular profile and enzymes in infectious pulmonary disorders*. Eur Respir J, 1999. **14**(3): p. 496-502.
268. Torun, S.H., et al., *Clinical and epidemiological characteristics of pandemic influenza A/(H1N1) in hospitalized pediatric patients at a university hospital, Istanbul, Turkey*. J Trop Pediatr, 2011. **57**(3): p. 213-6.
269. Venkataraman, C., et al., *Death receptor-6 regulates the development of pulmonary eosinophilia and airway inflammation in a mouse model of asthma*. Immunol Lett, 2006. **106**(1): p. 42-7.
270. Kruse, S., et al., *The Ile198Thr and Ala379Val variants of plasmatric PAF-acetylhydrolase impair catalytical activities and are associated with atopy and asthma*. Am J Hum Genet, 2000. **66**(5): p. 1522-30.
271. Ninio, E., et al., *Platelet-activating factor-acetylhydrolase and PAF-receptor gene haplotypes in relation to future cardiovascular event in patients with coronary artery disease*. Hum Mol Genet, 2004. **13**(13): p. 1341-51.

272. Hussell, T., A. Pennycook, and P.J. Openshaw, *Inhibition of tumor necrosis factor reduces the severity of virus-specific lung immunopathology*. Eur J Immunol, 2001. **31**(9): p. 2566-73.
273. de Jong, M.D., et al., *Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia*. Nat Med, 2006. **12**(10): p. 1203-7.
274. Rimmelzwaan, G.F., et al., *Pathogenesis of influenza A (H5N1) virus infection in a primate model*. J Virol, 2001. **75**(14): p. 6687-91.
275. Lipatov, A.S., et al., *Pathogenesis of Hong Kong H5N1 influenza virus NS gene reassortants in mice: the role of cytokines and B- and T-cell responses*. J Gen Virol, 2005. **86**(Pt 4): p. 1121-30.
276. Daidoji, T., et al., *H5N1 avian influenza virus induces apoptotic cell death in mammalian airway epithelial cells*. J Virol, 2008. **82**(22): p. 11294-307.
277. Styrt, B. and B. Sugarman, *Estrogens and infection*. Rev Infect Dis, 1991. **13**(6): p. 1139-50.
278. Verthelyi, D., *Sex hormones as immunomodulators in health and disease*. Int Immunopharmacol, 2001. **1**(6): p. 983-93.
279. Yeretssian, G., et al., *Gender differences in expression of the human caspase-12 long variant determines susceptibility to Listeria monocytogenes infection*. Proc Natl Acad Sci U S A, 2009. **106**(22): p. 9016-20.
280. Baba, A., T. Fujita, and N. Tamura, *Sexual dimorphism of the fifth component of mouse complement*. J Exp Med, 1984. **160**(2): p. 411-9.
281. Caslake, M.J. and C.J. Packard, *Lipoprotein-associated phospholipase A2 as a biomarker for coronary disease and stroke*. Nat Clin Pract Cardiovasc Med, 2005. **2**(10): p. 529-35.
282. Robinson, D.P., et al., *Elevated 17 β -Estradiol Protects Females from Influenza A Virus Pathogenesis by Suppressing Inflammatory Responses*. PLoS Pathog, 2011. **7**(7): p. e1002149.
283. De Keyser, D., et al., *Increased PAFAH and oxidized lipids are associated with inflammation and atherosclerosis in hypercholesterolemic pigs*. Arterioscler Thromb Vasc Biol, 2009. **29**(12): p. 2041-6.
284. Camussi, G., et al., *Acute lung inflammation induced in the rabbit by local instillation of 1-O-octadecyl-2-acetyl-sn-glycerol-3-phosphorylcholine or of native platelet-activating factor*. Am J Pathol, 1983. **112**(1): p. 78-88.
285. Soares, P.M., et al., *Role of platelet-activating factor in the pathogenesis of 5-fluorouracil-induced intestinal mucositis in mice*. Cancer Chemother Pharmacol, 2011. **68**(3): p. 713-20.
286. Garland, R.C., et al., *Noninvasive molecular imaging reveals role of PAF in leukocyte-endothelial interaction in LPS-induced ocular vascular injury*. FASEB J, 2011. **25**(4): p. 1284-94.
287. Wilensky, R.L., et al., *Inhibition of lipoprotein-associated phospholipase A2 reduces complex coronary atherosclerotic plaque development*. Nat Med, 2008. **14**(10): p. 1059-66.

288. Peng, T., et al., *Role of C5 in the development of airway inflammation, airway hyperresponsiveness, and ongoing airway response*. J Clin Invest, 2005. **115**(6): p. 1590-600.
289. Demos, C., M. Bandyopadhyay, and B. Rohrer, *Identification of candidate genes for human retinal degeneration loci using differentially expressed genes from mouse photoreceptor dystrophy models*. Mol Vis, 2008. **14**: p. 1639-49.
290. Yasuda, K. and J.M. Johnston, *The hormonal regulation of platelet-activating factor-acetylhydrolase in the rat*. Endocrinology, 1992. **130**(2): p. 708-16.
291. Blake, G.J., et al., *A prospective evaluation of lipoprotein-associated phospholipase A(2) levels and the risk of future cardiovascular events in women*. J Am Coll Cardiol, 2001. **38**(5): p. 1302-6.
292. Pan, G., et al., *Identification and functional characterization of DR6, a novel death domain-containing TNF receptor*. FEBS Lett, 1998. **431**(3): p. 351-6.
293. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform*. Bioinformatics, 2009. **25**(14): p. 1754-60.
294. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. Bioinformatics, 2009. **25**(16): p. 2078-9.
295. Li, H., *Improving SNP discovery by base alignment quality*. Bioinformatics, 2011. **27**(8): p. 1157-8.
296. Cingolani, P., et al., *A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3*. Fly (Austin), 2012. **6**(2): p. 80-92.
297. Cingolani, P., et al., *Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift*. Front Genet, 2012. **3**: p. 35.
298. Patel, S.N., et al., *C5 deficiency and C5a or C5aR blockade protects against cerebral malaria*. J Exp Med, 2008. **205**(5): p. 1133-43.
299. Nissan, M.H., et al., *Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence*. Cancer Res, 2014. **74**(8): p. 2340-50.
300. Blazejewska, P., et al., *Pathogenicity of different PR8 influenza A virus variants in mice is determined by both viral and host factors*. Virology, 2011. **412**(1): p. 36-45.
301. Gurevich, I., et al., *TNIP1, a retinoic acid receptor corepressor and A20-binding inhibitor of NF-kappaB, distributes to both nuclear and cytoplasmic locations*. J Histochem Cytochem, 2011. **59**(12): p. 1101-12.
302. Allanore, Y., et al., *Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis*. PLoS Genet, 2011. **7**(7): p. e1002091.
303. Nair, R.P., et al., *Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways*. Nat Genet, 2009. **41**(2): p. 199-204.

304. Bowes, J., et al., *Confirmation of TNIP1 and IL23A as susceptibility loci for psoriatic arthritis*. Ann Rheum Dis, 2011. **70**(9): p. 1641-4.
305. Gateva, V., et al., *A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus*. Nat Genet, 2009. **41**(11): p. 1228-33.
306. Han, J.W., et al., *Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus*. Nat Genet, 2009. **41**(11): p. 1234-7.
307. Weber, B.N., et al., *A critical role for TCF-1 in T-lineage specification and differentiation*. Nature, 2011. **476**(7358): p. 63-8.
308. Bollag, G., et al., *Loss of NF1 results in activation of the Ras signaling pathway and leads to aberrant growth in haematopoietic cells*. Nat Genet, 1996. **12**(2): p. 144-8.
309. Wallace, M.R., et al., *Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients*. Science, 1990. **249**(4965): p. 181-6.
310. Ding, L., et al., *Somatic mutations affect key pathways in lung adenocarcinoma*. Nature, 2008. **455**(7216): p. 1069-75.
311. Largaespada, D.A., et al., *Nf1 deficiency causes Ras-mediated granulocyte/macrophage colony stimulating factor hypersensitivity and chronic myeloid leukaemia*. Nat Genet, 1996. **12**(2): p. 137-43.
312. Tassabehji, M., et al., *Tandem duplication within a neurofibromatosis type 1 (NF1) gene exon in a family with features of Watson syndrome and Noonan syndrome*. Am J Hum Genet, 1993. **53**(1): p. 90-5.
313. Rajagopal, S. and J. Treanor, *Pandemic (avian) influenza*. Semin Respir Crit Care Med, 2007. **28**(2): p. 159-70.
314. Stafforini, D.M., *Biology of platelet-activating factor acetylhydrolase (PAF-AH, lipoprotein associated phospholipase A2)*. Cardiovasc Drugs Ther, 2009. **23**(1): p. 73-83.
315. Rosenson, R.S. and D.M. Stafforini, *Modulation of oxidative stress, inflammation, and atherosclerosis by lipoprotein-associated phospholipase A2*. J Lipid Res, 2012. **53**(9): p. 1767-82.
316. Jiang, Z., et al., *The effect of lipoprotein-associated phospholipase A2 deficiency on pulmonary allergic responses in Aspergillus fumigatus sensitized mice*. Respir Res, 2012. **13**: p. 100.
317. Wiltshire, S.A., G.A. Leiva-Torres, and S.M. Vidal, *Quantitative trait locus analysis, pathway analysis, and consomic mapping show genetic variants of Tnni3k, Fpgt, or H28 control susceptibility to viral myocarditis*. J Immunol, 2011. **186**(11): p. 6398-405.
318. Tew, D.G., et al., *Purification, properties, sequencing, and cloning of a lipoprotein-associated, serine-dependent phospholipase involved in the oxidative modification of low-density lipoproteins*. Arterioscler Thromb Vasc Biol, 1996. **16**(4): p. 591-9.
319. Smyth, G., *Limma: linear models for microarray data*, ed. R.C. Gentleman, V.; Dudoit, S.; Irizarry, R.; Huber, W. 2005, New York: Springer.

320. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics*. Genome Biol, 2004. **5**(10): p. R80.
321. He, S., et al., *Lp-PLA2 Antagonizes Left Ventricular Healing After Myocardial Infarction by Impairing the Appearance of Reparative Macrophages*. Circ Heart Fail, 2015. **8**(5): p. 980-7.
322. Kireeva, M.L., S.C. Lam, and L.F. Lau, *Adhesion of human umbilical vein endothelial cells to the immediate-early gene product Cyr61 is mediated through integrin alphavbeta3*. J Biol Chem, 1998. **273**(5): p. 3090-6.
323. Chen, N., C.C. Chen, and L.F. Lau, *Adhesion of human skin fibroblasts to Cyr61 is mediated through integrin alpha 6beta 1 and cell surface heparan sulfate proteoglycans*. J Biol Chem, 2000. **275**(32): p. 24953-61.
324. Chen, C.C., N. Chen, and L.F. Lau, *The angiogenic factors Cyr61 and connective tissue growth factor induce adhesive signaling in primary human skin fibroblasts*. J Biol Chem, 2001. **276**(13): p. 10443-52.
325. Walsh, C.T., et al., *Thrombin receptor and RhoA mediate cell proliferation through integrins and cysteine-rich protein 61*. FASEB J, 2008. **22**(11): p. 4011-21.
326. Choi, J.S., K.H. Kim, and L.F. Lau, *The matricellular protein CCN1 promotes mucosal healing in murine colitis through IL-6*. Mucosal Immunol, 2015. **8**(6): p. 1285-96.
327. Kim, S.C., et al., *TRIM72 is required for effective repair of alveolar epithelial cell wounding*. Am J Physiol Lung Cell Mol Physiol, 2014. **307**(6): p. L449-59.
328. Jennings, R.T., et al., *RhoA determines disease progression by controlling neutrophil motility and restricting hyperresponsiveness*. Blood, 2014. **123**(23): p. 3635-45.
329. Li, C., et al., *The Immune Adaptor ADAP Regulates Reciprocal TGF-beta1-Integrin Crosstalk to Protect from Influenza Virus Infection*. PLoS Pathog, 2015. **11**(4): p. e1004824.
330. Jolly, L., et al., *Influenza promotes collagen deposition via alphavbeta6 integrin-mediated transforming growth factor beta activation*. J Biol Chem, 2014. **289**(51): p. 35246-63.
331. Hansson, G.K. and P. Libby, *The immune response in atherosclerosis: a double-edged sword*. Nat Rev Immunol, 2006. **6**(7): p. 508-19.
332. Fish, E.N., *The X-files in immunity: sex-based differences predispose immune responses*. Nat Rev Immunol, 2008. **8**(9): p. 737-44.
333. Roberts, C.W., W. Walker, and J. Alexander, *Sex-associated hormones and immunity to protozoan parasites*. Clin Microbiol Rev, 2001. **14**(3): p. 476-88.
334. Straub, R.H., *The complex role of estrogens in inflammation*. Endocr Rev, 2007. **28**(5): p. 521-74.
335. Hatoum, I.J., et al., *Dietary, lifestyle, and clinical predictors of lipoprotein-associated phospholipase A2 activity in individuals without coronary artery disease*. Am J Clin Nutr, 2010. **91**(3): p. 786-93.

336. Yoshimura, T., et al., *Estrogen replacement therapy decreases platelet-activating factor-acetylhydrolase activity in post-menopausal women*. Maturitas, 1999. **31**(3): p. 249-53.
337. Ohshige, A., et al., *Effects of estrogen and progesterone on plasma platelet-activating factor-acetylhydrolase activity and low-density lipoprotein cholesterol concentration in men*. Artery, 1996. **22**(3): p. 115-24.
338. Camateros, P., et al., *Identification of novel chromosomal regions associated with airway hyperresponsiveness in recombinant congenic strains of mice*. Mamm Genome, 2010. **21**(1-2): p. 28-38.
339. Hu, D., et al., *Analysis of regulatory CD8 T cells in Qa-1-deficient mice*. Nat Immunol, 2004. **5**(5): p. 516-23.
340. Lu, L., M.B. Werneck, and H. Cantor, *The immunoregulatory effects of Qa-1*. Immunol Rev, 2006. **212**: p. 51-9.
341. Steen, D.L. and M.L. O'Donoghue, *Lp-PLA2 Inhibitors for the Reduction of Cardiovascular Events*. Cardiol Ther, 2013. **2**(2): p. 125-34.
342. Hassan, M., *STABILITY and SOLID-TIMI 52: Lipoprotein associated phospholipase A2 (Lp-PLA2) as a biomarker or risk factor for cardiovascular diseases*. Glob Cardiol Sci Pract, 2015. **2015**: p. 6.
343. Mohler, E.R., 3rd, et al., *The effect of darapladib on plasma lipoprotein-associated phospholipase A2 activity and cardiovascular biomarkers in patients with stable coronary heart disease or coronary heart disease risk equivalent: the results of a multicenter, randomized, double-blind, placebo-controlled study*. J Am Coll Cardiol, 2008. **51**(17): p. 1632-41.
344. Jang, Y., et al., *Carriage of the V279F null allele within the gene encoding Lp-PLA(2) is protective from coronary artery disease in South Korean males*. PLoS One, 2011. **6**(4): p. e18208.