2016

Transcriptome Analysis of Vaccine Responses to Francisella Tularensis or Venezuelan Equine Encephalitis Virus

Rebecca Ann Erwin-Cohen
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

Rebecca Ann Erwin-Cohen

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Ji Shen, Committee Chairperson, Public Health Faculty
Dr. Bin Cai, Committee Member, Public Health Faculty
Dr. James Rohrer, University Reviewer, Public Health Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2016

Abstract

Transcriptome Analysis of Vaccine Responses to *Francisella Tularensis* or Venezuelan Equine Encephalitis Virus

by

Rebecca A. Erwin-Cohen

MS, Hood College, 1995

BS, Gannon University, 1990

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health -- Epidemiology

Walden University

December 2016

Abstract

The lack of vaccines for emerging and re-emerging diseases highlights technical gaps and indicates a need for innovative approaches to produce new vaccines. Vaccines may be improved by knowledge of host responses to vaccination, disease pathogenesis, and the effect of age and genetics on vaccine outcome. This study's purpose was to quantitatively assess the molecular epidemiology of *Francisella tularensis* (Ft) and Venezuelan Equine Encephalitis Virus (VEEV). Study results support the Epidemiology Nexus model which holds that association of changes in gene expression to vaccination facilitate understanding the mechanisms of immune development and link public health and disease epidemiology. My research questions assessed the relationship between gene expression following vaccination, the relationship between age and vaccine response, and the association between Human Leukocyte Antigen (HLA) allele and vaccine response. The study was a novel secondary analysis of human data subjected to ANOVA to measure association between treatment and outcome, correlation to measure association of age with vaccine outcome, and Mann-Whitney U tests to measure association of HLA allele with vaccine outcome. Both Ft and VEEV vaccination elicited significant changes in gene expression. A highly positive relationship between age and vaccine outcome was shown for VEEV. The results may affect positive social change by contributing to a growing compendium of evidence of vaccine efficacy mechanisms that may function to assure the public of vaccine safety, combat vaccine hesitancy, and promote vaccine acceptance, as well as contribute mechanistic knowledge to reduce developmental costs of novel vaccines.

Transcriptome Analysis of Vaccine Responses to *Francisella Tularensis* or Venezuelan Equine Encephalitis Virus

by

Rebecca A. Erwin-Cohen

MS, Hood College, 1995

BS, Gannon University, 1990

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health - Epidemiology

Walden University

December 2016

Dedication

I dedicate this work to my family; my beloved husband, Jeffrey, and our wonderful daughter, Aislinn. Their love, support, and encouragement made it possible to begin this scholarly journey and to see it through to the conclusion. I am grateful for the love and companionship of Daisy, Sookie, Walter, and BabbaLooey. I am also grateful for the encouragement and support of many friends, family, and colleagues, who excused my absences and absent-mindedness, lifted my spirits when down, and at the end of it all still want to spend time with me. Thank you all for supporting me.

Acknowledgments

I will be eternally grateful for the support and guidance of my committee. I would like to express my sincere thanks and gratitude to Dr. Ji Shen for his patience, wisdom, and guidance throughout the process of writing this dissertation; Dr. Bin Cai for his guidance on statistical analyses and interpretations; and Dr. James Rohrer for guidance and a keen eye for adhering to the guidelines and expectations of Walden.

I would like to thank Dr. Luis DaSilva for his mentoring, for teaching me transcriptome analysis, and for being one of the finest human beings I've ever had the good fortune to meet. Finally, I would like to thank COL Shelley Honnold for her mentoring, encouragement, critical review of chapters, and friendship.

The complete dissertation has been cleared for public presentation and dissemination under USAMRIID approval PR-17-009. The following human use protocols were associated with the work described herein: FY-05-01, FY-06-17. USAMRIID work has been funded through Department of Defense grant #05-4-8I-052 and Defense Threat Reduction Agency grant # 8.10022_05_RD_B.

Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Table of Contents

List of Tables vi

List of Figures viii

Chapter 1: Introduction to the Study 1

Introduction 1

Background 3

General Overview of *Francisella Tularensis* 4

General Overview of Venezuelan Equine Encephalitis Virus 10

Problem Statement 15

Study Purpose 18

Research Questions and Hypotheses 19

Research Question 1 19

Research Question 2 19

Research Question 3 20

Research Question 4 20

Conceptual Framework 21

Nature of the Study 22

Definitions 24

Assumptions 26

Scope and Delimitations. 26

Limitations 27

Significance 28

Summary 30
Chapter 2: Literature Review 32
Problem Addressed by the Study 32
Purpose of the Study 32
Literature Relevant to the Problem 33
Literature Search Strategy 34
Conceptual Framework 35
Theoretical Concepts 35
How the Current Study May Benefit From the Epidemiology Nexus Concept 42
Literature Review Related to Key Variables and Concepts 44
Previous Transcriptomic Studies of Human Responses to Ft or VEEV: Relation to Key Study Variables 44
Strengths and Weaknesses of Previous Experimental Approaches and Controversial Issues 52
Previous Studies of Association of Age or Genetic Makeup with Vaccine Outcome: Relation to Key Study Variables 59
Rationale for Selection of Key Variable and Unanswered Questions 60
Summary of Previous Studies and Relation to Research Questions 62
Summary and Conclusions 65
Chapter 3: Research Method 67
Study Purpose 67

Research Design and Rationale 67
Use of Secondary Datasets to Reduce Time and Resource Constraints 68
Refined Study Parameters Retain Consistency with Previous Studies 68
Methodology 69
Study Populations from Original Data 70
Inclusion and Exclusion Criteria Used in Original Study 72
Power Calculations and Sampling Size 73
Sampling Procedures of Original Study 74
Sampling in Original Study 74
Use of Archival Data 76
Research Questions and Hypotheses 76
Data Analysis Plan 78
Interpretation of Results 85
Threats to Validity 85
Ethical Procedures 87
Summary 88
Chapter 4: Results 90
Brief Overview 90
Research Questions and Hypotheses 90
Data Collection and Handling of Data Files 92
Demographic and Descriptive Data 100
Results 104

Descriptive Statistics for Transcriptome Data 104
Inferential Statistics 108
Analysis of Variance 108
Correlational Analysis of Relationship Between Age and Neutralizing Titer 117
Association Between HLA Allele and Post-Vaccination Titer 122
Additional Ad Hoc Analyses: Ingenuity Pathway Analysis 131
Summary 140
Answers to Research Questions Based on Observed Results 140
Chapter 5: Discussion, Conclusions, and Recommendations 143
Overview 143
Interpretation of the Findings 144
Association of Level of Gene Expression with Time and Treatment Following Vaccination 144
Analysis Interpretations for Ft Vaccinated Versus Controls 144
Analysis Interpretations for VEEV Vaccinated versus Controls 146
Analysis Interpretations for Comparison of Ft to VEEV Vaccinated Responses 147
Analysis Interpretations Relative to Epidemiology Nexus Conceptual Theory 151
Analysis Interpretations for Comparison of Age to Vaccine Responses 152
Analysis Interpretations for Comparison of HLA to Vaccine Responses 153

Limitations of the Study.....153

Recommendations.....155

Implications.....156

Positive Social Change 156

Conclusion 157

References.....160

Appendix A: Memorandum for Record, Determination of Not Human Subject Research.....179

List of Tables

Table 1. Ft Data Files Missing or Removed from Analysis 97

Table 2. VEEV Data Files Missing or Removed From Analysis 98

Table 3. Descriptive and Demographic Characteristics of Ft Study Participants..... 101

Table 4. Descriptive and Demographic Characteristics of VEEV Study Participants ... 102

Table 5. Descriptive Statistics for Ft Study: Log2 Transcript Expression Values by Study Subject..... 105

Table 6. Descriptive Statistics for VEEV Study: Log2 Transcript Expression Values by Study Subject 106

Table 7. Descriptive Statistics: Log2 Expression Values by Time Point (Includes Both Ft and VEEV)..... 107

Table 8. Descriptive Statistics: Log2 Expression Values by Treatment (Includes both Ft and VEEV)..... 108

Table 9. False Discovery Rate Report: ANOVA Analysis Contrasting Ft Vaccinated With Ft Control Subjects 111

Table 10. False Discovery Rate Report: ANOVA Analysis Contrasting VEEV Vaccinated With VEEV Control Subjects 113

Table 11. False Discovery Rate Report: ANOVA Analysis Contrasting Ft Vaccinated Against VEEV Vaccinated Subjects..... 115

Table 12. List of Genes in Top Canonical Pathways (Day 2-7) in VEEV Vaccine Study 134

Table 13. List of Genes in Top Canonical Pathways (Day 14) in VEEV Vaccine Study 136

Table 14. List of Genes in Top Canonical Pathways in Ft Compared to VEEV Vaccine Study 138

Table 14. List of Genes in Top Canonical Pathways in Ft Compared to VEEV Vaccine Study (Con't). 139

List of Figures

Figure 1. Schematic of original studies for Ft and VEEV vaccination.....93

Figure 2. Log2 expression of microarray hybridization controls 94

Figure 3. Log2 expression of microarray labeling controls.....95

Figure 4. 3'/5' *in vitro* transcription (IVT) ratio of microarray samples.....96

Figure 5. Principle component analysis.....99

Figure 6. Principle component analysis, by scan date 100

Figure 7. Correlation between age group and titer in Ft study 118

Figure 8. Correlation between age and titer in Ft study.....119

Figure 9. Correlation between age group and titer in VEEV study.....120

Figure 10. Correlation between age and titer in VEEV study 121

Figure 11. Correlation between age and titer in VEEV study (outlier removed)……122

Figure 12. HLA DRB1 allele and post-vaccination titer from the Ft study group 123

Figure 13. HLA DQB1 allele and post-vaccination titer from the Ft study group 124

Figure 14. HLA DRB1 allele and post-vaccination titer from the VEEV study group ...125

Figure 15. HLA DQB1 allele and post-vaccination titer from the VEEV study group...126

Figure 16. Pathway analysis of significant genes from Ft study 132

Figure 17. Top five canonical pathways for Days 2 and 7 in VEEV study.....133

Figure 18. Top five canonical pathways for Day 14 in VEEV study 135

Figure 19. Top five canonical pathways in the comparison of Ft to VEEV study 137

Chapter 1: Introduction to the Study

Introduction

Global public health efforts to prevent disease from both *Francisella tularensis* (Ft) and Venezuelan equine encephalitis virus (VEEV) have been hampered by the lack of approved human vaccines and the lack of knowledge of the human molecular responses to infectious disease agents (Foley & Nieto, 2010; Wolfe et al., 2013; Wolfe et al., 2014). Molecular epidemiology studies in other infectious diseases have provided innovative information in humans and animal models; however, there is little understanding or information of the molecular consequences of human infection with either Ft or VEEV. In this dissertation project, I sought to make a novel contribution to the field of public health by studying the relationship between gene expression and the temporal effects of vaccination, the potential relationship between age and vaccine outcome, and the potential association between genetic make-up of individuals with vaccine outcome.

This project may contribute to positive social change at multiple levels. First, transcriptomic analysis may provide evidence to support the Epidemiology Nexus conceptual model. I developed and named the model based on the collected works of several subject matter experts in the field of vaccinology and systems biology. I chose the word nexus as a descriptor to link the molecular epidemiology of the host responses to the traditional public health epidemiology of disease and specifically to link the molecular responses with a concrete social impact. The Epidemiology Nexus concept holds that examining changes in gene expression can contribute to a broad knowledge

base for understanding the molecular mechanisms of immune development and expand knowledge of the fundamental mechanisms of action of successful vaccination (Li et al., 2015; Kennedy et al., 2014; Nakaya, Li, & Pulendran, 2011; Pulendran et al., 2013). Secondly, my use of molecular epidemiology data for the explicit purpose of creating rationally-designed vaccines supports a positive social change agenda relating to administrative policy; financial incentives are often absent with regard to development of neglected tropical or zoonotic diseases and the World Health Organization (WHO) recognizes that developmental costs for vaccines and lack of correlates of protection are unmet challenges (Hortez et al., 2016). Studying the common and unique changes to gene expression that are elicited by vaccination may make it possible to develop multi-agent vaccines that share similar immune-related mechanisms. By using Ft and VEEV to compare and contrast molecular epidemiology, I worked to contribute new knowledge to develop vaccines which can meet rigorous standards for licensure. Licensed vaccines could have a positive social impact on society on both and individual and family levels through reduced disease incidence, and impact on communities through reduced mortality and morbidity (Hortez et al., 2016; Reichert et al., 2009, Wolfe et al., 2013, Wolfe et al., 2014). The compendium of evidence I have built regarding the mechanisms of vaccine efficacy may also serve to assure the public that vaccines are well-tested and safe. Such assurance has become increasingly important as public health professionals seek to understand and combat vaccine hesitancy and to promote vaccine acceptance (Eskola, Duclos, Schuster, MacDonald, & the SAGE Working Group on Vaccine Hesitancy, 2015).

In this chapter, I introduce the background and basic epidemiology of both Ft and VEEV, including descriptions of the disease that each agent causes, the impact of disease on morbidity and mortality, and what is known regarding the economic impact of disease. A discussion of the research questions that I sought to address will follow the discussion of the epidemiology, and I conclude with a brief introduction to the type of data that I used in the analysis.

Background

Many global infectious diseases exist without viable vaccines for prevention or therapeutic treatment to resolve the infection. Two examples of diseases which fit this description include tularemia, caused by the gram-negative bacterium Ft, and Venezuelan equine encephalitis (VEE), caused by VEEV. According to the U.S. Centers for Disease Control and Prevention (CDC), each year roughly 200 cases of tularemia were reported to the National Notifiable Disease Surveillance System (NNDSS) in the United States alone, although the niche for tularemia infection is not limited to the United States (NNDSS, 2016). Tularemia infection remains problematic in other parts of the world as well, with many human cases having been diagnosed in the past decade particularly in Norway, Russia, Australia, Germany, and Spain (HealthMap, 2016). VEE is mosquito-borne disease that is found primarily in Latin and South America, but has been detected in the southern United States during large epizootic outbreaks (Gubler, 2002). Since January of 2010, there have been more than 100 confirmed cases of VEE in the United States, with four confirmed deaths (HealthMap, 2016). The number of cases of VEE is believed to be greatly underreported because of the similarity of symptoms and presentation of disease

with those of Dengue fever. Indeed, Aguilar et al. (2011) reported that up to 10% of cases diagnosed as Dengue may actually be VEE cases which are not detected due to lack of specialized tests or limited resources for testing. One of the largest outbreaks of VEE occurred in 1995 in Venezuela and Columbia, and resulted in more than 75,000 cases of disease and approximately 300 deaths (Operational Biosurveillance, 2010). Both Ft and VEEV are categorized by the CDC as Category B biological threat agents and present a global public health threat.

General Overview of *Francisella Tularensis*

Basic epidemiology: The effect of Ft on disease, deaths, disability, and economic impact. The species of bacteria *Francisella tularensis* subspecies tularensis was first described in 1919 by Edward Francis as a gram-negative intracellular coccobacillus (Francis, 1919). The bacterium was named after both the researcher and the county in which it was studied as an endemic disease of rodents (Tulare County, CA) (Francis, 1919). The disease caused by infection with the bacteria was termed Tularemia, and the basic epidemiology of the disease has been studied extensively over the last century (Sjöstedt, 2007).

Tularemia disease has been described in several forms, each of which is dependent on the route of exposure. Forms of diseases have included ulceroglandular tularemia (the most common form), glandular tularemia, oculoglandular tularemia, oropharyngeal tularemia, intestinal tularemia, respiratory or inhalational tularemia (the most serious form), and typhoidal tularemia (NNDSS, 2016; Nelson, Kugeler, Petersen, & Mead, 2013). Ulceroglandular tularemia disease has been described as having been

produced by the bite of an infected tick, and disease presentation included localized cutaneous ulcers at the site of infection (Nelson et al., 2013). Researchers subsequently identified ticks as a common vector for transmission of percutaneous disease (Keim, Johansson, & Wagner, 2007) including three species of genus *ixodid* ticks, *Amblyomma americanum*, *Dermacentor variabilis*, and *Dermacentor andersoni*, which were shown to be responsible for disease transmission in North America in a predominantly terrestrial disease cycle (Keim et al., 2007). In other parts of the world, where researchers have found the Ft subspecies *holarctica* to be the predominant form of bacteria, the transmission vector has been shown to be more often mosquitos than ticks due to the aquatic nature of the transmission cycle (Ulu-Kilic & Doganay, 2014). Researchers have shown that the larval forms of mosquitos were infected with Ft *holarctica*, and thus adult mosquitos were primed with the ability to pass the bacteria to vertebrate hosts (Ulu-Kilic & Doganay, 2014).

Respiratory tularemia has been shown to be caused by accidental inhalation of bacteria from an infected host (typically rabbits or rodents; Staples, Kubota, Chalcraft, Mead, & Petersen, 2006). Researchers have described pneumonic tularemia, a complication of respiratory tularemia, as one of the most serious forms of the disease in humans, and death from infection has been shown to occur with inhalation of as few as 10 bacteria or colony forming units (CFU). The pathogenesis of human disease has included primary propagation of bacteria in macrophages, but also included widespread infection of lymphocytes, erythrocytes, and in plasma causing an intense anti-

inflammatory response in the localized pulmonary environment (Dennis et al., 2001; Ellis, Oyston, Green, & Titball, 2002; Horzempa et al., 2011).

The ecological niche where Ft bacteria have been found has had profound effects upon the incidence of disease. Early research efforts regarding the life-cycle of the bacteria were largely unsuccessful due to low incidence of human disease and small, pocket localizations of endemic areas of bacterial occurrence (Sjöstedt, 2007). Studies conducted in the United States, Europe, and parts of the middle east by Ariza-Miguel et al. (2014), Nakazawa et al. (2007, 2010), and Ulu-Kilic & Doganay (2014) showed results which were in agreement that the Ft bacteria persistently remained in localized environments, and that outbreaks were likely the result of human contact with the local foci, largely through hunting and outdoor activities. Nakazawa et al. (2007, 2010) used changes in climate patterns over the course of 50 years to construct a computer model of changing climate that paired disease incidence based on national surveillance reporting with environmental information consisting largely of variables that included minimum and maximum temperatures, average temperature, amount of rainfall, evaporation, and moisture surplus. This computer algorithm, termed the Genetic Algorithm for Rule-Set Prediction, used by Nakazawa et al. (as cited in Stockwell & Noble, 1992, Stockwell & Peters, 1999) had previously been tested and validated (Nakazawa et al., 2007).

Death and disability caused by Ft. In the United States from 2004 through 2013, there were on average 137.8 cases of tularemia per year, with an accumulated total of 1,378 (CDC, 2015a). The annual number of cases in recent times (~138 cases per year) was fewer than in the middle part of the 20th century when the annual number of reported

cases was between 500-1000 per year (CDC, 2015b; Staples et al., 2006). The reduced incidence of tularemia (predominantly manifested as ulceroglandular disease) was likely due to the gradual shift from a more agrarian society to a more urban society, and a reduction in animal to human contact thru small mammal hunting (Eliasson et al., 2006; Ellis et al., 2002). In contrast to the somewhat limited outbreaks of tularemia in the United States, there have been large outbreaks in Europe and the former Soviet Union during the mid-part of the 19th century to present; these outbreaks were largely manifested as ulceroglandular and oropharyngeal (Tärnvik, Priebe, & Grunow, 2006). An outbreak of more than 100,000 cases was reported in the Soviet Union during the late 1940s, but it was unclear whether the source of the epidemic was a result of biological warfare or local conditions (e.g., contaminated water or tick infestation; Tärnvik et al., 2006). An outbreak in Kosovo during the 1999-2000 conflict in that region resulted in over 300 cases of tularemia which were traced back to rodent infestation from unharvested crops; infected rodents contaminated local water supplies and most cases of tularemia observed were oropharyngeal infections (Tärnvik et al., 2006). In Spain over the period of 1997 through 2008, there were two distinct periods of outbreaks reported (1997-1998 and 2007-2008), each of which affected less than 1000 people (Ariza-Miguel et al., 2014). The first Spanish outbreak in 1997 was associated with small mammal hunting, with disease manifestation primarily observed as glandular or ulceroglandular tularemia (Ariza-Miguel et al., 2014). In contrast, a second major tularemia outbreak occurred in Spain between 2007-2008 following a population expansion of the common vole, a rodent amplification host commonly associated with typhoidal tularemia (Ariza-

Miguel et al., 2014; Rossow et al., 2014). Because the disease presentation in the majority of patients was characterized as primarily typhoidal infection, epidemiology investigators concluded that the infections were likely transmitted via the inhalational route (Ariza-Miguel et al., 2014). Surveillance conducted in France between 2002 and 2012 discovered small clusters of naturally occurring cases of tularemia infection, with a total of 433 cases reported, comprising roughly ten clusters of disease (Mailles & Vaillant, 2014). The majority of cases observed were glandular/ulceroglandular tularemia disease, which represented 72% of the disease cases and was associated most frequently with actions which involved handling of hares (41% of cases), typhoidal and pneumonic presentations of tularemia each represented 10% of the disease cases but were not associated with a specific exposure event (Mailles & Vaillant, 2014).

The disease burden caused by naturally occurring tularemia in the United States has been documented to be predominantly one of morbidity rather than of mortality (CDC, 2015a). In contrast, the impact of tularemia infection when used as a bioterrorism pathogen during World War II has been reported to have primarily been a lethal disease (Dennis et al., 2001; Oyston, Sjostedt, & Titball, 2004). The threat of Ft has been documented as a continued biological threat problem, particularly for military personnel who may deploy into areas where infection is endemic or re-emerging, into areas with shifting political/military stability, or in areas with potential for biological terrorism (Wolfe, Florence, & Bryant, 2013).

Economic impact of Ft infections. I found no published information regarding the incurred costs associated with naturally-occurring Ft infections. However, researchers

have conducted computer modeling and simulation studies for three biological threat agents to estimate the economic impact of an intentional large-scale aerosol delivery of *Bacillus anthracis*, *Brucella melitensis*, or *Francisella tularensis* in large city or suburban environments (Kaufmann, Meltzer, & Schmid, 1997). Kaufmann Meltzer, and Schmid (1997) noted that "the impact of a bioterrorist attack depends on the specific agent or toxin used, the method and efficiency of dispersal, the population exposed, the level of immunity in the population, the availability of effective postexposure and/or therapeutic regimens, and the potential for secondary transmissions" (p. 83). Intentional exposure to aerosolized Ft was estimated to cause a high rate of illness and death; for every 100,000 persons exposed to aerosolized Ft, Kaufmann et al. (1997) estimated that exposure would cause 82,500 cases of pneumonic tularemia and result in 6,188 deaths. In the assessment of costs, Kaufmann et al. (1997) included such variables as cost of premature human death, cost of hospitalization, cost of post-hospitalization outpatient visits, and cost of diagnostic tests, and then compared the burden of infection against the cost of intervention. Hospital costs for Ft infection were estimated to be on average $6,338 to $7,582 over the course of an average 20-day stay (Kaufmann et al., 1997). Non-hospitalization costs incurred for Ft infection (e.g., outpatient services, laboratory tests, lost productivity, etc.) were estimated to average between $722 and $1,120 per patient (Kaufmann et al., 1997). The cost to provide prophylaxis consisting of doxycycline and gentamicin treatment for Ft infection during a biological attack was estimated to average between $418,094 and $1,488,037, depending on the number of people who would need to be treated in suburban versus highly populated urban areas (Kaufmann et al., 1997).

Kaufmann et al. (1997) hypothesized that the cost of prophylaxis would provide a net savings over the economic losses that were simulated in the absence of treatment, with total costs of a biological attack with Ft ranging from $5.402 million to $5.507 million, with the simulated assumption that 100,000 people would be exposed.

General Overview of Venezuelan Equine Encephalitis Virus

Basic epidemiology. Venezuelan equine encephalitis is caused by infection with a positive-strand RNA virus, Venezuelan equine encephalitis virus. Discovery of the virus in the early part of the 20th century was preceded by large enzootic outbreaks of viral infection in equids (Beck &Wyckoff, 1938; Kubes & Rios, 1939). The occurrence of disease among agricultural animals was occasionally observed to spill over into human populations as well, particularly amongst agricultural workers; the infection events that caused disease in humans were distinguished from outbreaks that only affected animals (e.g., enzootic) and termed epizootic events (Adams et al., 2012). Researchers have discovered several subtypes of VEEV and have characterized them into categories enumerated IA through IF (Steele & Twenhafel, 2010). Individual strains have been described according to whether they were enzootic or epizootic (Steele & Twenhafel, 2010). Enzootic strains have been shown to circulate through mosquito-rodent cycles and include strain subtypes ID, IE, and IF (Weaver & Reisen, 2010). Horses and other equids have not been demonstrated to amplify enzootic viral strains, and enzootic strains are generally considered to be avirulent (Sahu, Pederson, Jenny, Schmitt, & Alstad, 2003; Steele & Twenhafel, 2010). In contrast, examination of epizootic strains that included subtypes IA/B and IC has shown that the lifecycle of epizootic subtypes included an

amplification cycle in equids, and resulted in high susceptibility to infection in both equids and humans (Steele & Twenhafel, 2010).

There have been three major epidemic outbreaks (1962-1964, 1992-1993, and 1995), which have resulted in human case fatalities; each of these events was associated with epizootic IC strains of VEEV (Brault et al., 2001). Human disease caused by VEEV has been described to range from very mild flu-like symptoms to severe clinical and neurological syndromes (Bowen, Fashinell, Dean, & Gregg, 1976; Ehrenkranz & Ventura, 1974). The most common symptoms observed or reported included complaints of sickness similar to influenza-like illness including fever, malaise, vomiting, sore throat, and lethargy (Rivas et al., 1997; Steele & Twenhafel, 2010). These symptoms comprised most mild cases of infection, which were typically resolved within three to five days after the initial symptoms appeared (Rivas et al., 1997). VEEV infection was reported to produce symptoms in nearly all human cases, regardless of the severity of infection (Steele & Twenhafel, 2010). Severe cases of infection, which were noted to have occurred predominantly in the very young or very old, were associated with case-fatality rates of approximately 0.5% of adults and up to 4% of children (Wolfe et al., 2014). Symptoms observed in cases of illness which progressed into serious clinical infection included fever, headache, photophobia, biphasic fever, seizures, ataxia, dysphasia, paresis, encephalitis, and pneumonia (Franz et al., 1997; de la Monte, Castro, Bonilla, Gaskin de Urdaneta, & Hutchins, 1985). Extensive involvement of lymphoid and endothelial tissues was also noted by de la Monte et al. (1985). The reported rates of overall severe infection varied in the literature. De la Monte et al. (1985) reported rates of

neurological disease ranging from 4-14% of VEEV cases. In contrast, Franz et al. (1997) reported a smaller percentage of VEEV cases that advanced to the severe neurological state (0.5% to 4%). However, those estimations appeared to be based on epidemic data from up to the 1970s and did not include the VEEV outbreaks in 1992-1993 or 1995.

Death and disability caused by VEEV. The observed case-fatality rate for VEEV in humans has been low (estimated to be between 0.5 and 4%); however, the economic impact of disability caused by infection remains largely unknown and unstudied (Wolfe, 2014). Little has been reported on the neurological outcomes associated with non-fatal infection. Bowen et al. (1976) and Ehrenkranz and Ventura (1974) both reported on the extent of clinical signs during the fulminant and encephalitic phases of illness in both children and adults infected during the 1969-1971 VEEV epidemic. Among the clinical signs observed, “confusion, hallucinations, or gait abnormality in association with excessive drowsiness” and serious consequences which included “grand mal seizures, temporary paralysis, or coma” were described for serious encephalitic forms of infection (Bowen et al., 1976, p. 49). Longer-term complications were reported in 12 patients (13.6% of patients) at 1 month and 9-12 months after initial infection. The complications reportedly included unilateral muscle weakness in extremities; reduced sensory perception of taste, smell, and hearing; recurrent headaches; easy fatigue; poor concentration/forgetfulness; and depression (Bowen et al., 1976). The long-term symptoms reported are in keeping with the general long-term neurological sequelae that have been reported for a similar encephalitic disease, West Nile virus (Weatherhead et al., 2015). Weatherhead et al. (2015) conducted an observational study

in which they examined the progression of neurological complications in 60 patients following West Nile infection at 1-3 years post-infection, and 8-11 years post-infection. A total of 35 patients were classified as having had encephalitis during acute phase infection; of those, 86% (or 30 patients) displayed abnormal neurological issues at the first post-infection assessment (Weatherhead et al., 2015). At the second post-infection assessment, "57% (4/7) of West Nile fever (WNF), 33% (2/6) of West Nile meningitis (WNM), and 36% (5/14) of West Nile encephalitis (WNE) had developed new neurological complications" (Weatherhead et al., 2015, p. 1006). The complications described by Weatherhead et al. (2015) were similar to those described by Bowen et al. (1976) following VEEV infection, and likely represented general long-term consequences of encephalitis. The common sequelae included unilateral abnormalities in gait, loss of hearing, loss of motor reflexes, and persistent weakness in muscles. Neither study was able to differentiate whether the long-term issues experienced by patients were due to neuroinvasive infection or if the abnormalities occurred as a result of comorbidities (Bowen et al., 1976; Weatherhead et al., 2015).

Economic impact of VEEV infections. The predominant economic losses associated with or due to enzootic infection in animals have been confined to the loss of agricultural animals, namely horses, donkeys, and mules (Navarro et al., 2005). Loss of production, productivity, and potential income for agricultural workers could have serious economic impact, depending on the local environment, but particularly in developing countries where farm animals may be the primary source of working the land (Aréchiga-Ceballos & Aguilar-Setién, 2015; Nara et al., 2010). I found no economic

information on the number of cases of humans affected post-infection with VEEV in terms of neurological sequelae or what, if any, economic impact their long-term care may have imposed. However, in one report on the economic burden sustained by survivors of a related alphavirus, eastern equine encephalitis virus (EEEV), researchers discussed the costs incurred by three individuals who experienced mild infection as well as three who survived infection-induced encephalitis (Villari, Spielman, Komar, McDowell, & Timperi, 1995). The cost/benefit analysis conducted by Villari et al. (1995) described the short-term (approximately one month) costs associated with the mild, transient cases of EEEV, including an average 20-day stay in the hospital, and calculated the average cost to be $21,051 per individual. For the three surviving individuals who experienced severe infection, all of whom were children, the neurological sequelae were remarkably similar in each case, and similar to other descriptions of VEEV long-term effects in children (Villari et al., 1995). For the encephalitis cases with long-term residual effects, the care costs began with longer, more extensive-care hospital stays (which averaged 86 days) than the milder cases of EEEV infection, rehabilitation hospital care (average stay 238 days), round-the-clock nursing care once the individuals were well enough to go home, educational costs (e.g., remedial therapies, transportation), and loss of lifetime earnings as the afflicted children had such severe disabilities that it was suggested that none would ever attain productive employment (Villari et al., 1995). In light of the severity and longevity of caring for such afflicted individuals, Villari et al. (1995) estimated that the average lifetime costs resulting from EEEV infection would be $2.95 million per person, ranging from a little as $2.54 million to as much as $4.28 million.

The economic impact of VEEV infection from an intentional biological attack has never been estimated.

The purpose of this dissertation project was to address a gap in knowledge by quantitatively assessing the molecular epidemiology of Ft and VEEV in a human vaccine model. I also worked to quantitatively assess the potential relationship between gene expression over time in response to treatment (vaccination) with either Ft or VEEV; to assess the potential relationship between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination or vaccine failure); and to identify a potential correlation between HLA phenotype and vaccine response. My research was intended to add to the limited published information about the changes in gene expression that each vaccine elicits, and address the gaps in knowledge regarding the association of age and genetic makeup on vaccine outcome.

Problem Statement

Vaccination has been a fundamental pillar for the prevention of infectious diseases and improvement of public health for roughly two centuries (Centlivre & Combadière, 2015; Kennedy, Ovsyannikova, Lambert, Haralambieva, & Poland, 2014; Wilson & Karp, 2015). However, emerging and re-emerging diseases such as tularemia, dengue fever, Ebola, equine encephalomyelitis, and Zika highlight gaps in the development and approval of new vaccines as well as the need to incorporate innovative approaches to address global health challenges to produce new vaccines (Hotez et al., 2016; Kennedy et al., 2014; Wilson & Karp, 2015). For many emerging or re-emerging diseases, traditional methods of vaccine development have sometimes failed to produce

effective vaccines. In these cases, many subject matter experts believe that knowledge of additional factors such the host innate and adaptive responses to vaccination, differences in genetic makeup, gender and age-related differences, and differences among diverse ethnic populations may provide evidence-based data needed to facilitate development strategies to overcome technological limitations, including identifying novel correlates of protection (Angel, Steele, & Franco, 2014; Kennedy et al., 2014; Hotez et al., 2016; Pulendran, Oh, Nakaya, Ravindran, & Kazmin, 2013). The lack of (a) knowledge surrounding the pathogenesis of disease in humans, (b) knowledge of vaccine outcome among diverse populations, and (c) knowledge of the effect of genetic makeup in relation to vaccine outcome have emerged as critical barriers to producing effective vaccines, particularly against emerging and re-emerging infectious diseases (Klein and Poland, 2013). Researchers have conducted transcriptome studies of human responses to yellow fever vaccine, smallpox, and patients with naturally occurring upper-respiratory infection to quantitatively assess changes in gene expression and associate specific patterns of gene expression with molecular signatures for each of the viruses (Scherer et al., 2007). Similarly, in a study to assess transcriptional changes induced by infection with nontypeable *Haemophilus influenza* bacteria in humans, Baddal et al. (2015) highlighted pathogen-induced regulatory changes that facilitated pathogenesis and inflammation, and identified potential targets for infection control. In recent work with West Nile virus neuroinvasive disease, researchers have shown specific gene variants associated with seroconversion and HLA alleles that were associated with immunoprotection (Long, Deng, Singh, Loeb, Lauring, & Seielstad, 2016). For Ft and VEEV, two important re-

emerging pathogens, there have been small studies published describing gene expression profiles in peripheral blood mononuclear cells (PBMCs); however, neither study addressed the issue of the host response to vaccination from whole blood and the sampling time frames are dissimilar for comparison purposes (Andersson et al., 2006; Erwin-Cohen, Porter, Pittman, Rossi, & DaSilva, 2012).

There are several studies which have shown the association of age with vaccine outcome (Pawelec & Derhovanessian, 2010). The response of elderly people to influenza vaccination is greatly reduced (17-53% efficacy) in comparison to young adults who demonstrate between 70-90% efficacy in response to influenza vaccination (Pawelec & Derhovanessian, 2010). In a study of naïve aging individuals immunized with Hepatitis B vaccine (HBV), age was significantly associated with vaccine response; vaccinees aged 40 years or younger responded better to vaccination than did individuals who were 65 years of age or older (Fourati et al., 2015). The mechanisms of reduced vaccine response may be related to declining T cell populations in older individuals, as elderly men were found to have reduced T cells populations, in comparison to younger individuals (De Benedetto et al., 2015). There are no studies, to date, that have addressed a potential relationship between vaccine outcomes and age for vaccination with Ft. Pittman, Liu, Cannon, Mangiafico, and Gibbs (2009) discussed a small study in which age was investigated as a demographic factor for vaccine response to VEEV; however the authors reported only that there was no evidence of age being associated with primary vaccine failure in individuals who had previously been vaccinated with two other vaccines which are related to VEEV, eastern and western equine encephalitis viruses. Several researchers

have reported the association of specific genetic factors with vaccine outcome in response to vaccination with childhood vaccines (HBV, diptheria, tetanus, pertussis [DTaP]), smallpox, influenza, anthrax, and rubella (Li, Nie, & Zhuang, 2013; Narwaney et al., 2013; Ovsyannikova, Jacobson, Ryan, Dhiman, Vierkant, & Poland, 2007; Ovsyannikova et al., 2013; Ovsyannikova, Pankratz, Salk, Kennedy, & Poland, 2014; Yucesoy et al., 2013). HLA are encoded by the major histocompatibility complex (MHC) genes (Li et al., 2013). HLA class II molecules are responsible for presentation of antigens to ultimately elicit antigen-specific B cell responses (Li at al., 2013). The genes for HLA-DR beta 1 (DRB1) and HLA-DQ beta 1 (DQB1) have been associated with immune dysregulation or autoimmunity (NLM, 2016; NLM, 2016b). There have been no studies, to date, which directly address the potential contribution of HLA phenotype with vaccine outcome in response to vaccination with either Ft or VEEV.

Study Purpose

The purpose of this dissertation project was to quantitatively assess the molecular epidemiology of Ft and VEEV in a human vaccine model. I worked to quantitatively assess the potential relationship between gene expression over time in response to treatment (vaccination) with either Ft or VEEV; to assess the potential relationship between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination or vaccine failure); and to identify a potential correlation between HLA phenotype and vaccine response. I conducted this research to add to the limited published information about the changes in gene expression that each vaccine elicits, and

to address the gaps in knowledge regarding the association of age and genetic makeup on vaccine outcome.

Research Questions and Hypotheses

Research Question 1

Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft or VEEV? To answer this question I used a quantitative analysis utilizing three-way analysis of variance (ANOVA), where the dependent variable was level of gene expression for a given transcript and the independent variables which affect the dependent variable were time following vaccination, treatment status (which vaccine was administered or if the subject was a mock-vaccinated control), and the combined effect of time*treatment which I used to consider the role of both time and treatment as a single variable. The null hypothesis was that there is no association between changes in the level of gene transcription in response to vaccination with Ft or VEEV, and the alternative hypothesis was that there is association between changes in the level of gene transcription and vaccination with Ft or VEEV.

Research Question 2

Is there a significant association between changes in the level of gene transcription over the time course of vaccination with Ft in comparison to VEEV? The null hypothesis was that there is no association between changes in the level of gene transcription in response to vaccination with Ft in comparison to VEEV, and the alternative hypothesis was that there is association between changes in the level of gene

transcription and vaccination with Ft in comparison to VEEV. To answer this question, I used a quantitative analysis utilizing ANOVA, where the dependent variable was level of gene expression for a given transcript and the independent variables which affect the dependent variable were time following vaccination, treatment status (which vaccine was administered or if the subject was a mock-vaccinated control), and the combined effect of time*treatment which I used to consider the role of both time and treatment as a single variable.

Research Question 3

Is there a correlation between age of the vaccinated subject and vaccination outcome? This question can be addressed by conducting a correlation analysis. The null hypothesis was that there is no correlation between age of the vaccinated individuals and vaccine outcome (measured as the log10 neutralizing antibody titer), whereas the alternative hypothesis was that there is a statistical correlation between age and vaccine outcome.

Research Question 4

Is there a significant association between HLA phenotype and vaccination outcome? I answered this by employing Mann-Whitney U tests to examine the association of immune response with HLA-DRB1 or HLA-DQB1alleles. The null hypothesis was that there is no association between HLA allele and vaccine outcomes (measured as the log10 neutralizing antibody titer), whereas the alternative hypothesis was that there is a statistical association between HLA allele and vaccine outcome.

Conceptual Framework

I based the conceptual framework for this study upon theories of rational vaccine design which have emerged in recent years. Historically, vaccine design has relied upon two tried and true methodologies – that of attenuation or inactivation of the biological agent (Plotkin, 2003; Plotkin, 2008). New methods of generating vaccines include efforts to produce vectored vaccines (i.e., vaccines that are delivered in a "shell" virus such as an adenoviral vector), and DNA-based vaccines, but no single ideal strategy has emerged (Plotkin, 2003). Genomic-based efforts provide an attractive alternative for modern vaccine design because they can provide mechanistic information that was previously unattainable (Seib, et al., 2009). Seib et al. (2009) discussed several ways in which genomics can contribute to efforts to overcome emerging infectious diseases, including the use of immunogenetics to understand the host immune response to infection, and cited the use of such information to understand the host responses to mumps virus vaccine. Indeed, Ovsyannikova et al. (2008) found that variation in response to mumps vaccine was associated with both single nucleotide polymorphisms (SNPs) in several cytokine genes and with specific human leukocyte antigen (HLA) phenotypes. Hoft, Brusic, and Sakala (2011) reviewed the use of immunogenetics information to predict differential development of specific T helper subsets, as well as the use of inhibitors such as rapamycin and metformin which can influence the generation of long-term memory T cells following vaccination. Hoft et al. (2011) also discussed their unpublished preliminary human data from their study on the molecular responses to TB vaccine. Immunogenetics can be used to understand host cell proteins and signaling processes

which are involved in both the pathogenesis and development of immunity against a specific biological agent (Seib et al., 2009). Additionally, such studies may also identify molecular correlates of protection (Pulendran and Ahmed, 2011). Understanding the underlying mechanisms of disease process and development of immunity can facilitate our understanding of the molecular inner armor that functions to protect that human host against infection (Nara et al., 2010). I have termed this collection of concepts the Epidemiology Nexus concept. I applied the notion of a nexus of information to connect the impact of disease upon individuals, families, and communities with the potential impact that such information could exert when applied to policy issues regarding vaccine design.

Nature of the Study

My basic goal was to study the association of transcriptional responses of two disease-causing agents with the temporal effects of treatment, and additionally to examine the correlation or association of demographic factors with vaccine outcome. By examining and comparing the transcriptional changes, genes and/or pathways that are important to both types of infection as well as genes and pathways that are unique to each type of infection can be identified. These data can contribute to a foundation of knowledge of the key molecular mediators and events that lead to the development of an appropriate immune response as well as the development of immune status through the identification of molecular correlates of protection. I leveraged existing data sets of human microarray data, and used two human vaccine study data sets (*in vivo*; one each for Ft and VEEV).

The key study variables I used to test association of vaccination with changes in the level of transcription included transcript expression level as the dependent or outcome variable, with independent variables such as time following vaccination and treatment. For the association of HLA phenotype or allele with vaccine outcome, the dependent variable was neutralizing antibody titer, and the independent variable was HLA allele.

Given that this project was a secondary analysis, I used previously collected data. The data sets were available from the US Army Medical Research Institute of Infectious Diseases (USAMRIID) as electronic data comprised of fluorescence expression signals in the form of a pixelated image; the expression data are averaged to give a single intensity value for each probe set on each microarray chip; each Affymetrix Human Genome U133 Plus 2.0 Array chip contains more than 54,000 probe sets which correspond to approximately 38,500 human genes (Affymetrix, 2004). The original vaccine studies conducted by researchers at USAMRIID included 39 participants between the two studies. The treatment groups were separated into 4 sub-groups which were comprised of 10 control individuals who were given a placebo vaccination for Ft, 10 individuals who received the Ft vaccine, 10 control individuals who were given a placebo vaccination for VEEV, and 9 individuals who received the VEEV vaccine. Control and vaccinee volunteers were age-matched (+/- 5 years) and all study participants were males. Blood samples were collected at various time points before (0 hr) and after vaccination (or mock-vaccination) at 1, 4, 8, 24 hrs, as well as through days 2, 7, 14, 21, and 28. One study volunteer for the VEEV vaccine portion of the study was removed from the data set

due to primary vaccine failure (the volunteer failed to develop neutralizing antibody titer after vaccination and was thus reclassified as a vaccine non-responder).

For generation of transcript expression profiles, each whole blood sample was used to isolate RNA which was subsequently hybridized to an Affymetrix Human Genome U133 Plus 2.0 Array chip (Affymetrix, 2004). Power calculations for the original studies (Statistical Applications Software [SAS] Institute Inc., 2011) were performed by a USAMRIID statistician to estimate that using 10 individuals per group should typically yield 95% power to detect 3.5-fold differences in gene expression using a two-sample *t* test at the 0.001 two-sided significance level, assuming that the coefficient of variation is 0.50. To maximize the number of transcripts that are examined, I decreased the cut off for fold expression to 2-fold up or down to broaden the number of transcripts examined. I applied this change and recalculated the power that could be achieved in the secondary analysis which resulted in anestimated reduction in the power to 90% (M. D. Anderson, 2016).

Definitions

Gene expression: "The process by which a gene's sequence is converted into a mature gene product or products (proteins or RNA)" (GO, 2012). Gene expression includes "the production of an RNA transcript as well as any processing to produce a mature RNA product or an mRNA (for protein-coding genes)" (GO, 2012). Gene expression and transcript expression are terms which are often used synonymously.

Probe set: A set of synthetic oligonucleotide probes, typically 20 nucleotides in length, which are used to interrogate gene expression. The synthetic probes make up a set

that has been validated to recognize a specific gene of interest. A single gene of interest may be represented by multiple probe sets which can bind to different portions of the transcript RNA (Affymetrix, 2004).

Microarray analysis: A technique used to measure changes in gene expression. Gene expression for any given transcript is the base 2 logarithm of the fluorescent intensity signal from a microarray chip; the RNA from each individual subject at each time point was hybridized to a single gene chip. A base 2 logarithm of the fluorescent intensity signal expression value is generated for every transcript at each experimental time point (Partek, 2016). The temporal trend of gene expression is assessed by analysis of variance (ANOVA), minimally comparing transcripts over time, treatment, and in deference to the combined effects of Time*Treatment. Individual gene expression levels do follow linear trends (as opposed to exponential, for example).

Human leukocyte antigen (HLA): This system, as part of the Major Histocompatibility Complex, is comprised of four genes (A, B, C, and D) encoded on human chromosome 6. The HLA genes encode specialized cell surface molecules which present antigenic peptides to the T cell receptor (TCR) on T cells in order to initiate a cell-based immune response. The HLA phenotype of an individual, particularly for HLA-DR, HLA-B and HLA-A, is a critical determinant for tissue typing to match organ donors and recipients. However, there are also a number of immune and autoimmune diseases which have been associated with specific HLA alleles (Delves, 2016).

Assumptions

I made some assumptions that may have affected the study. First, I assumed in the analysis of microarray data that the number of genes being assessed was large enough (typically more than 10,000) and that only a small number of genes would actually have differences in gene expression when comparing different points (either across time or treatment) (Wang & Xi, 2013). Although the chip that each of the data sets I used for the secondary analysis contained more than 54,000 RNA transcripts, only several hundred or thousand will exhibit differential expression. This was an important assumption which provided a basis for the data to be normalized by a process called robust multi-array analysis (RMA). A second assumption, related to the overall goal of the analysis, was that my evaluation of the data would provide information about the molecular epidemiology of the cells from which the RNA originated. These assumptions were critical to both the process by which the data from two independent data sets was combined and normalized, and for the implications for molecular epidemiology in cataloging the changes to gene transcription in response to vaccination.

Scope and Delimitations.

The researchers who conducted the two original vaccine studies limited the scope of the study temporally to immediately prior to vaccination and ending at day 28 following vaccination. The researcher also limited the volunteers to those who had not previously received any vaccines within 30 days of the start of the Ft or VEEV vaccination protocol, previous vaccination to either Ft, VEEV, or to other, related alphavirus vaccines. Volunteers also had to be negative for infection with human

immunodeficiency virus (HIV), hepatitis B, and hepatitis C. Volunteers were excluded if a review of their medical history included abnormal findings in immune system function, evidence of anemia, or blood cell populations outside of normal ranges. The delimitations of the study included restricting participants to males of an adult age range (21-48 years of age). The researchers of the original studies placed no limitations on geographic origin of the participant pool; all volunteers were drawn from within the scope of active duty members of the US military.

The above-mentioned factors may impact the overall generalizability of the secondary analyses, particularly in light of the fact that the study population did not include women.

Limitations

The secondary analyses of the two data sets I conducted for this study were limited by the design and quality of each of the original studies. In particular, it should be noted that each original data set did not contain any female study volunteers. In addition, only 10 volunteers were included in each treatment group (e.g., control and vaccinee). Sample size was likely not a limitation for the research questions related to association of transcript expression with treatment. My calculations for power and sample size demonstrated that 10 volunteers in each group would yield 90% power to detect 2 fold changes in gene expression, so in terms of the statistical analysis there was no limitation. Rather, the potential limitation may stem from a perception of just how generalizable the data are. The study volunteers were overwhelmingly of Caucasian, not Hispanic or Latino race and ethnicity; therefore, the data were also limited in that there can be no correlation

of vaccine response with any racial or ethnic group.

The live attenuated strain of Ft that was used in the original study may not produce the same host responses to infection that a wild-type pathogenic strain of bacteria may elicit. The live attenuated strain of VEEV that was used to vaccinate volunteers belongs to the IA/B strain of VEEV; the changes in gene expression that are associated with this particular strain may or may not be the same as what would be observed if using a virus from one of the other VEEV subtypes or with fully virulent strains.

The researchers implemented measures during the collection of the original data to control bias, including blinding of sample collection personnel to the treatment groups, and collection of all samples within a pre-defined range of acceptable collection (e.g., 4 hours post-vaccination ± 30 minutes, day 2 post-vaccination ± 2 hours, or day 21 post-vaccination ± 2 days). Some types of bias could not be controlled in the study, including attrition bias (study participants had the right to terminate participation in the study at any time) and exclusion bias—in particular the exclusion of women from the study in order to control variation in transcript expression that could be due to fluctuations in female hormone signaling.

Significance

Global public health efforts to prevent disease from both Ft and VEEV have been hampered by the lack of approved human vaccines and the lack of knowledge of the human molecular responses to infectious disease agents (Foley & Nieto, 2010; Wolfe et al., 2013; Wolfe et al., 2014). Molecular epidemiology studies of other infectious

diseases have provided innovative information in humans and animal models; however, there is little understanding or information of the molecular consequences of human infection with either Ft or VEEV. In this project, I sought to make a novel contribution to the field of public health by studying the relationship between gene expression and the temporal effects of vaccination, the potential relationship between age and vaccine outcome, and the potential association between genetic make-up of individuals with vaccine outcome.

My study may contribute to positive social change at multiple levels. First, transcriptomic analysis may provide evidence to support the Epidemiology nexus conceptual model which holds that examining changes in gene expression can contribute to a broad knowledge base for understanding the molecular mechanisms of immune development and expand knowledge of the fundamental mechanisms of action of successful vaccination (Li et al., 2015; Kennedy et al., 2014; Nakaya, Li, & Pulendran, 2011; Pulendran et al., 2013). Secondly, my use of molecular epidemiology data for the explicit purpose of creating rationally-designed vaccines supports a positive social change agenda relating to administrative policy; financial incentives are often absent with regard to development of neglected tropical or zoonotic diseases and the World Health Organization (WHO) recognizes that developmental costs for vaccines and lack of correlates of protection are unmet challenges (Hortez et al., 2016). Studying the common and unique changes to gene expression that are elicited by vaccination may make it possible to develop multi-agent vaccines that share similar immune-related mechanisms. By using Ft and VEEV to compare and contrast molecular epidemiology, I worked to

contribute new knowledge to develop vaccines which can meet rigorous standards for licensure. Licensed vaccines could have a positive social impact on society on both and individual and family levels through reduced disease incidence, and impact on communities through reduced mortality and morbidity (Hortez et al., 2016; Reichert et al., 2009, Wolfe et al., 2013, Wolfe et al., 2014). The compendium of evidence I have built regarding the mechanisms of vaccine efficacy may also serve to assure the public that vaccines are well-tested and safe. Such assurance has become increasingly important as public health professionals seek to understand and combat vaccine hesitancy and to promote vaccine acceptance (Eskola, Duclos, Schuster, MacDonald, & the SAGE Working Group on Vaccine Hesitancy, 2015).

Summary

Global public health efforts to prevent disease from both Ft and VEEV have been hampered by the lack of approved human vaccines and the lack of knowledge of the human molecular responses to infectious disease agents (Foley & Nieto, 2010; Wolfe et al., 2013; Wolfe et al., 2014). Molecular epidemiology studies in other infectious diseases have provided innovative information in humans and animal models; however, there is little understanding or information of the molecular consequences of human infection with either Ft or VEEV. In this dissertation project, I sought to make a novel contribution to the field of public health by studying the relationship between gene expression and the temporal effects of vaccination, the potential relationship between age and vaccine outcome, and the potential association between genetic make-up of individuals with vaccine outcome.

In the next chapter, I review the relevant published literature for each infectious agent to provide readers with a comprehensive understanding of the subject matter, and to identify the gaps in the public health knowledge regarding each infectious agent.

Chapter 2: Literature Review

Problem Addressed by the Study

Emerging and re-emerging diseases such as tularemia, dengue fever, Ebola, equine encephalomyelitis, and Zika highlight gaps in the development and approval of new vaccines as well as the need to incorporate innovative approaches to address global health challenges to produce new vaccines (Hotez et al., 2016; Kennedy et al., 2014; Wilson & Karp, 2015). Vaccine subject matter experts believe that knowledge of factors such the host innate and adaptive responses to vaccination, differences in genetic makeup, gender and age-related differences, and differences among diverse ethnic populations may provide evidence-based data needed to facilitate development strategies to overcome technological limitations, including identifying novel correlates of protection (Angel, Steele, & Franco, 2014; Hotez et al., 2016; Kennedy et al., 2014; Pulendran et al., 2013). The lack of knowledge (a) surrounding the pathogenesis of disease in humans, (b) knowledge of vaccine outcome among diverse populations, and (c) knowledge of the effect of genetic makeup in relation to vaccine outcome have emerged as critical barriers to producing effective vaccines, particularly against emerging and re-emerging infectious diseases (Klein & Poland, 2013).

Purpose of the Study

The purpose of this dissertation project was to quantitatively assess the potential relationship between gene expression over time in response to treatment (vaccination) with either Ft or VEEV; to assess the potential relationship between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination

or vaccine failure), and to identify the potential correlation between HLA phenotype and vaccine response. I conducted the study to add to the limited published information about the changes in gene expression that each vaccine elicits, and to address the gaps in knowledge regarding the association of age and genetic makeup on vaccine outcome.

Literature Relevant to the Problem

Global public health efforts to prevent disease from both Ft and VEEV have been hampered by the lack of approved human vaccines and the lack of knowledge of the human molecular responses to infectious disease agents (Foley & Nieto, 2010; Wolfe et al., 2013; Wolfe et al., 2014). Worldwide the annual incidence rates of infection with Ft and VEEV vary from year to year, but are often thought to be underdiagnosed, particularly for VEEV, due to poor diagnostic capabilities in remote areas and the general symptoms of disease which are often misdiagnosed as other common diseases, such as dengue fever (Aguilar et al., 2011; NNDSS, 2016; HealthMap, 2016). There have been few studies to date which have examined human immune or molecular responses to infection with either Ft or VEEV. Changes in gene expression have been studied by microarray analysis following both naturally occurring infections as well as in purified blood cells populations following vaccination (Andersson et al., 2006; Erwin-Cohen et al., 2012; Fuller et al., 2007; Paranavitana et al., 2008b). However, researchers to date have not assessed changes in gene expression from whole blood as a sample source; the use of whole blood as source material has been postulated to be important to gene expression analysis because it would leave *in vivo* cell to cell signaling mechanisms intact rather than studying one type of blood cell population (Joehanes et al., 2012).

In this chapter, I review relevant literature related to the conceptual framework that guided my formulation of the research questions, literature that provides examples of how the Epidemiology Nexus concept has been applied previously to research efforts, and literature relating to how the framework can benefit the my dissertation research.

Literature Search Strategy

I executed multiple literature searches, and modified search strategies each time. The first searches were relatively narrow in terms of both years searched and the generality of search terms. The primary literature search for VEEV encompassed years ranging from 2010 to 2016, using the search terms *Venezuelan equine encephalitis, VEE, epidemiology, effects, gene expression profiling, or biomarker* separately and in conjunction. I conducted this search using the National Center for Biotechnology Information (NCBI) database, also known as PubMed. I filtered the search results by restricting them to human data. The same search strategy was subsequently used to conduct searches using the same search terms, but I expanded the scope of literature from years 2005 to 2016 in order to decrease stringency, capture more published articles, and maintain updated literature during the process of literature review synthesis. In the final search, I further expanded the range of references to include articles published between the years 2000 and 2016. Throughout the process, I found that there was less human-relevant literature from which to draw for VEEV, in comparison to literature available related to Ft. Many recent VEEV studies reported results from studies of animal models, but the literature I chose for the review was restricted to only human data.

I also conducted several literature searches for Ft. The initial literature search was a relatively narrow search utilizing the terms *Francisella tularensis, epidemiology, gene expression profiling, epidemiology, or biomarker,* separately and in conjunction and filtering results for those applicable to humans only. I limited the initial search of the PubMed database to articles published between 2010 and 2016. The search strategy was subsequently modified to expand the scope of the search to include publications spanning 2005-2016, but restricted to articles only relevant to human infection and published in English. I repeated literature searches periodically to ensure that up-to-date information was incorporated into the literature review. Many recent publications reported results from studies of animal models rather than reports of human studies or studies using human cell lines.

Throughout my searches, I found a profound lack of literature published within the last 5 years on human disease with Ft or VEEV, particularly with regard to human responses to infection. Published reports extending beyond the past 5 years were deemed to be highly relevant to the dissertation study.

Conceptual Framework

Theoretical Concepts

The theoretical concept that informed my recognition of the public health problem and formulation of quantitative research questions is that a comprehensive understanding of the human immune responses to vaccination may provide key insight into the molecular epidemiology of infectious diseases. Though many vaccines have been established following traditional paradigms of development, some diseases have proven

to be more difficult to address due to the nature of the disease and the lack of natural immunity (Nabel, 2013). Many emerging and re-emerging diseases such as tularemia, dengue fever, Ebola, equine encephalomyelitis, and Zika highlight gaps in the development and approval of new vaccines as well as the need to incorporate innovative approaches to address global health challenges to produce new vaccines (Hotez et al., 2016; Kennedy et al., 2014; Wilson & Karp, 2015). Many subject matter experts believe that knowledge of additional factors such the host innate and adaptive responses to vaccination, differences in genetic makeup, gender and age-related differences, and differences among diverse ethnic populations may provide evidence-based data needed to facilitate development strategies to overcome technological limitations, including identifying novel correlates of protection (Angel, Steele, & Franco, 2014; Kennedy et al., 2014; Hotez et al., 2016; Pulendran, Oh, Nakaya, Ravindran, & Kazmin, 2013). The lack of (a) knowledge surrounding the pathogenesis of disease in humans, (b) knowledge of vaccine outcome among diverse populations, and (c) knowledge of the effect of genetic makeup in relation to vaccine outcome have emerged as critical barriers to producing effective vaccines, particularly against emerging and re-emerging infectious diseases (Klein & Poland, 2013). For diseases in which it has been difficult or impossible to demonstrate natural immunity, a comprehensive understanding of immunopathogenesis is necessary to inform vaccine design (Nabel, 2013).

I leveraged one aspect of a systems approach, the study of transcriptomics, to understand the connection between the complex molecular epidemiology of human responses to pathogen-derived disease and the mechanisms of action of successful

vaccination (Nakaya, Li, & Pulendran, 2011).

I formulated the Epidemiology Nexus concept from leading subject matter experts in the field of systems vaccinology. This ides has influenced my dissertation research in that associating changes to the transcriptome that occur as a result of the host response to infection with bacterial or viral infectious agents may provide a foundation of knowledge for multiple disease paradigms which may be leveraged to influence combinatorial vaccine design, therapeutic targets, and diagnostic markers. The research actions which facilitate the acquisition of epidemiological knowledge include identification of genes (transcripts) which are modulated, and identification of the known signaling pathways which are significantly associated with the transduction of a biological signal (Kellam, 2006; Kellam & Weiss, 2006; Nakaya et al., 2011). Collectively, these data provide a portrait of the host response to the vaccine (which is used as a surrogate for virulent infection) from which models of mechanisms of action, biological signatures, and predictive responses can be built (Nakaya et al., 2011). The overall philosophical concept of an Epidemiology Nexus has been described in consistent terms by several prominent figures in vaccinology including Angel, Steele, and Franco (2014), Kennedy et al. (2014), Li et al. (2013), Nabel (2013), Plotkin (2003), Pulendran and Ahmed (2011), and Pulendran et al. (2013).

Key statements and definitions of the framework. Systems vaccinology has been defined as a field of study which combines system-wide measures, signaling pathways and networks, and contributes to predictive models within the context of vaccines (Nakaya et al., 2011). Transcriptomics is defined as the complete set of RNA

transcripts which are produced under specific circumstances or in a specific cell type and are assessed techniques such as microarray analysis (Affymetrix, 2004).

Application in previous research. Several studies have applied theoretical concepts that are consistent with an Epidemiology Nexus connecting disease with alterations in the transcriptome in a manner similar to what I did for this dissertation research project.

There have been three independent studies which have applied the theoretical concept to the yellow fever vaccine in a manner similar to my approach. Scherer et al. (2007) first investigated the association between disease and gene expression profiles of peripheral blood mononuclear cells (PBMCs) from patients who were vaccinated with either smallpox, yellow fever, or from patients with naturally occurring upper-respiratory infection (URI). Healthy, naïve volunteers were recruited to receive either the smallpox vaccine Wetvax (*n*=24), or the yellow fever vaccine YF-VAX (*n*=20; Scherer et al., 2007). The third experimental group was comprised of patients who presented with clinical symptoms including "fever, cough, chills, myalgia, pharyngitis, and rhinorrhea" and were presumed to have viral (URI; Scherer et al., 2007, p. 6). Each patient with presumed URI had experienced symptoms for 3.7 days on average (Scherer et al., 2007).

The timeframes for sample collection were not uniform; PBMCs from the volunteers vaccinated with Wetvax (smallpox) were collected at four general points corresponding to a pre-vaccination sample, a sample at 2-4 days post-vaccination, a sample at 5-7 days post-vaccination, and a convalescent sample at days 50-60 post-vaccination (Scherer et al., 2007). For volunteers vaccinated with YF-VAX, the samples

were collected only at pre-vaccination and at days 5-7 post-vaccination (Scherer et al., 2007). Finally, the samples collected for the URI group were collected at the initial visit, presumed to be approximately 3-7 days post-infection and then again four weeks later which corresponded to approximately 35 days post-infection (Scherer et al., 2007). Association of gene expression with disease was assessed by hybridization of samples to cDNA microarrays; smallpox samples were hybridized to Agilent human 1 cDNA microarrays and the yellow fever and URI samples were hybridized to Agilent human 1A oligo microarrays (Scherer et al., 2007).

Gaucher et al. (2008) and Querec et al. (2009) both employed transcriptional profiling to assess gene expression signatures in order to predict vaccine outcome and innate immune responses in response to vaccination with yellow fever virus vaccine. Gaucher et al. isolated RNA from whole, unfractionated blood from vaccinated volunteers (n=15) both prior to vaccination (day 0) and at several time points following vaccination (days 3, 7, 10, 14, 28, and 60). RNA samples were hybridized to Illumina Human RefSeq-8 BeadChips v2 to assess the transcriptional levels of more than 24,000 unique probe sets (Gaucher et al., 2008). The researchers made Comparisons of statistically significant changes in gene transcription between the post-vaccination samples (days 3, 7, 10, 14, 28, and 60) and the pre-vaccination sample (day 0) using BioConductor software and the linear models for microarray analysis (LIMMA) protocol (Gaucher et al., 2008). Gaucher et al. discovered that the majority of changes to gene transcription occurred on days 3 and 7 post-vaccination. Several important functional categories were suggested to be modulated by vaccination including the interferon

pathway, complement systems, and dendritic cell-associated genes; natural killer cell-associated genes were also up-regulated (Gaucher et al., 2008). Among the genes that were discovered to be down-regulated were those that belong to ribosomal protein families.

The association of gene expression with disease from the study by Scherer et al. (2007) included 28 genes that overlapped between all three disease states (smallpox, yellow fever, or URI) and included several interferon-response genes (IFI27, IFI30, IFIT1, IFIT2, IFIT3, IFITM1, IFITM2, ISG15, ISGF3G) and genes involved in viral innate immunity (MX1, MX2), as well as genes associated with proteasome and transporter functions (PSMA4, PSMB9, PSME1, PSME2, TAP1). The study conducted by Querec et al. (2009) also assessed changes in gene expression in PBMCs of vaccinated individuals and reported transcript expression alteration in response to vaccination of genes belonging to the interferon family including ISRE, IRF7, SREBF1; genes associated with antiviral responses (OAS1, OAS2, OAS3, OASL); genes involved in viral recognition including TLR7, DDX58, IFIH1, DHX58, and EIF2AK; and genes involved in the mediation of viral immunity (CXCL10 and MX1). Many of the genes reported by Querec et al. (2009) overlapped with those reported by Scherer et al. (2007) and indicated a generalized immune response to viral infection that depended on early signals through interferon pathways including signals which indicated that early activation of the inflammasome is a key common response to viral presentation, regardless of the viral species (smallpox, yellow fever, or suspected viral URI). The experimental approach executed by Gaucher et al. (2008) was slightly different in

comparison to those of Scherer et al. (2007) or Querec et al. (2009) in that instead of assessing transcription in a purified cell population (i.e., PBMCs), the researches monitored changes in expression from whole, unfractionated blood. Importantly, the timeframes of sample collection overlapped similarly between all three studies with a pre-vaccination sample, and then post-vaccination samples collected at days 3, 7, 10, 14, 28, and 60 (Gaucher et al., 2008).

Gaucher et al. (2008) found most of the changes in gene expression occurred at days 3-7 post-vaccination and that composition of the aggregate changes included genes of the toll-like receptor pathway (TLR7, MYD88, and IRF7), the interferon stimulated pathway (IFI27, IFI30, and OAS-1), genes involved in antigen processing and transport (TAP1 and TAP2) and the complement system (C1QA, C1QB, C3AR1, and SERPING1). Overall, assessing molecular epidemiology through the analysis of changes in gene transcription produced remarkably similar responses in three independent studies using the yellow fever vaccine as a surrogate for endemic infection.

The approach has continued to be relevant as shown by the recent study by Baddal et al. (2015) who used RNA-seq technology to assess transcriptional changes induced by infection with nontypeable *Haemolphilus influenza*. In their study, Baddal et al. (2015) well-differentiated, normal primary human bronchial epithelial cells and infected the apical side of the cultured cells with nontypeable Haemophilus influenza strain Hi176. Cells were harvested at 1, 6, 24, and 72 hours post-infection (in three independent experiments) were used to isolate RNA which was subjected to RNA-seq analysis (Baddal, et al., 2015). Results reported by Baddal et al. (2015) included a rapid

alteration of host cell junctional complexes, including decreased transcription of PCDH8, CLDN3, CLDN8, ADD3γ, and ARPC2. Analysis of functional groups revealed significant differential expression at 1-6 hours post-infection of genes involved in hypoxia response, regulation of cell death, epithelial cell differentiation, anchoring, and adhering junctions (Baddal, et al., 2015). In contrast, the predominant functional groups observed at the later 24-72 hour time points indicated involvement of immune, defense, and stress responses; wound response; inflammatory response; and cytoskeletal and matrix gene alterations (Baddal et al., 2015).

The conceptual framework can be extended to include information relating to the utility of exploring potential relationships between factors such as age or genetic makeup with vaccine outcome. HLA are encoded by the MHC genes (Li et al., 2013). HLA class II molecules are responsible for presentation of antigens to ultimately elicit antigen-specific immune-cell responses (Li at al., 2013). Recent work with West Nile virus neuroinvasive disease identified specific gene variants associated with seroconversion and HLA alleles that were associated with immunoprotection (Long et al., 2016). Several studies have reported the association of specific genetic factors with vaccine outcome in response to vaccination with childhood vaccines [HBV, diptheria, tetanus, pertussis (DTaP)], smallpox, influenza, anthrax, and rubella (Li et al., 2013; Narwaney et al., 2013; Ovsyannikova, Jacobson, Ryan, Dhiman, Vierkant, & Poland, 2007; Ovsyannikova et al., 2013; Ovsyannikova, Pankratz, Salk, Kennedy, & Poland, 2014; Yucesoy et al., 2013).

How the Current Study May Benefit From the Epidemiology Nexus Concept

I based the Epidemiology Nexus concept upon theories of rational vaccine design

which have been described in recent years; particularly the need to understand the molecular epidemiology in order to influence rational vaccine designs. Historically, vaccine design has relied upon two tried and true methodologies – that of either attenuation or inactivation of the biological agent (Plotkin, 2003). New methods of generating vaccines include efforts to produce vectored vaccines (i.e., vaccines that are delivered in a "shell" virus, such as an adenoviral vector) and DNA-based vaccines, but no single ideal strategy has emerged (Plotkin, 2003). Genomic-based efforts provide an attractive alternative for modern vaccine design as they can provide mechanistic information that was previously unattainable (Seib et al., 2009). Seib et al. (2009) discussed several ways in which genomics can contribute to efforts to overcome emerging infectious diseases, including the use of immunogenetics to understand the host immune response to infection, and cited the use of such information to understand the host responses to mumps virus vaccine. Indeed, Ovsyannikova et al. (2008) found that variation in response to mumps vaccine was associated with both single nucleotide polymorphisms (SNPs) in several cytokine genes as well as associated with specific human leukocyte antigen (HLA) phenotypes. Hoft et al., (2011) reviewed the use of immunogenetics information to predict differential development of specific T helper subsets as well as the use of inhibitors such as rapamycin and metformin which can influence the generation of long-term memory T cells following vaccination. Hoft et al. (2011) also discussed their unpublished preliminary human data in which the molecular responses to TB vaccine were studied. Molecular epidemiology can be used to understand host cell proteins and signaling processes which are involved in both the

pathogenesis and development of immunity against a specific biological agent (Seib et al., 2009). Additionally, such studies may also identify novel molecular correlates of protection (Pulendran & Ahmed, 2011).

Global public health efforts to prevent disease from both Ft and VEEV have been hampered by the lack of approved human vaccines and the lack of knowledge of the human molecular responses to infectious disease agents (Foley & Nieto, 2010; Wolfe et al., 2013; Wolfe et al., 2014). Molecular epidemiology studies in other infectious diseases have provided innovative information in humans and animal models; however, there is little understanding or information of the molecular consequences of human infection with either Ft or VEEV. This dissertation project seeks to make a novel contribution to the field of public health by studying the relationship between gene expression and the temporal effects of vaccination, the potential relationship between age and vaccine outcome, and the potential association between genetic makeup of individuals with vaccine outcome.

Literature Review Related to Key Variables and Concepts

Previous Transcriptomic Studies of Human Responses to Ft or VEEV: Relation to Key Study Variables

There have been few human studies to date which have addressed the state of transcription in humans following infection with either Ft or VEEV. Andersson et al. (2006) analyzed the transcriptional responses of seven individuals who presented at a university hospital with naturally occurring Ft infection in central Sweden. The patient age ranged from 50-76 years of age, with an average age of 60.7 years old and were

approximately equal for male and female (four males, three females) (Andersson et al., 2006). All patients were administered antibiotic treatment either at the first hospital visit or upon follow-up (approximately 6 days after the initial visit) (Andersson et al., 2006). Transcriptional changes were examined in RNA isolated from peripheral blood mononuclear cells of the seven patients, collected at various points beginning with the initial visit to the hospital (presumed to be roughly 2-3 days following the initial infection event), at days 6-7, 8-9, 10-11, and 13 following infection, and finally at about 3-4 months following presumed initial infection to provide a convalescent sample (Andersson et al., 2006). Using Affymetrix HG-U133A microarrays, levels of gene transcription from patient samples were compared to the levels of transcription in RNA isolated from PBMCs of eight healthy adult donors who were demonstrated to be naïve for Ft infection; the RNA from the control donors was pooled, analyzed, and used for comparison as a single reference control (Andersson et al., 2006). The researchers in that study identified seven transcripts as potential biomarkers of early infection with Ft to include the genes STAT1, SECTM1, TNFAIP6, TNFSF10, CD3E, MSRB2, and IL2RB (Andersson et al., 2006). There were several limitations of the Andersson study. The number of infected study subjects (n= 7) presented a limitation to statistical analysis and there was variation in the number of samples collected from each volunteer at each time point (Andersson et al., 2006). Ft samples were compared with RNA isolated from eight uninfected control subjects, and the control RNA samples were subsequently pooled together; there was no mention of the age of the control donors and the samples were not time-matched to the infected samples (Andersson et al., 2006). Finally, each of the Ft study subjects was

treated with antibiotics either upon the first presentation at the university hospital or at the follow-up visit one week later; the transcriptional results obtained may not be authentically representative of Ft infection due to the antibiotic treatment which was administered (Andersson et al., 2006).

The effects of Ft on gene transcription were examined *in vitro* at very early times post -infection by Paranavitana, Pittman, Velauthapillai, Zelazowska, and DaSilva (2008b). Paranavitana et al. (2008b) isolated peripheral blood mononuclear cells (PBMCs) from eight healthy human volunteers with no previous Ft infection. PBMC cells were subsequently exposed to the live-attenuated vaccine strain of Ft (LVS) *in vitro*; cells were harvested and RNA isolated at 1, 4, 8, 16, and 24 hours post infection (Paranavitana et al., 2008b). The researchers confirmed the *in vitro* study results for four of the early disease biomarkers of the Andersson et al. (2006) study, including TNFAIP6, STAT1, TNFSF10, and SECTM1 and additionally observed the genes to be expressed in a temporal manner (Paranavitana et al., 2008b). Paranavitana et al. (2008b) also describe changes in novel pathways, such as changes in genes associated with tryptophan metabolism, lipid metabolism, and genes involved in activation of the inflammasome. Infection was also noted to induce a complex pattern of pro-inflammatory cytokines as well as anti-inflammatory cytokine responses, including IL-1Ra, IL2, GCSF, VEGF, IL-17 and IL-22 (Paranavitana, Pittman, Velauthapillai, & DaSilva, 2008a; Paranavitana, Zelazowska, DaSilva, Pittman, & Nikolich, 2010). Some of the limitations associated with the *in vitro* study include a restriction to only very early transcriptional events (the time course of the exposures ranged from 1 hour to 24 hours post infection) and make

comparison to the Andersson et al. (2006) study difficult since the earliest estimated time point for that study was 2 to 3 days following presumed initiation of infection.

In another study, the molecular epidemiology events following vaccination with the LVS Ft vaccine were examined from the PBMCs of healthy naïve donors who were subsequently vaccinated (Fuller et al., 2006; Fuller et al., 2007). The PBMCs of six adults (male and female) ranging in age from 22-54 years of age were collected at five time points: 6 days prior to vaccination, 18 hours post-vaccination, and then at days 2, 8, and 14 post-vaccination to correspond with early (18 hours and day 2 post-vaccination) and late (days 8 and 14 post-vaccination) stages of infection (Fuller et al., 2006; Fuller et al., 2007). Only samples from five of the six vaccinees were used in the analysis (Fuller et al., 2007). RNA isolated from the PBMCs was hybridized to Affymetrix HG U133 Plus 2.0 microarray chips (Fuller et al., 2007). The data sampled from individual volunteers were compared and contrasted temporally between pre-vaccination and post-vaccination; no unvaccinated control volunteers were used in this study (Fuller et al., 2006; Fuller et al., 2007). The predominant pattern responses to infection were revealed through cellular processes which included “immune-related, cell cycle-related, apoptosis-related, and biosynthesis/metabolism” related processes (Fuller et al., 2007, p. 3177). Fuller et al. (2007) examined temporal patterns of transcript expression and used similar patterns to categorize transcripts; two patterns showed interesting activity. The first pattern, where the temporal pattern of expression showed reduced expression early on contained many genes related to immunity including CD96, CCL5, PTPRCAP, TNFRS25, CD3D, ZAP70, KLF2, GSPT1, RHOH, PPP2R5C (Fuller et al., 2007). Down-regulation of the

expression of these transcripts was postulated by the researchers to represent an early negative regulation of pro-inflammatory immune responses by LVS Ft (Fuller et al., 2007). The second pattern in which the temporal pattern of expression indicated an early increased expression contained genes associated with activation of the innate immune response, including LAMP2, DCL-1, RNASE2, IL18, LY96, LILRA6, CD39, IL13RA1, CSAP2, and IRAK3 (Fuller et al, 2007). Fuller et al. (2007) distinguished their work from the similar studies conducted both by Andersson et al. (2006) and Paranavitana et al., (2008b) by examining and reporting specific responses that were linked to the activation of dendritic cells, which have been suggested as key modulators of long-term responses to Ft (Katz, Shang, Martin, Vogel, & Michalek, 2006). To that end, Fuller et al. (2007) reported the increased transcription and involvement of several MHC class I (B2M and TAP1) and MHC class II genes, and particularly toll-like receptors (TLR4, TLR5, and TLR8) and related genes (IRAK3, CARD8, and MYD88). These results were corroborated by a second human study conducted by Paranavitana et al. 2014, in which the recall responses to Ft were examined in human PBMCs isolated from individuals who had previously been vaccinated with Ft LVS and which then were subsequently re-stimulated (for 24 hours) *in vitro*. Transcripts related to dendritic cell maturation were among the most significant responses to the re-challenge (Paranavitana et al., 2014). The pathways that were predominantly affected by re-stimulation with Ft included dendritic cell maturation, TREM-1 signaling, cytotoxic T-cell mediated apoptosis, and IL-4 signaling (Paranavitana et al., 2014). Additionally, Paranavitana et al. (2010) demonstrated that the recall immune responses to Ft were driven by Th17 responses (IL-

17 and IL-22 cytokine secretion and signaling) in addition to the previously described responses observed involving Th1 cytokines IFN-γ and IL-2.

There have been two additional relevant studies which examined changes in gene transcript expression in response to biological threat agents other than Ft. One study was conducted as a proof of concept experiment that transcription expression patterns could be used as molecular signatures to differentiate between exposure to or infection with different biological threat agents that often presented clinically with similar flu-like symptoms (Das et al., 2008). Das et al., (2008) collected PBMCs from at least three independent donors, but did not report sex or ages for the donors who contributed to the VEEV portion of the study; the study population was merely described as male and female health volunteers between the ages of 19-61 years of age. The PBMCs were isolated and further separated into lymphocyte and monocyte subgroups by Ficoll gradient centrifugation, then re-mixed at a ratio of 1:4 monocytes to lymphocytes, which represented substantially enriched populations for both cellular subsets relative to normal levels (Das et al., 2008). PBMCs were then infected *in vitro* with biological threat agents including *Bacillus anthracis*, Venezuelan equine encephalitis virus, dengue fever virus, *Brucella melitensis*, *Yersinia pestis*, cholera toxin, staphylococcal enterotoxin B, or botulinum neurotoxin A at target doses that were appropriate for each agent to ensure optimal biological activity upon infection; infection was allowed to proceed for 30 minutes then cells were washed and further incubated for various periods of time, relevant to each pathogen (Das et al., 2008). Researchers isolated RNA samples at time points specific for each biological threat agent and subjected the samples to microarray

analysis on custom arrays with genes corresponding to Human Atlas 1.2 (Das et al., 2008). Changes in gene expression in the treated samples were compared by the researchers against the expression profiles of uninfected control samples, and gene transcripts that demonstrated variability in expression in the control samples were subtracted from the analysis leaving only stable baseline transcripts for comparison within the control groups (Das et al., 2008). The results of the expression analyses indicated a common transient induction of early response genes related to monocyte recruitment, whereas later time points demonstrated induction of genes consistent with DNA damage, hypoxia-inducible proteins, and proteases (Das et al., 2008). The results reported skewed heavily towards changes in expression for bacterial pathogens; no specific results for VEEV infection were noted (Das et al., 2008).

Finally, an *in vitro* study of PBMCs isolated from unvaccinated, previously VEEV-vaccinated, and VEEV-vaccination nonresponders was conducted to examine the similar and diverging molecular responses (Erwin-Cohen et al., 2012). The study population consisted all male volunteers ranging from 18-45 years of age: The control subjects were demonstrated to be naïve to any previous alphavirus exposure (n=10); responder subjects were previously vaccinated with the live-attenuated VEEV vaccine (TC-83) and had a demonstrated positive titer in response to vaccination (n=10); nonresponder subjects were previously vaccinated with the live-attenuated VEEV vaccine (TC-83) but did not develop a positive titer after vaccination (n=3) (Erwin-Cohen et al., 2012). The researchers harvested PBMCs from each volunteer and infected the cells *in vitro* with VEEV TC-83 and then harvested samples at 1, 4, 12, or 24 hours post-

infection (Erwin-Cohen et al., 2012). The RNA isolated from the PBMC samples was hybridized to Affymetrix HG U133 Plus 2.0 Genechip arrays which contain 54,000 probe sets representing over 38,500 genes (Affymetrix, 2004; Erwin-Cohen et al., 2012). This study was unique among the other, similar studies which examined molecular epidemiological changes induced by vaccination (as a surrogate for virulent infection) as it included a group of subjects that had failed to respond to previous vaccination, yet the changes in gene expression upon re-challenge with the vaccine strain of virus were not the same as the control group (i.e., representing an innate response) which suggested that while the nonresponders did not produce an adequate neutralizing antibody response, they did mount some type of response – albeit an uncharacterized response (Erwin-Cohen et al., 2012). The canonical pathways that were most notably shown to be involved in recall responses included the pattern recognition receptors in recognition of bacteria and viruses, the IL-12 signaling pathway, the IL-15 production pathway, and the TREM-1 signaling pathway (Erwin-Cohen et al., 2012). Altered transcription among the naïve and responder populations included up regulation of a number of interferon response genes (OAS1, OAS2, OAS3, IFNB1, IRF7), genes associated with activation of the inflammasome (DDX58, MYD88, IFIH1), as well as cytokine (IL-6) and toll-like receptor transcripts (TLR3) (Erwin-Cohen et al., 2012). The changes in transcription elicited in the nonresponder samples were predominantly observed to be decreased levels of transcription and included down regulation of TLR4, TLR8, C5AR1, NLCR4, CLEC7A, C1QA, C1QB, and C1Q3 (Erwin-Cohen et al., 2012).

Strengths and Weaknesses of Previous Experimental Approaches and Controversial Issues

There are several variables from previous studies which presented either strengths or weaknesses in relation to the data, including the sample type (whole blood vs PBMCs), time of sampling, the number of human volunteers, the gender of the human volunteers, the controls used, and the type of array used to assess changes in gene transcription.

Sample type. Previous attempts to catalog the human molecular epidemiology of infection with Ft have been conducted in explanted PBMC cells or lymphoid-derived subsets purified from PBMCs (Andersson et al., 2006; Fuller et al., 2006; Fuller et al., 2007; Paranavitana et al., 2008a, Paranavitana et al., 2008b; Paranavitana et al., 2014). The use of explanted PBMCs was employed by Das et al. (2008) and Erwin-Cohen et al. (2012) to study the molecular epidemiology of VEEV. Gene expression analyses performed on whole-blood samples offer a distinct advantage over PBMCs for the ability to suggest which cell types are important in different phases of the infectious process and development of immunity. RNA samples "derived from whole blood capture RNA profiles of all cell types in whole blood, including erythrocytes, granulocytes (neutrophils, eosinophils, and basophils), lymphocytes, monocytes, and platelets" (Joehanes et al., 2012, p. 59). In contrast, PBMC samples isolated from Ficoll-gradient centrifugation contain primarily lymphocytes and monocytes and are devoid of granulocytes, platelets, and reticulocytes (Joehanes et al., 2012). The use of isolated PBMCs can provide insight into the changes in gene transcription within these cellular subsets; however, using whole unfractionated blood to examine transcriptional changes

which occurred in vivo included critical regulation and influence of important cell to cell signaling pathways that may not have been captured in the PBMC populations.

Time points of sample collection. Time can be viewed as an important limitation of the previous studies that may have impacted the utility and interpretation of the data collected. In the PBMC studies of Ft-induced changes in gene expression conducted by Fuller et al. (2006, 2007) RNA from PBMCs was isolated following vaccination according to the following schedule: 6 days prior to vaccination, 18 hours following vaccination, and then at 2, 8, and 14 days after vaccination. Transcription profiles were compared back to a single pre-vaccination time point; no control (unvaccinated) subjects were used in the study (Fuller et al., 2006, 2007). The time points were chosen to represent early (18 hours, day 2), intermediate (day 8), and late (day 14) stages of infection (Fuller et al., 2006, 2007). Paranavitana et al. (2008b) isolated PBMCs and infected cells from each individual (n=8) with Ft then harvested RNA at 1, 4, 8, 16, and 24 hours post-infection; a second set of PBMC cells from each individual was mock-infected with saline to provide time-matched control samples. To describe the transcriptional responses of a memory or recall response, Paranavitana et al. (2014) isolated PBMCs from ten volunteers who had been previously vaccinated with LVS Ft (approximately 6-9 months prior) and developed a positive titer in response to vaccination. PBMCs were then subjected to in vitro infection with LVS Ft and harvested at 24 hours post-infection for transcriptional profiling or at both 24 and 72 hours post-infection for transcriptional analysis of Th17 cytokine expression (Paranavitana et al., 2010; Paranavitana et al., 2014). The early time points from the in vitro studies contrasted

with the time points estimated from the single in vivo transcriptional study of Ft whereby the time of sample collection could only be estimated relative to the date of presentation/assessment for medical treatment (Andersson et al., 2006). In the Andersson et al. (2006) study, patients with clinical signs of tularemia infection had six blood samples drawn over the course of diagnosis and treatment for ulceroglandular tularemia that were equivalent to days 2-3, 6-7, 8-9, 10-11, 13, and 90 days following first presentation at hospital.

The sampling time frames for transcriptional analyses of VEEV were similar to several of the Ft studies. Das et al. (2008) used PBMCs from at least three individual and infected the cells with virulent VEEV then collected the cells for RNA isolation at 1, 4, and 8 hours post infection. Erwin-Cohen et al. (2012) used PBMCs isolated from three independent groups of subjects to study the transcriptional responses to VEEV vaccination; a naïve group, responder group which had previously been vaccinated with VEEV and developed an appropriate neutralizing antibody response, and a non-responder group which had previously been vaccinated with VEEV but failed to develop an adequate neutralizing antibody response. PBMCs isolated from each individual were infected in vitro with VEEV and samples were subsequently harvested at 1, 4, 12, and 24 hours post infection (Erwin-Cohen et al., 2012).

Number of subjects in study. The number of study participants has been limited for several of the previous studies. Relevant transcriptional profiling studies have incorporated as few as three subjects (Das et al., 2008) to as many as ten subjects in each sampling category (Erwin-Cohen et al., 2012; Paranavitana et al., 2014) to assess changes

in transcription to either Ft or VEEV. Andersson et al. (2006) recruited a total of seven patients which presented with naturally-occurring infection, comparing transcriptional profiles of infected samples against control samples collected at a single, unrelated time point from eight healthy volunteers. The number of participants per group in all previous relevant studies was fewer than or equal to ten; however, the observance of differentially expressed transcripts indicated that the sample sizes were sufficient to generate statistically significant data. The issue of sample size and how that relates to powered studies will be discussed further in the Materials and Methods chapter.

Age and sex of study and control subjects. The ages of study participants for the studies relevant to the dissertation research ranged within each study. All volunteers were adult (>18 years of age) however some of the study participants were middle-aged (45-64 years of age) or elderly (> 65 years of age). The range of age of volunteering subjects was an important factor, particularly for VEEV, due to the manifestation of clinical disease and the greater threat to very young (children) and older populations for severe disease (Weaver & Reisen, 2010). Among the Ft studies, the volunteers who donated PBMCs for the Fuller et al. (2006, 2007) studies ranged in age from 22-54 years of age and consisted of a total of five subjects, three of whom were male and two were female. The subjects who donated PBMC cells for the examination of naïve responses to Ft conducted by Paranavitana et al. (2008b) ranged from 21-44 and were all male. The studies of recall responses conducted by Paranavitana et al. (2010, 2014) consisted of PBMCs isolated from all male volunteers who ranged in age between 18-50 years. The study participants for the study conducted by Andersson et al. (2006) were roughly equal

in terms of gender, with four male and three females who ranged in age between 50-76 years of age. In the previous human VEEV transcriptional studies, the ages of subjects who donated PBMCs ranged from 19-61 for the Das et al. (2008) study and included both male and female subjects; the exact numbers of male versus female participants was not disclosed. Finally, the study conducted by Erwin-Cohen et al. (2012) consisted of all male volunteers whose ages ranged from 18-45 years of age.

One important feature that distinguishes several of the studies is the use of control volunteers which are matched to the vaccinated volunteers by age and sex. Matching volunteers by gender and age was reported as a strengthening refinement to the study design by Paranavitana et al. (2010, 2014) and Erwin-Cohen et al. (2012).

As a single variable, the sex of research volunteers has been discussed as an important consideration in the analysis of immunological and microarray data due to the confounding effects that hormonal signaling may have on gene transcription and immunity (Cook, 2008; Giefing-Kröll, Berger, Lepperdinbger, & Grubeck-Loebenstein, 2015; Klein, Jedlicka, & Pekosz, 2010; Klein, Marriott, & Fish, 2015). Few studies have examined the disparity in responses in depth, however, Cook (2008), Klein et al. (2010), and Giefing-Kröll et al. (2015) conducted meta-analyses of the reported results of vaccine studies that demonstrated variable results in terms of humoral response or transcriptional response. The most relevant aspects of those studies to the dissertation research were the results in males versus females with regard to response to viral vaccination; transcriptomic data from a Yellow Fever Virus vaccine (17D) were obtained and analyzed independently by Klein et al. (2010) but the novel analysis included the contrast

of male-female responses, which was lacking in the first analysis reported by Querec et al. (2009). Klein's re-analysis of the transcriptome data described a total of 660 genes that were differentially expressed in response to YF vaccination in women, but only 67 genes differentially expressed in men only (of which, 64 were common to both men and women) (Klein et al., 2010). Additionally, adverse event reports were more frequent and severe in women than in men (Klein et al., 2010). Cook (2008) also reported on the differences observed between men and women regarding vaccine response in a meta-analysis study that examined sex-differences in antibody production following vaccination. Cook (2008) found that in most of the viral vaccine studies, the antibody response from males was greater than that reported for females, including vaccines for measles, yellow fever, Venezuelan equine encephalitis, and in two of three rabies studies. However, females produced a more robust antibody response against vaccination for influenza, rubella, and in one of three rabies studies (Cook, 2008). Differences in mortality were noted to be greater in men than women in response to infection with hepatitis B, hepatitis C, and rabies (Giefing-Kröll et al., 2015) Differences in immune and transcriptional responses have been postulated to be attributable to a greater stimulatory effect of sex-related hormones in women, particularly for genes that stimulate and are regulated by interferon pathways which may lead to rapid clearance of viruses (Cook, 2008; Klein et al., 2010). To that end, the studies which employed only male volunteers may have had an advantage by reducing any bias due to sex (and fluctuating hormonal patterns) particularly within such small sampling sizes, but overall the use of male-only study populations reduces the generalizability of the study. Interestingly, a meta-analysis

study conducted to address potential sex-specific variability in gene expression was conducted to address concerns of sex bias in preclinical studies, where the use of male study subjects has been disproportionate to the use or inclusion of female study subjects (Itoh & Arnold, 2015). The researchers compiled 293 microarray datasets, encompassing both human and mouse data from various tissues including brain, spleen, adrenal glands, kidney, and muscle; however, notably did not include blood or blood cells as a tissue source (Itoh & Arnold, 2015). The results of the analysis demonstrated only very small, minor variations in gene expression in the target tissues suggesting that any observed differences between sexes were likely to be regulated downstream of transcription and were likely subject to temporal effects of gonadal hormones (Itoh & Arnold, 2015). The issue of sex in relation to aging was discussed by Gubbels-Bupp (2015) and summarized to note changes in immune cell phenotypes in humans over the age of sixty to sixty-five years particularly for natural killer (NK) cell populations, which were dramatically increased in females, but with markedly decreased B and T cells numbers and correspondingly decreased production of IL-2, IFN-gamma, IL-17 cytokines in men.

The effects age also represented a contentious variable, as discussed by a number of studies (Arlt & Hewison, 2004; Giefing-Kroöll et al., 2015; Gubbels-Bupp, 2015). In addition to the alterations noted in immune cell phenotype and function related to gonadal hormone signaling, Gubbels-Bupp (2015) noted many changes in B and T cells that were decreased in older populations, as well as decreases in interferon stimulated responses. The issue of immune responses in relation to age may be a serious weakness to the transcriptional study conducted by Andersson et al, 2006; samples from both males

and females combined with advanced age (ages ranged from 50-76 years of age in the study) of the subjects constitutes a weakness to the model of Ft molecular epidemiology due to the effects of immunosenescence, as has been suggested by Arlt and Hewison (2004).

Type of array used to assess transcription. The type of microarray chip used by researchers in previous studies has predominantly been the Affymetrix human genome U133 2.0 Plus array chip, with two exceptions: The analysis conducted by Andersson et al. (2006) utilized the Affymetrix HG U133 GeneChip, which was an earlier version of the Affymetrix HG U133 2.0 array chip; and the analysis conducted by Das et al. (2008) employed custom cDNA arrays.

Previous Studies of Association of Age or Genetic Makeup with Vaccine Outcome: Relation to Key Study Variables

There are several studies which have described the association of age with vaccine outcome (Pawelec & Derhovanessian, 2010). The response of elderly people to influenza vaccination is greatly reduced (17-53% efficacy) in comparison to young adults who demonstrate between 70-90% efficacy in response to influenza vaccination (Pawelec & Derhovanessian, 2010). In a study of naïve aging individuals immunized with hepatitis B vaccine (HBV), age was significantly associated with vaccine response; vaccinees aged 40 years or younger responded better to vaccination than did individuals who were 65 years of age or older (Fourati et al., 2015). The mechanisms of reduced vaccine response may be related to declining T cell populations in older individuals as elderly men were found to have reduced T cells populations, in comparison to younger individuals (De

Benedetto, Derhovanessian, Steinhage-Thiessen, Goldeck, Müller, & Pawelec, 2015). There are no studies to date that have addressed a potential relationship between vaccine outcomes and age for vaccination with Ft. Pittman, Liu, Cannon, Mangiafico, and Gibbs (2009) discussed a small study in which age was investigated as a demographic factor for in relation to vaccination with VEEV; however the authors reported only that there was no evidence of age being associated with primary vaccine failure in individuals who had previously been vaccinated with two other vaccines which are related to VEEV, eastern and western equine encephalitis viruses.

Several studies of infectious diseases have reported the association of specific genetic factors with vaccine outcome in response to vaccination with childhood vaccines [HBV, diptheria, tetanus, pertussis (DTaP)], smallpox, influenza, anthrax, and rubella (Li et al., 2013; Narwaney et al., 2013; Ovsyannikova et al., 2007; Ovsyannikova et al., 2013; Ovsyannikova et al., 2014; Yucesoy et al., 2013). HLA are encoded by the MHC genes (Li et al., 2013). HLA class II molecules are responsible for presentation of antigens to ultimately elicit antigen-specific B-cell responses (Li at al., 2013). There have been no studies to date which address the potential contribution of HLA phenotype with vaccine outcome in response to vaccination with either Ft or VEEV.

Rationale for Selection of Key Variable and Unanswered Questions

For each of the relevant previous transcriptional studies of both Ft- and VEEV-induced changes in gene expression, the dependent variable was common to all of the studies: changes to gene transcription (Andersson et al., 2006; Das et al., 2008; Erwin-Cohen et al., 2012; Fuller et al., 2007; Paranavitana et al., 2008b; 2014). The independent

variables previously employed have predominantly been time and treatment; however, the study conducted by Erwin-Cohen et al. (2012) also included the additive effect of time*treatment as an independent variable.

Random variables included age and gender of the study subjects. The relevant transcriptional studies varied in the use of study population; while three utilized study subjects of both sexes (Andersson et al., 2006; Fuller et al., 2007;, Das et al., 2008), three other studies used male-only subjects (Erwin-Cohen et al., 2012; Paranavitana et al., 2008b; Paranavitana et al., 2014). The effect of sex in relation to changes in gene transcription in response to vaccination was discussed briefly in a study which examined the gene transcription induced by yellow fever virus vaccination, but sex as an independent variable remains to be studied in depth (Klein et al., 2010). It has been suggested that elimination of females from study populations may be a method to limit sex as a covariate and changing levels of hormones in females as a time-dependent covariate (Erwin-Cohen et al., 2012; Paranavitana et al., 2008b; Paranavitana et al., 2014). Age was limited in several of the studies to include only adult study subjects (but not elderly subjects) (Erwin-Cohen et al., 2012; Paranavitana et al., 2008b; Paranavitana et al., 2014). The effect of age on transcription in response to vaccination has not been addressed directly and thus remains somewhat controversial. The potential relationship between age and the outcome of vaccination to either Ft or VEEV has not been addressed directly; with the caveat that Pittman et al. (2009) briefly discussed age of VEEV vaccinated individuals as a potential factor in primary vaccine failure following

sequential administration of other, related alphaviral vaccines [against Eastern equine encephalitis virus (EEEV) and western equine encephalitis virus (WEEV)].

Summary of Previous Studies and Relation to Research Questions

Previous studies of the transcriptional responses to vaccination with live-attenuated Ft or VEEV or to responses from naturally occurring Ft infection were groundbreaking glimpses into the molecular epidemiology of human infection with either agent. Both studies conducted by Andersson et al. (2006) and Fuller et al. (2007) incorporated assessments of the molecular epidemiology of Ft infection, via naturally occurring infection and live-attenuated vaccination, respectively. Several experimental conditions were similar, including the small study size (between 5-7 persons), the inclusion of both male and female study subjects, and overlapping time points (Day 2-3, Day 8-9, Day 13-14) however the results reported only a single transcript that overlapped at any time point, TAP-1 (Andersson et al., 2006, Fuller et al., 2007). Limitations that may have affected the results were the comparison of experimental results against a single pre-vaccination time point (Fuller et al., 2007), comparison against healthy controls from only a single time point (Andersson et al., 2006), the diverse ages of the study subjects, and the small sample sizes that each study employed. In contrast, in an *in vitro* PBMC Ft study, Paranavitana et al. (2008b) recruited eight male subjects with an age range of 21-44 years but also compared the experimental transcriptional responses against uninfected PBMCs at the same time points as the infected cells (1, 4, 8, 16, and 24 hours post infection). The reported results corroborated results from Andersson et al. (2006) with the observation of increased transcription of CASP1, PSME2, TAP-1, GBP1,

and GCH1 predominantly at time points between 16-24 hours post infection (Paranavitana et al., 2008b). Comparison of transcriptional results reported by Fuller et al. (2007) revealed an overlap in increased transcript expression of LY96, TAP-1, CD80, ICAM-1, and TLR2 at 16-24 hours post infection (Paranavitana et al., 2008b).

The transcriptional changes observed in response to *in vitro* infection with the live-attenuated strain of VEEV were deduced from a slightly larger sample set than was used for most of the Ft studies; Erwin-Cohen et al. (2012) examined PBMCs from individuals in three groups, previously vaccinated responders, vaccinated nonresponders, and naïve individuals. The sample size for each group was ten subjects with the exception of the nonresponder group, which only consisted of three volunteers (Erwin-Cohen et al., 2012). The naïve and response groups were infected with the live-attenuated strain of VEEV *in vitro* and followed for 1, 4, 12, or 24 hours post infection; analysis was conducted by comparison to time-matched naïve PBMCs (Erwin-Cohen et al., 2012). The temporal changes in gene expression were described and included differentially expressed transcription in response to infection as well as pathway analysis and biomarker assessment for each response group (Erwin-Cohen et al., 2012). A previous study in human PBMCs by Das et al. (2008) included *in vitro* infection of PBMCs with a variety of biological agents, including VEEV but did not report transcriptional profiling results for VEEV.

These results highlighted the need to understand the temporal changes in gene expression in vivo, in a manner that reduced potential transcriptional variability due to sex and hormonal signaling but that also could compare experimental results with control

results, at equally matched time points with control-experimental subjects matched for age and within a purely adult age range. The use of whole, unfractionated blood, as opposed to PBMCs purified from whole blood, was an important refinement to the dissertation study due to the complex cell-to-cell signaling and regulation constraints that would only exist *in vivo*. These issues prompted and informed the research questions for this dissertation work, including the questions of what are the changes to transcription, what molecular functions do the transcripts play, and what biological processes are involved in the development of immunity following vaccination with live-attenuated Ft or live-attenuated VEEV. The time points to be considered in the dissertation research were also influenced by previous results, thus the temporal range of the dissertation study will span from pre vaccination (0 hours) to 1, 4, 8, 24 hours post vaccination and then from days 2, 7, 14, 21, and 28 post vaccination.

There are no studies, to date, that have addressed a potential relationship between vaccine outcomes and age for vaccination with Ft. There have been no studies to date which address the potential contribution of HLA phenotype with vaccine outcome in response to vaccination with either Ft or VEEV. Molecular epidemiology studies in other infectious diseases have provided innovative information in humans and animal models; however, there is little understanding or information of the molecular consequences of human infection with either Ft or VEEV. This dissertation project seeks to make a novel contribution to the field of public health by studying the relationship between gene expression and the temporal effects of vaccination, the potential relationship between age

and vaccine outcome, and the potential association between genetic make-up of individuals with vaccine outcome.

Summary and Conclusions

There have been few published studies which have described the molecular epidemiology of human infection with either Ft or VEEV; many studies directly relevant to the dissertation research were conducted more than five ago. The lack of recently published reports is an important part of the knowledge gap that surrounds the research questions. Knowledge of the molecular epidemiology of each infectious disease may provide critical data which can be used to design modern, rational vaccines (since no licensed vaccine currently exists for either infectious agent) but may also provide important epidemiological information about common host responses to infection or novel host targets for diagnostic tests. A broad review of the literature demonstrated that there remains a profound lack of knowledge and published literature in recent years regarding the effects of either Ft or VEEV infection in humans.

My dissertation research study addresses gaps in the literature by examining potential association between changes in the level of gene transcription over time following vaccination with either Ft or VEEV. My study also investigates the potential correlation of age with vaccine out come and the relationship between HLA phenotype and vaccine outcome.

The methodology used to assess changes in the transcriptome include a temporal assessment of RNA collected at specific time points following vaccination with a live-attenuated virus or mock-vaccination (either Ft or VEEV) in human volunteers who had

been matched for age and gender. Samples of whole, unfractionated blood were collected from which RNA was isolated. Samples of whole, unfractionated blood were collected from which RNA was isolated. I leverage the technique of examining changes in the transcription of genes following an event (i.e., vaccination) to provide a snapshot of gene activity as it changes in response to vaccination. Some genes may be triggered to increase transcriptional levels, and in theory protein levels as well, while other genes may be induced to decrease transcription. Some of the changes are in response to the infectious invasion, whereas others are triggered to elicit development of immunity. Comparison of expression levels were normalized and assessed for analysis of variance using Partek Genomics Suite software. Analysis of biological process and molecular functions were conducted with Partek Genomics Suite Pathways Module and Ingenuity Pathway Analysis Software.

Chapter 3: Research Method

Study Purpose

The purpose of this dissertation project was to quantitatively assess the molecular epidemiology of Ft and VEEV in a human vaccine model. I worked to quantitatively assess the potential relationship between gene expression over time in response to treatment (vaccination) with either Ft or VEEV, to assess the potential relationship between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination or vaccine failure), and to identify the potential correlation between HLA phenotype and vaccine response. This research was intended to add to the limited published information about the changes in gene expression that each vaccine elicits, and to address the gaps in knowledge regarding the association of age and genetic makeup on vaccine outcome.

In the following sections, I describe the research design, including the use of pre-existing data sets and refinements to the independent variables; the methodology of how the samples were generated, including details of the study populations recruited, rules for inclusion and exclusion, and the data analysis plan; and threats to validity, including ethical concerns.

Research Design and Rationale

This dissertation study was a novel, quantitative, secondary analysis. The study variables included the changes in the level of transcript expression as a dependent variable. The independent variables, which related to Research Questions 1 and 2 were: Time of blood sampling post vaccination; treatment – either vaccination of study subjects

with live-attenuated Ft or VEEV; or a third independent variable which combined the effect of time*treatment as a single variable. The study design was developed to produce data analyses to directly answer Research Questions 1: Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft or VEEV? and Research Question 2: Is there a significant association between changes in the level of gene transcription over the time course of vaccination with Ft in comparison to VEEV?

The dependent variable of vaccine outcome (measured by neutralizing antibody titer), and the independent variable of HLA alleles were the study variables which I used to address Research Question 4: Is there a significant association between HLA allele and vaccination outcome?

Use of Secondary Datasets to Reduce Time and Resource Constraints

The data sets I used in this novel secondary analysis were previously generated (USAMRIID FY05-01, 2012; USAMRIID FY06-17, 2012). By using these existing data, I eliminated time and resource constraints that would otherwise have be prohibitive to the execution of the study design due to the time and expense of writing/approval of a human use study protocol; recruitment, consent, and treatment of human volunteers; access and approval to use investigational new drug vaccines; and costs associated with microarray processing and hybridization.

Refined Study Parameters Retain Consistency with Previous Studies

The study design was consistent with other published studies (Andersson et al., 2006, Erwin-Cohen et al., 2012; Fuller et al., 2007; Paranavitana et al., 2008b, 2014)

relevant to the dissertation topic, but included further refinements to the study design through the use of whole, unfractionated blood as a source of RNA; incorporation of time points that covered very early (within the first 24 hours), intermediate (2-14 days), and late (21-28 days) events following vaccination; and implementation of an equal number of age-matched control study subjects that were mock-vaccinated but had blood samples drawn on the same schedule and at the same time as the vaccinated study subjects. These additional measures to refine the study design expanded the improvements employed by Paranavitana et al. (2014), which included using PBMCs sampled from an equal number of study volunteers. With regard to the VEEV portion of the study, my design choice for the dissertation research was consistent with previously published human study reports (Erwin-Cohen et al., 2012) in terms of sample size and age-matched controls, but additional refinements to the study included the use of whole, unfractionated blood rather than PBMCs, and the time points of blood sample collection were expanded to cover the very early events following vaccination (less than one day), intermediate events (2, 7, and 14 days) and late transcriptional events (21-28 days).

Methodology

In this study, I used secondary data sets which were collected through previous studies conducted at USAMRIID (USAMRIID protocol FY05-01, 2012, and USAMRIID FY06-17, 2012). The original studies used research protocols for both independent studies conducted under good clinical practice (GCP) quality systems, approved by the USAMRIID institutional review board (IRB), and under which volunteers signed a written informed consent document (ICD) prior to enrollment in each study which

described the purpose of the study and the manner in which samples would be collected, used, and disposed. The original study protocols described the purpose and goals of the studies and how the microarray data would be used. No personally-identifiable information was contained in the data sets or supporting documents, and the data were available to open publication in medical or military journals. I used the microarray data generated by the previous two studies, and combined that data to generate a novel secondary analysis. In the next sections, I describe procedures regarding how the data were originally collected and generated.

Study Populations from Original Data

The selection processes for both of the original Ft and VEEV studies were designed in a similar manner. Ft study subjects consisted of male volunteers between the ages of 21 and 44 years. Male volunteers were used to reduce potential variability in global gene expression that could be attributed to hormonal changes occurring during the menstrual cycle, rather than changes induced by the Ft vaccine. All potential study participants were screened for previous exposure to Ft and demonstrated to be negative by microagglutination titer assay. Ten study volunteers received 0.5 ml of attenuated Ft live vaccine strain (LVS) (NDBR-101 vaccine) administered by skin scarification with a bifurcated needle. The same skin scarification process was used to administer 0.5 ml of saline to 10 control subjects. Whole blood was collected from control and vaccine volunteers at specific time points prior to (0 h) and post-vaccination (1, 4, 8 hours, and days 1, 2, 14, 21, and 28) for isolation of RNA. On day 28, serum was tested from members of the vaccine group to assess development of a titer against tularemia. RNA

samples from ten Ft vaccinees and ten control subjects were used for the microarray analysis. This study was conducted in accordance with USAMRIID Human Use Protocol FY05-01.

The population for the VEEV portion of the study consisted of 20 male volunteers between the age of 23 and 48 years. To eliminate potential variables in the study which could be attributed to estrogen signaling, only male volunteers were selected for the study. In addition, each study volunteer was age-matched to a control volunteer. Study participants included individuals who had not previously received any alphavirus IND vaccines (e.g., EEE, WEE, or VEE). Prior to enrollment and participation in the study, all study participants were screened for antibodies by enzyme-linked immunosorbent assay (ELISA) and plaque reduction neutralization titer (PRNT) assays for previous exposure to new world alphaviruses (VEE, EEE, and WEE) and demonstrated to be negative for previous exposure. Whole blood was collected from control and vaccine volunteers at specific time points prior to (0 h) and post-vaccination (1, 4, 8, and 24 hours, and days 2, 7, 14, 21, and 28) for isolation of RNA. On day 56, serum was tested from members of the vaccine group to assess development of a neutralizing antibody titer against VEEV. RNA samples from nine vaccinees and 10 control subjects were used for the microarray analysis; one vacinnee was removed from the study analysis as the individual failed to develop a positive neutralizing antibody response following vaccination and was deemed to be a vaccine nonresponder. This study was conducted in accordance with USAMRIID Human Use Protocol FY06-17.

Inclusion and Exclusion Criteria Used in Original Study

For inclusion in the original Ft study, clinical researchers required volunteer and control group volunteers to be male, between 18-49 years of age, willing to forego blood donation for at least 56 days following completion of participation in this protocol, and have signed the ICD. In addition, volunteers in the vaccine group were required to be participants in the special immunization protocol program (SIP) and be scheduled to receive the tularemia vaccine because of occupational risk of exposure to Ft.

To be eligible for inclusion in the original VEEV study, researchers required volunteer and control group volunteers to be male between the ages of 18-49 years of age, willing to allow study personnel to review their most recent HIV and hepatitis B and C test results and VEE vaccination records including titers, willing to undergo all blood draws including blood draws for complete blood counts (CBC), VEE titer (control group), ELISA for alphavirus exposure, and HIV and hepatitis screening tests if these latter two tests had not been performed within the past year, forego blood donation of greater than 125 mL for at least 30 days following completion of participation in the protocol, and have signed the ICD. Additionally, volunteers in the vaccine group were required to be participants in the SIP and be scheduled to receive TC-83 vaccine (NDBR-102) because of occupational risk of exposure to VEE virus.

Criteria which researchers used to exclude volunteers from participation in either original study protocol included a history of immune system abnormalities (e.g., positive HIV test, positive hepatitis B or C test), less than 18 years of age, hematology screening with parameters out of normal laboratory range (e.g., anemia), previous vaccination with

Ft or any new world alphavirus or immunity from naturally-occurring infection, receipt of any vaccine 30 days prior to or following vaccination with either Ft, VEEV, or placebo; donation of 125 mL or more of blood within 30 days prior to or following vaccination; having received antiviral medication within 7 days prior to vaccination; and taking immunosuppressive therapy (excluding topical steroids).

Power Calculations and Sampling Size

The goal of my study using the microarray experiment as a secondary data source was to detect significantly differential expression between two types of samples, namely naïve and vaccinated. The Affymetrix human HG-133 Plus 2.0 microarray chips contain 47,000 transcripts representing 38,500 genes (Affymetrix, 2004). I reassessed the power that could be achieved by combining the two datasets using the M. D. Anderson bioinformatics sample size for microarray experiments calculator (M. D. Anderson, 2016). I set the number of genes to be assessed at 47,000 with an acceptable number of false positives set to 4700 (a 10% false discovery rate). The minimum fold difference desired for detection between two samples was set to 2 and the desired power was set to 90%. I used a suggested standard deviation of 0.75, and the sample size for each group was calculated to be 10 subjects to generate a per-gene value of alpha of 0.1 (M. D. Anderson, 2016). Power calculations for each original study were conducted using Statistical Applications Software (SAS) and estimated that using 10 individuals per group would typically yield 95% power to detect 3.5-fold differences in gene expression using a two-sample t-test at the 0.001 two-sided significance level, assuming that the coefficient of variation was 0.50. My dissertation research refined the power calculations to include

transcripts that displayed 2-fold changes in gene expression to capture a broader view of transcriptional changes and to remain consistent with other relevant published reports.

Sampling Procedures of Original Study

Each original study protocol reported that study samples were collected as a stratified random sampling of adult males, stratified by treatment (mock-vaccination with saline or vaccination with live-attenuated Ft or live-attenuated VEEV vaccines).

Blood samples were drawn by percutaneous venous puncture using a 19- or 21-gauge straight or butterfly needle. To minimize local effects of vaccination/placebo administration, blood samples were drawn from the arm that did not receive the vaccine/placebo (contralateral arm) whenever possible. Up to 10 mL of whole blood was drawn for transcriptome studies at each time point and collected in PAXgene tubes. The sampling timeline included a pre vaccination sample on the day of vaccination (0 hour), then subsequent sampling at 1, 4, 8, and 24 hours post vaccination, and days 2, 7, 14, 21, and 28 post vaccination. Samples were collected along an identical timeline, with the exception that the Ft study did not collect a sample at day 7 post vaccination. RNA was isolated from samples using the PAXgene RNA isolation kit (QIAGEN, Inc., Valencia, California). Samples were hybridized to Affymetrix HG U133 2.0 Plus microarray chips according to manufacturer's protocol.

Sampling in Original Study

The VEEV *in vivo* study, conducted under the guidelines of an approved human use protocol (USAMRIID FY06-17, 2012), included 10 vaccinees who received 0.5 ml live-attenuated TC-83 VEE virus (NDBR-102 vaccine), roughly equivalent to 1.7×10^5

plaque forming units (PFU) of virus, administered subcutaneously (SC) in the upper outer aspect of the arm, as well as 10 control subjects who were administered 0.5 ml saline via the same procedure. Researchers collected whole, unfractionated blood at specific time points immediately prior to (0 hours) and following vaccination (1, 4, 8 hours and days 1, 2, 7, 14, 21 and 28) and from which RNA was isolated. On day 56 post vaccination, a serum sample was collected from volunteers to test for development of neutralizing antibody titer against VEE virus; one vaccinee failed to respond to vaccination, as measured by the lack of neutralizing antibodies against VEEV, and was removed from the analysis. Thus the data set for my secondary analysis was comprised of RNA samples from the 9 responsive vaccinees and 10 control subjects which were then subjected to microarray analysis.

The procedures conducted by clinical researchers under an approved human protocol for the Ft study included administration of 0.5 ml of attenuated Ft [live vaccine strain (LVS) (NDBR-101 vaccine)] to 10 study volunteers, in which the vaccine was given via skin scarification with a bifurcated needle. The same skin scarification process was used to administer 0.5 ml of saline to 10 control subjects. Whole blood was collected from control and vaccine volunteers at specific time points prior to (0 hours) and post vaccination (1, 4, 8, 24 hours, and days 2, 14, 21, and 28) for isolation of RNA. On day 28, a serum sample was collected and tested from members of the vaccine group to assess development of an antibody titer against Ft. RNA samples from 10 vaccinees and 9 control subjects were used for the microarray analysis; an incomplete sample set disqualified the samples from one control volunteer.

Use of Archival Data

Procedures for data collection, including recruitment, consent, and participation, have been described above. I used archival data in the form of Affymetrix .CEL files for my secondary analysis that were provided by USAMRIID. I received permission to access and use data as well as publish results relating to the human data sets from USAMRIID FY05-01 and FY06-17 which was granted by United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Office of Human Use and Ethics (OHU&E), Human Use Committee (HUC) on 11 March 2016 (Appendix A). The Exempt Determination Official (EDO) made the determination that the data do not meet the definition of "human subject" under 32 CFR 219.102 (f) (1) nor (f) (2).

Research Questions and Hypotheses

Research Question 1. Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft or VEEV?

Null Hypothesis 1. There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), followed by assessment of fold-change (< 2.0 in either direction)] in the mean gene expression level for any given transcript in the vaccinated group (for Ft or VEEV) compared to the control group, or between vaccinated and control at any specific time point.

Alternative Hypothesis 1. There is a significant change [first assessed by Step Up-corrected p-value ($p<0.1$), followed by assessment of fold-change (> 2.0 in either direction)] in the mean gene expression level for any given transcript in the vaccinated group (for Ft or VEEV) compared to the control group, or between vaccinated and

control at any specific time point.

Research Question 2. Is there a significant association between changes in the level of gene transcription over the time course of vaccination with Ft in comparison to VEEV?

Null Hypothesis 2. There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), then assessed by fold-change in expression (< 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between the vaccinated groups at any specific time point.

Alternative Hypothesis 2. There is a significant change [first assessed by Step Up-corrected p-value ($p<0.1$), then assessed by fold-change in expression (> 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between the vaccinated groups at any specific time point.

Research Question 3. Is there a correlation between age of the vaccinated subject and vaccination outcome?

Null Hypothesis 3. There is no correlation between age and postvaccination titer.

Alternative Hypothesis 3. There is correlation between age and postvaccination titer.

Research Question 4. Is there a significant association between Human Leukocyte Antigen (HLA) phenotype and vaccination outcome?

Null Hypothesis 4. There is no significant difference in postvaccination titer

between HLA phenotypes.

Alternative Hypothesis 4. There is significant difference in postvaccination titer between HLA phenotypes.

Data Analysis Plan

Software. I conducted quantitative analysis of transcriptional data using Partek Genomics Suite software (version 6.6) I made analyses of pathways through use of the Pathways Module within Partek Genomics Suite or Ingenuity Pathway Analysis (IPA) (IPA, 2016; Partek, 2016). Statistical tests for pathway analysis included the use of Fisher's exact test to examine the statistical likelihood of a gene being represented within a pathway (Li et al., 2013, IPA, 2016). Tests to measure the correlation of age with vaccine outcome or the association of HLA allele with vaccine outcome were conducted with GraphPad software.

Normalization of data. Data consisting of Affymetrix .CEL files were imported into Partek Genomics Suite (version 6.6). I used a data normalization configuration that included a background correction using the Robust Multiarray Algorithm (RMA) background correction algorithm on the perfect-match (PM) probe values (Bolstad et al., 2003; Partek, 2016) and data were normalized by quantile normalization, Log base 2 transformed, and underwent median polish summarization (Partek, 2016). As part of the statistical software program, I conducted quality assurance and quality check procedures which included conducting principal component analysis for the visualization of data variance.

Statistical tests and procedures. Following importation of individual Affymetrix

data files into Partek Genomics Suite software, background correction using the RMA algorithm, quantile normalization, and median polishing, I subjected the data to a batch-removal procedure to remove any biases to the data, which may have been due to unwanted effects of factors such as processing batches, and which may have occurred on different days or at different fluidic stations, for example (Irizarry et al., 2003).

After the batch-removal step has been executed, the cleaned data will be examined to elucidate the amount of variation and the sources of variation in the gene expression data by principal component analysis (PCA) (Partek, 2016). The parameters for assessing variation will include Time (hour or day of blood sample collection) and Treatment (control or vaccinated).

I then assessed changes in the levels of gene transcription mixed models analysis of variance (ANOVA) of the normalized data. The change in gene expression was expressed as a ratio of the mean expression level for any particular transcript at a specific time point in the vaccinated group relative to the mean expression level of the same transcript at the same time in the mock-vaccinated control group. The parameters for a 4-way mixed models ANOVA included comparisons of time, treatment, and the intersection of time*treatment using the following formula and methodology of restricted maximum likelihood (REML), rather than Method of Moments (MoM), as the variance component estimation method (Baayen, Davidson, & Bates, 2008; Krueger & Tian, 2004):

Model: $Y_{ijklm} = \mu + \text{scan date}_{i} + \text{time point}_{j} + \text{treatment}_{k} + \text{subject (treatment)}_{kl} + \text{time point} * \text{treatment}_{jk} + \varepsilon_{ijklm}$

Where Y_{ijklm} represented the m^{th} observation of gene expression on the i^{th} scan date j^{th} time point k^{th} treatment l^{th} subject; μ was the common effect for the whole experiment. ε_{ijklm} represented the random error present in the m^{th} observation of gene expression on the i^{th} scan date j^{th} time point k^{th} treatment l^{th} subject. The errors ε_{ijklm} were assumed to be normally and independently distributed with mean 0 and standard deviation δ for all measurements (Partek, 2016). Scan date and subject were considered random effects. I used the REML approach for estimation of variance components in favor of the alternative method of moments because the data sets were not balanced (Partek, 2016).

Although the data from each data set represented repeated measures from the same individuals, analysis of the data by repeated-measures ANOVA I could not employ analysis of the data by repeated-measures ANOVA because the data sets were not balanced (in each data set, there was one subject that had to be removed from the study, resulting in an unbalanced dataset for both Ft and VEEV). I chose a more robust method for analysis of repeated measures, complex, multivariate, temporal data analysis through use of a mixed models ANOVA, which is a type of repeated-measures ANOVA, with contrast analysis to answer the question of changes in level of gene expression (the dependent variable) as a function of the independent variables of time, treatment (vaccination) and the effect produced by the interaction of time and treatment. Contrasts

were calculated using Fisher's Least Significant Difference (LSD) (Partek, 2016; Tamhane & Dunlop, 2000). I performed the following contrast(s) to compare, (1) differential expression of transcripts induced by Ft vaccination; (2) differential expression of transcripts induced by VEEV vaccination; and (3) differential expression between Ft compared with VEEV:

Research Question 1: Part A- Differential expression induced by Ft vaccination:

0 hr * Ft Vaccine vs. 0 hr * Control

1 hr * Ft Vaccine vs. 1 hr * Control

4 hr * Ft Vaccine vs. 4 hr * Control

8 hr * Ft Vaccine vs. 8 hr * Control

day 01 * Ft Vaccine vs. day 01 * Control

day 02 * Ft Vaccine vs. day 02 * Control

day 14 * Ft Vaccine vs. day 14 * Control

day 21 * Ft Vaccine vs. day 21 * Control

day 28 * Ft Vaccine vs. day 28 * Control

Research Question 1: Part B- Differential expression induced by VEEV vaccination:

0 hr * VEEV Vaccine vs. 0 hr * Control

1 hr * VEEV Vaccine vs. 1 hr * Control

4 hr * VEEV Vaccine vs. 4 hr * Control

8 hr * VEEV Vaccine vs. 8 hr * Control

day 01 * VEEV Vaccine vs. day 01 * Control

day 02 * VEEV Vaccine vs. day 02 * Control

day 07 * VEEV Vaccine vs. day 07 * Control

day 14 * VEEV Vaccine vs. day 14 * Control

day 21 * VEEV Vaccine vs. day 21 * Control

day 28 * VEEV Vaccine vs. day 28 * Control

Research Question 2: Differential expression induced by Ft vaccination in comparison to VEEV vaccination:

(0 hr * Ft Vaccine + 0 hr * Ft Control) - (0 hr * VEEV Vaccine + 0 hr * VEEV Control)

(1 hr * Ft Vaccine + 1 hr * Ft Control) - (1 hr * VEEV Vaccine + 1 hr * VEEV Control)

(4 hr * Ft Vaccine + 4 hr * Ft Control) - (4 hr * VEEV Vaccine + 4 hr * VEEV Control)

(8 hr * Ft Vaccine + 8 hr * Ft Control) - (8 hr * VEEV Vaccine + 8 hr * VEEV Control)

(day 01 * Ft Vaccine + day 01 * Ft Control) - (day 01 * VEEV Vaccine + day 01 * VEEV Control)

(day 02 * Ft Vaccine + day 02 * Ft Control) - (day 02 * VEEV Vaccine + day 02 * VEEV Control)

(day 14 * Ft Vaccine + day 14 * Ft Control) - (day 14 * VEEV Vaccine + day 14 * VEEV Control)

(day 21 * Ft Vaccine + day 21 * Ft Control) - (day 21 * VEEV Vaccine + day 21 * VEEV Control)

(day 28 * Ft Vaccine + day 28* Ft Control) - (day 28* VEEV Vaccine + day 28* VEEV Control)

Or

(0 hr * VEEV Vaccine + 0 hr * VEEV Control) - (0 hr * Ft Vaccine + 0 hr * Ft Control)

(1 hr * VEEV Vaccine + 1 hr * VEEV Control) - (1 hr * Ft Vaccine + 1 hr * Ft Control)

(4 hr * VEEV Vaccine + 4 hr * VEEV Control) - (4 hr * Ft Vaccine + 4 hr * Ft Control)

(8 hr * VEEV Vaccine + 8 hr * VEEV Control) - (8 hr * Ft Vaccine + 8 hr * Ft Control)

(day 01 * VEEV Vaccine + day 01 * VEEV Control) - (day 01 * Ft Vaccine + day 01 * Ft Control)

(day 02 * VEEV Vaccine + day 02 * VEEV Control) - (day 02 * Ft Vaccine + day 02 * Ft Control)

(day 14 * VEEV Vaccine + day 14 * VEEV Control) - (day 14 * Ft Vaccine + day 14 * Ft Control)

(day 21 * VEEV Vaccine + day 21 * VEEV Control) - (day 21 * Ft Vaccine + day 21 * Ft Control)

(day 28 * VEEV Vaccine + day 28 * VEEV Control) - (day 28 * Ft Vaccine + day 28 * Ft Control)

I refined the results of the mixed model ANOVA analysis by conducting a multiple test correction in the form of False Discovery Rate (FDR), and set an acceptable limit of not more than 10% false discovery. As part of the FDR analysis, the p-value for each contrast was corrected by the Step-up method (Benjamini & Höchberg, 1995; Partek, 2016).

Finally, results of the mixed model ANOVA analysis were further refined to only include those transcripts where the change in gene expression level was 2-fold or greater in either direction (increasing transcription or decreased transcription) (Partek, 2016).

I performed Pathway analysis as part of the Partek Genomics Suite software (or Ingenuity Pathway Analysis software) on the down-selected list of significantly modified transcripts identified by the criteria of significant FDR-corrected Step-up p-value and 2-fold or greater change in expression. Significance of pathways was determined by enrichment score, calculated as the negative natural log of the enrichment p-value derived from the contingency table (Fisher's Exact test) and p-value < 0.05 (IPA, 2016; Partek, 2016).

I conducted statistical tests to address Research Question 3 with GraphPad software and tested the correlation between age as either a continuous variable or a dichotomous variable in relation to vaccine outcome (as a dichotomous variable).

I conducted Mann-Whitney U tests and one-way ANOVA statistical tests to address Research Question 4 on the association of HLA-DRB1 or HLA-DQB1alleles with vaccine outcome.

Interpretation of Results

The interpretation of transcript expression results as significant were determined first by a given transcript having an expression value with FDR-corrected Step-up p-value ≤ 0.1 and a 2-fold or greater change in level of expression (in either direction) in order to reject the Null hypothesis and accept the Alternative hypothesis. Interpretation of Pathway analysis data was considered significant based on both on a minimum of 2 genes in any given pathway and the pathway enrichment p-value ≤ 0.05. Correlation of age with vaccine outcome was interpreted by commonly accepted "rule of thumb" rules where correlation coefficients between 0.0 to 0.3 would be considered a negligible correlation; correlation coefficient between 0.3 to 0.5 would be considered a low correlation; correlation coefficient between 0.5 to 0.7 would be interpreted as a moderate correlation; 0.7 to 0.9 interpreted as a high correlation; and correlation coefficient between 0.9 to 1.0 would be interpreted as a very high correlation (Mukaka, 2012).

Threats to Validity

There were three factors that I determined could be perceived to be threats to internal validity in the study. First, the selection of the study subjects was limited to males only, restricted to ages between 18-48 years of age, and participants were not randomly assigned into either the treatment or control group; rather, the participants who received the vaccine were identified as being at risk of infection due to their work

environment(s). The restriction of study participants to males only was done as a way to control the extraneous variable of hormonal fluctuation.

The selection of only male volunteers to the study may be considered an external threat to population validity for the study. The general population is comprised of both males and females in roughly equal numbers, so the use of only male study volunteers make threaten the generalizability of the study results. Demographic information that was collected on the study participants showed that the original study populations were almost completely white, non-hispanic or latino; in addition to the wide range of ages of the volunteers, there was an assumption that the participants were represented the general population in terms of race and ethnicity which was not quite met. Each volunteer was treated the same with regard to the administration of the test materials (either as a mock-vaccination with saline or vaccination with one of the two live-attenuated vaccines); however, I determined that the repeated blood sample collection could have posed a threat to external validity through multiple-treatment interference. The threat was minimized because the control subjects also experienced the same repeated blood sample collection manipulations. Finally, vaccine failure in one of the VEEV-vaccinated test subjects resulted in that individual being eliminated from the study; the loss of the one volunteer may reduce the power of the statistical tests to detect significantly different transcripts in the VEEV-vaccinated group.

I incorporated measures in the data analysis to control threats to the validity of the statistical conclusions. Overall, microarray experiments are powered at a level that is seldom observed in other types of studies, with power calculations typically exceeding

90-95%, even in small sample sizes (e.g., less than 10 study subjects) (SAS, 2011). The novel secondary study as planned accounted for an adequate power to detect differences between the two study groups. I used the False Discovery Rate test as a multiple test correction with Step-up correction of p-values is to decrease Type I error. The heterogeneity of the human genome could be perceived as a variable that could contribute to Type II error, as could influences of nutrition, sleep habits, and fasting or non-fasting on blood sample collection days.

Ethical Procedures

The original studies were approved by the USAMRIID institutional review board and met all federal regulations described in 45 CFR 46 Subpart A entitled The Federal Policy for the Protection of Human Subjects, 21 CFR 56 on Institutional Review Boards, 21 CFR 50 on Informed Consent, and Health Insurance Portability and Accountability Act of 1996 (HIPAA). The Archival data used in the dissertation research, in the form of Affymetrix .CEL files, were provided by USAMRIID. Permission to access and use data as well as publish results relating to the human data sets from USAMRIID FY05-01 and FY06-17 was granted by United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Office of Human Use and Ethics (OHU&E), Human Use Committee (HUC) on 11 March 2016. The Exempt Determination Official (EDO) made the determination that the data did not meet the definition of “human subject” under 32 CFR 219.102 (f) (1) nor (f) (2). There were no ethical issues related to using de-identified archived data; the data, once de-identified, were no longer subject to human-use protections. Permission was granted to proceed with the study by the Walden University

institutional review board, under IRB approval # 08-23-16-0185842, granted on 23 August 2016.

Each original raw data sample file was identified solely by a volunteer identification number (VIN) that contained no personally-identifiable information; the key to the coded VIN numbers was retained by the original institutional study principal investigator (PI)I, but was not available nor conveyed as part of the access to the archival data. The data are free to be openly published in military or scientific journals.

The original study was conducted within the work environment of the dissertation student at USARMIID.

Summary

The dissertation research involved the use of previously collected, archived human vaccine data to conduct a novel secondary analysis of the molecular epidemiology of human response to vaccination. The data modeled a comparison between a live-attenuated bacterial vaccine and a live-attenuated viral vaccine, both of which can be used as surrogates for virulent infection. The analysis I conducted incorporated commonly used statistical measures that were consistent with other, similar human studies in terms of the procedures employed for background correction of the raw transcript data, normalization procedure, analysis of variance, and multiple test corrections (Erwin-Cohen et al. 2012, Paranavitana et al., 2008b, Paranavitana et al., 2014). The pathways analysis I employed leveraged new capabilities of the Partek Genomics Suite software; the software module was a visualization tool for pathway enrichment which utilized the KEGG database for human data and significance was

achieved by applying Fisher's Exact Test to transcripts that were differentially expressed and matched to known pathways (Partek, 2016).

In Chapter 4, I report the results of my data analysis. Results directly address the research questions, including which transcripts were significantly expressed in response to vaccination with Ft, with VEEV, or as a comparison between the between the two vaccines. I will also report the results of the tests for correlation of age with vaccine outcome and tests to measure the association of genetic makeup with vaccine outcome.

Chapter 4: Results

Brief Overview

In this chapter I discuss the management of archival datasets, , data analysis procedures are discussed, and I present the results and interpretation of the data analyses. The purpose of my research project is to quantitatively assess the association between gene expression over time in response to treatment (vaccination) with either Ft or VEEV, or to Ft in direct comparison to VEEV; to assess the correlation between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination or vaccine failure), and the association between HLA phenotype and vaccine response. My research is intended to add to the limited published information about the changes in gene expression that each vaccine elicits, and address the gaps in knowledge regarding the association of age and genetic makeup on vaccine outcome.

Research Questions and Hypotheses

Research Question 1. Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft or VEEV?

Null Hypothesis 1. There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), followed by assessment of fold-change (< 2.0 in either direction)] in the mean gene expression level for any given transcript in the vaccinated group (for Ft or VEEV) compared to the control group, or between vaccinated and control at any specific time point.

Alternative Hypothesis 1. There is a significant change [first assessed by Step Up-corrected p-value ($p<0.1$), followed by assessment of fold-change (> 2.0 in either

direction)] in the mean gene expression level for any given transcript in the vaccinated group (for Ft or VEEV) compared to the control group, or between vaccinated and control at any specific time point.

Research Question 2. Is there a significant association between changes in the level of gene transcription over the time course of vaccination with Ft in comparison to VEEV?

Null Hypothesis 2. There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), then assessed by fold-change in expression (< 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between the vaccinated groups at any specific time point.

Alternative Hypothesis 2. There is a significant change [first assessed by Step Up-corrected p-value ($p<0.1$), then assessed by fold-change in expression (> 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between the vaccinated groups at any specific time point.

Research Question 3. Is there a correlation between age of the vaccinated subject and vaccination outcome?

Null Hypothesis 3. There is no correlation between age and postvaccination titer.

Alternative Hypothesis 3. There is correlation between age and postvaccination titer.

Research Question 4. Is there a significant association between Human

Leukocyte Antigen (HLA) phenotype and vaccination outcome?

Null Hypothesis 4. There is no significant difference in postvaccination titer between HLA phenotypes.

Alternative Hypothesis 4. There is significant difference in postvaccination titer between HLA phenotypes.

Data Collection and Handling of Data Files

Archival data files for the secondary analyses were obtained following approval of my research proposal from the USAMRIID Human Use Committee and the Walden University institutional review board, under IRB approval number 08-23-16-0185842. The dissertation research project involved a secondary analysis combining two studies that were previously conducted at USAMRIID. Figure 1 illustrates the experimental design schema of the two original studies, which were conducted under GCP standards of quality.

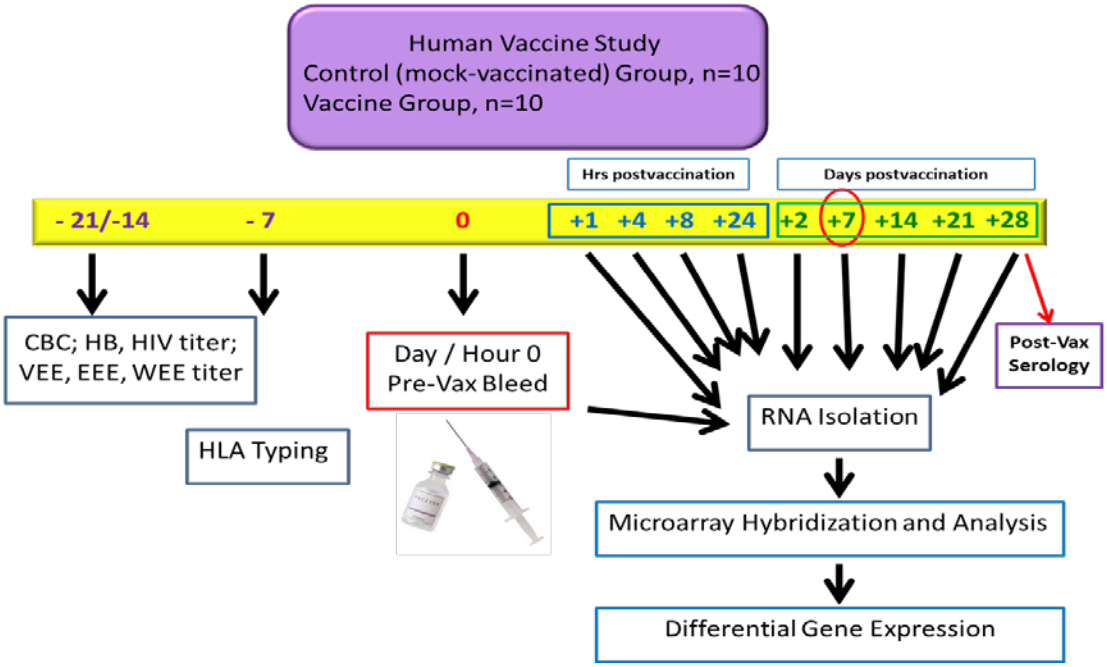


Figure 1. Schematic of original studies for Ft and VEEV vaccination.

Original study volunteers were screened for health and demonstrated to be naïve for previous exposure to either vaccine or disease. Blood sampling occurred prior to vaccination and then at 1, 4, and 8 hours following vaccination, and at days 1, 2, 7, 14, 21, and 28 (with the exception that there was no day 7 sample in the Ft study). At the final blood sample collection, serum was collected to conduct post-vaccination serology tests to assess vaccine response (neutralizing titer). Whole blood was used as a source of RNA and the RNA was subjected to microarray hybridization and analysis which resulted in differential gene expression data.

The resulting transcriptome data files were provided to me for my secondary analysis as Affymetrix .CEL files. Sample data file descriptions including de-identified sample names with demographic data, results of vaccine outcome, and results of HLA typing were provided in a separate Excel file. Transcriptome data were imported into

Partek Genomics Suite software (version 6.6) from 375 Affymetrix .CEL files and were subjected to background correction using a robust multi-array algorithm, quantile normalization, and median polish probe set summarization. The next figures visually illustrate the QC metrics I used to assess the fitness of the data.

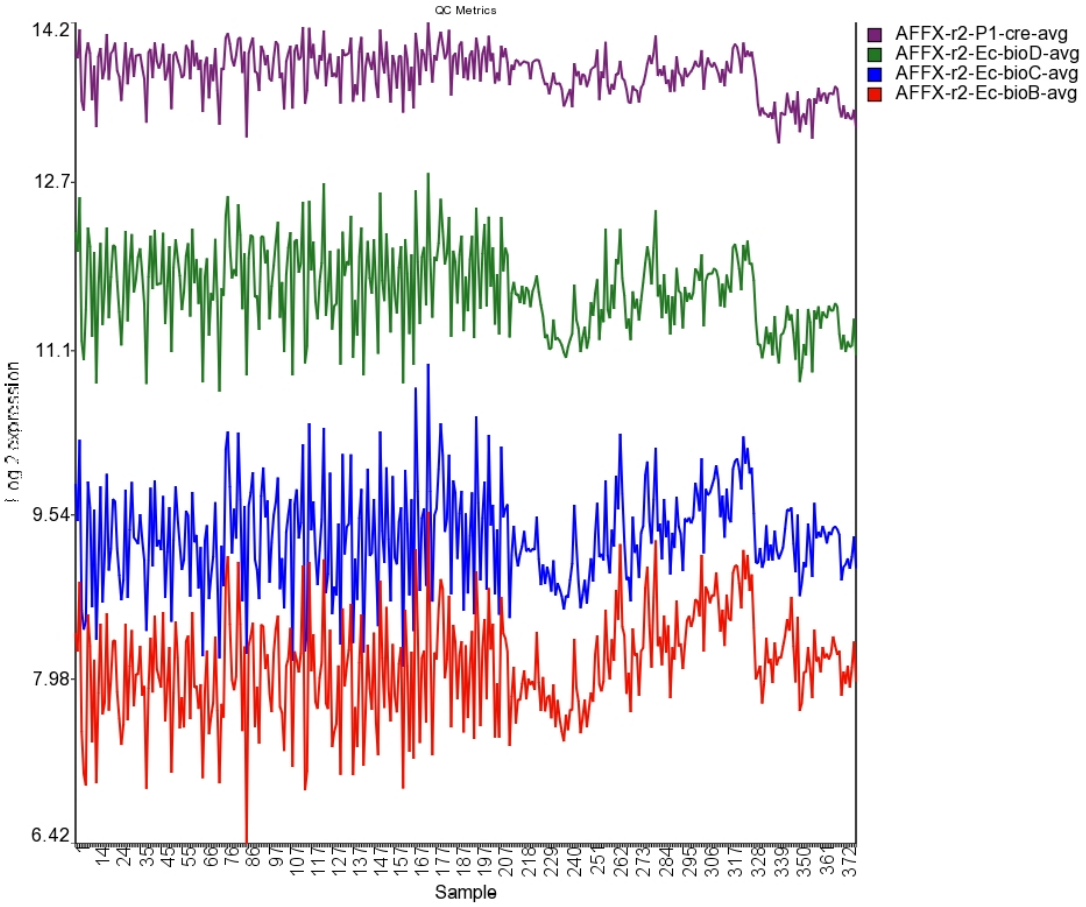


Figure 2. Log2 expression of microarray hybridization controls.

The hybridization control metrics are listed in the expected order from high to low. For the hybridization controls, AFFX-rs-P1-cre-avg should be higher than AFFX-r2-Ec-bioD-avg, which should be higher than AFFX-r2-Ec-bioC-avg, which should be

higher than AFFX-r2-Ec-bioB-avg, as indicated by the red arrow. Ft and VEEV samples are on X axis and the y-axis represents log2 values of the normalized spike in control probe-sets.

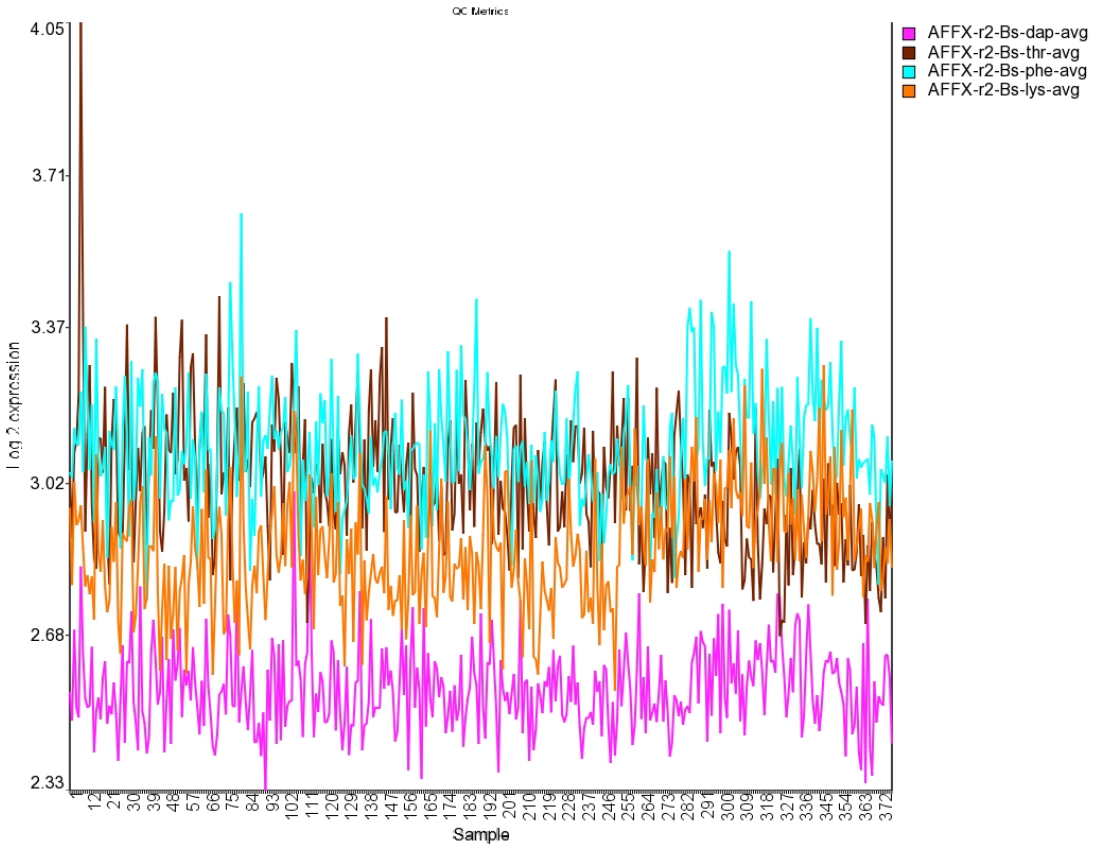


Figure 3. Log2 expression of microarray labeling controls.

The labeling control metrics are listed in the expected order from high to low. Four unlabeled polyA control spikes were inserted into the study samples prior to labeling and the resulting detection of the control spikes is dependent on the labeling reaction that labels the biological samples. The labeling controls are spiked in at increasing concentrations of Lys < Phe < Thr < Dap. The values are represented in log2

expression (y-axis) and the x-axis represents the study samples in the order that they appear on the post-import spreadsheet.

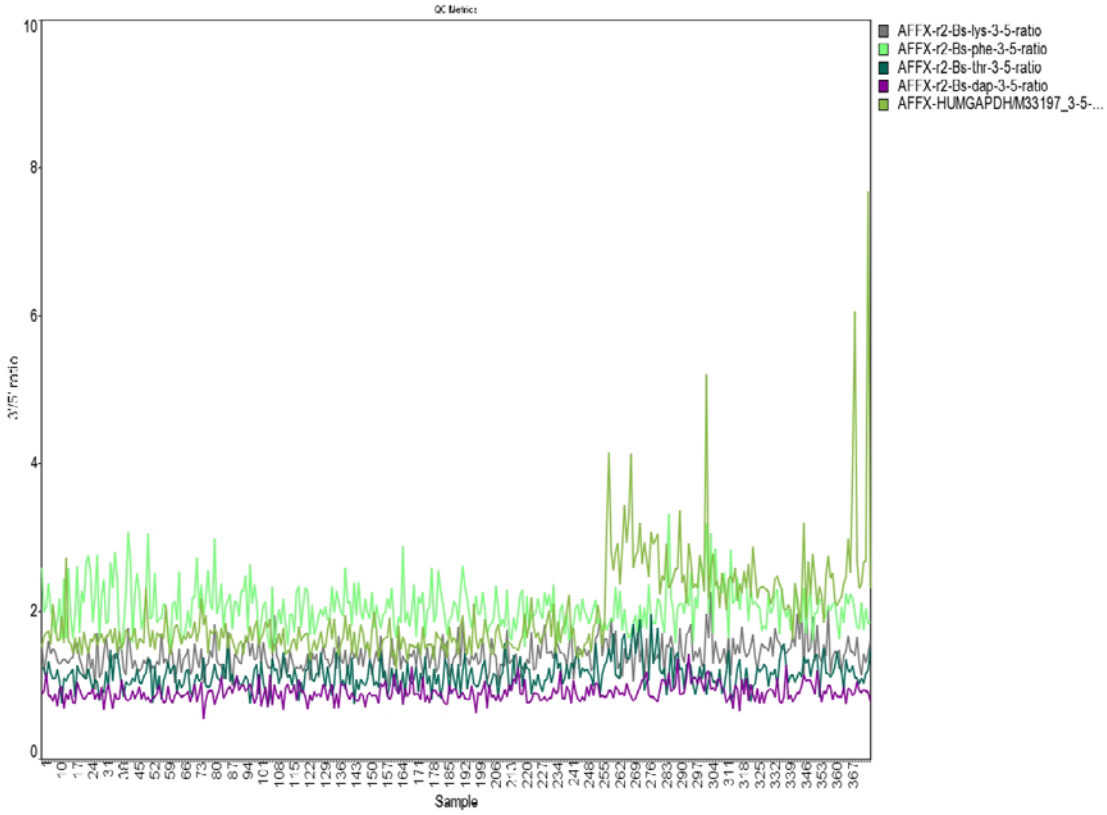


Figure 4. 3’/5’ *in vitro* transcription (IVT) ratio of microarray samples.

The reverse transcriptase reaction that transcribes RNA into cDNA should process from the 3’ through towards the 5’ end. The 3’ / 5’ ratio compares the abundance of the signal at the 3’ end over the abundance at the 5’ end, as a measure of the quality of the transcription reaction. A ratio of 3 or less is considered acceptable. It is apparent that there are several samples in which the 3’/5’ IVT ratio is far greater than the recommended cutoff of 3.

Post-import quality control (QC) of each data file was assessed by microarray chip hybridization, labeling, 3’/5’ in vitro transcription (IVT) ratio, as well as perfect

match and mismatch mean probe intensity values. I removed data files that were either missing or failed QC metrics prior to statistical analyses (listed in Tables 1 and 2).

Table 1

Ft Data Files Missing or Removed From Analysis

Subject ID	Sample time	Sample file removed	Reason for removal
Ft-0005	Day 28	FY05-01-0005-D28.CEL	Failed post-import QC, hybridization
Ft-0006	Day 28	No sample	Lost to follow-up
Ft-0025	1 Hr	No data file	Failed RNA integrity check
Ft-0026	Day 1	No data file	Failed RNA integrity check

For the Ft samples, from a total possible 180 samples, 4 were eliminated from the final analysis either due to loss of the study volunteer prior to that study collection point, poor quality isolated RNA, failure to hybridize properly.

Table 2

VEEV Data Files Missing or Removed From Analysis

Subject ID	Sample time	Sample file removed	Reason for removal
VEEV-0005	1 Hr	FY06-17-0005-D0H1.CEL	Failed 3'/5' IVT ratio metrics
VEEV-0006	1 Hr	FY06-17-0006-D0H1.CEL	Failed 3'/5' IVT ratio metrics
VEEV-0013	0 Hr	FY06-17-0013-D0H0.CEL	Non-responder
	1 Hr	FY06-17-0013-D0H1.CEL	Non-responder
	4 Hr	FY06-17-0013-D0H4.CEL	Non-responder
	8 Hr	FY06-17-0013-D0H8.CEL	Non-responder
	Day 1	FY06-17-0013-D1.CEL	Non-responder
	Day 2	FY06-17-0013-D2.CEL	Non-responder
	Day 7	FY06-17-0013-D7.CEL	Non-responder
	Day 14	FY06-17-0013-D14.CEL	Non-responder
	Day 21	FY06-17-0013-D21.CEL	Non-responder
	Day 28	FY06-17-0013-D28.CEL	Non-responder
VEEV-0021	4 Hr	FY06-17-0021-D0H4.CEL	Failed 3'/5' IVT ratio metrics
	8 Hr	FY06-17-0021-D0H8.CEL	Failed 3'/5' IVT ratio metrics
	Day 28	FY06-17-0021-D28.CEL	Failed 3'/5' IVT ratio metrics
VEEV-0027	0 Hr	FY06-17-0027-D0H0.CEL	Failed post-import QC, labeling
VEEV-0033	0 Hr	No data file	Failed RNA integrity check
VEEV-0034	0 Hr	No data file	Failed RNA integrity check

For the VEEV samples, from a total possible 200 samples, 18 were eliminated from the final analysis due to poor quality of isolated RNA, failure to label appropriately, failure of 3’/5’ IVT ratio, or in the case of VEEV study volunteer VEEV-0013, after the in-life portion of the study was completed, the study subject was found to be a primary vaccine failure and did not produce a neutralizing antibody titer.

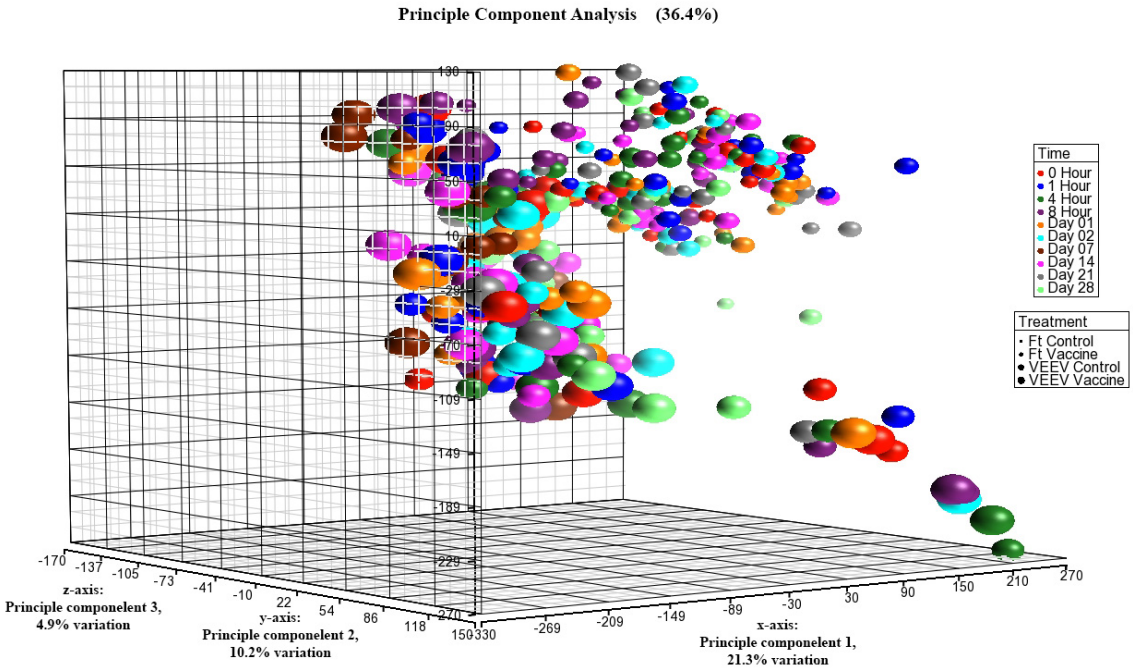


Figure 5. Principle component analysis.

Principle component analysis is a method to visualize highly dimensional data by reducing it to just three dimensions (x, y, and z axes) that can explain the variance in the data. The analyzed data are ordered by the amount of variance observed (x > y > z), the reduced variables are uncorrelated, and the reduced dimensionality of the data explain all of the variance in the dataset (Partek, 2016). From the combined Ft and VEEV datasets, the factor that is responsible for the greatest amount of variance in the data is that of treatment (Principle component 1, x-axis, represented by the size of the spheres). The next greatest source of variation is that of time (Principle component 2, y-axis, represented by the color of the spheres). The third axis represents the variance attributed to the interaction of time*treatment (Principle component 3, z-axis).

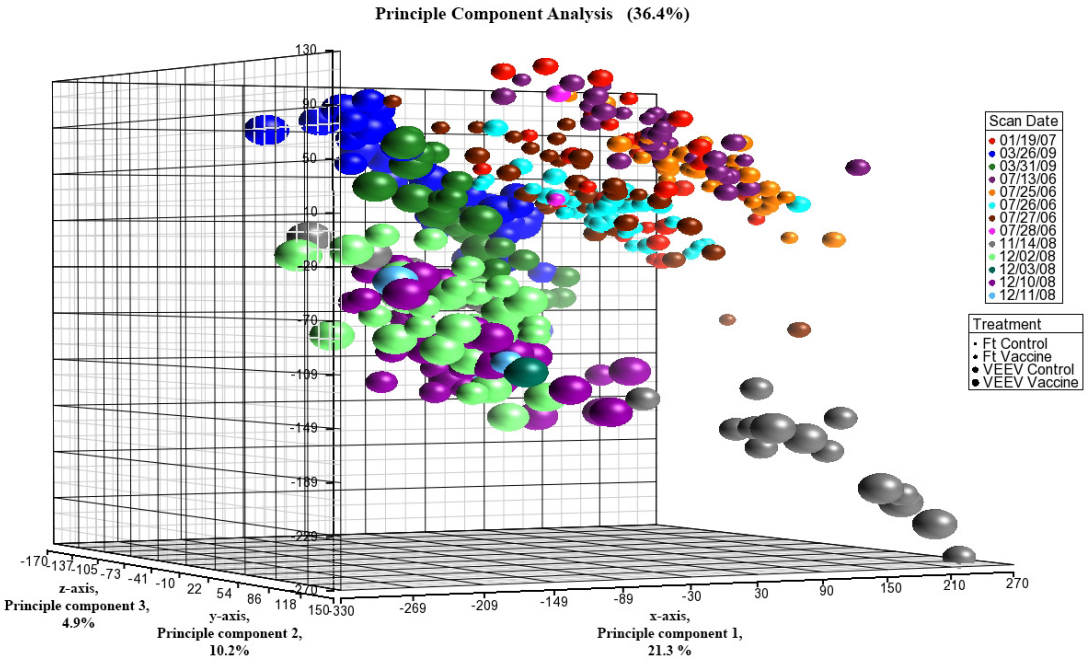


Figure 6. Principle component analysis, by scan date.

When principle component analysis was conducted and data were separated by scan date, it became apparent that the batch effects of RNA samples that were hybridized and read on different scan dates accounted for a great amount of the sample to sample variation.

Demographic and Descriptive Data

Each original study was planned to include twenty study volunteers; 10 pf which would be assigned to as mock-vaccination controls that would receive saline instead of the live-attenuated vaccine, and 10 of which would receive the live-attenuated vaccine. Tables 3 and 4 display demographic and descriptive characteristics of the de-identified subjects of each study.

Table 3

Descriptive and Demographic Characteristics of Ft Study Participants

Subject ID	Age	Gender	Race	Ethnicity	HLA DRB1	HLA DQB1	Log_{10} post-vaccination titer
Ft-0001	32	M	Caucasian	Non-Hispanic/Latino	11 / 04	03 / 03	1.9
Ft-0002	24	M	Caucasian	Non-Hispanic/Latino	01 / 04	03 / 05	0
Ft-0003	29	M	Caucasian	Non-Hispanic/Latino	04 / 11	03 / 03	2.2
Ft-0004	24	M	Caucasian	Non-Hispanic/Latino	03 / 13	02 / 06	0
Ft-0005	37	M	Caucasian	Non-Hispanic/Latino	11 / 13	03 / 06	3.11
Ft-0006	25	M	Other	Hispanic	04 / 14	03 / 05	0
Ft-0007	24	M	Caucasian	Non-Hispanic/Latino	04 / 07	02 / 03	1.6
Ft-0009	31	M	Caucasian	Non-Hispanic/Latino	04 / 15	03 / 06	2.51
Ft-0010	39	M	Caucasian	Non-Hispanic/Latino	11 / 13	03 / 06	0
Ft-0011	39	M	Caucasian	Non-Hispanic/Latino	14 / 15	05 / 03	1.9
Ft-0013	26	M	Caucasian	Non-Hispanic/Latino	11 / 15	03 / 06	3.11
Ft-0014	33	M	Caucasian	Non-Hispanic/Latino	03 / 15	02 /06	0
Ft-0016	40	M	Caucasian	Non-Hispanic/Latino	13 / 15	06 / 06	2.2
Ft-0017	39	M	Caucasian	Non-Hispanic/Latino	07 / 10	02 / 05	0
Ft-0018	44	M	Caucasian	Non-Hispanic/Latino	01 / 07	03 / 05	0
Ft-0019	21	M	Caucasian	Non-Hispanic/Latino	04 / 07	03 / 03	0
Ft-0020	43	M	Caucasian	Non-Hispanic/Latino	07 / 13	02 / 06	0
Ft-0021	40	M	Caucasian	Non-Hispanic/Latino	03 / 15	02 / 06	2.2
Ft-0025	25	M	Caucasian	Non-Hispanic/Latino	13 / 15	05 / 06	0
Ft-0026	31	M	Caucasian	Non-Hispanic/Latino	10 / 15	05 / 06	2.2

Table 4

Descriptive and Demographic Characteristics of VEEV study participants

Subject ID	Age	Gender	Race	Ethnicity	HLA DRB1	HLA DQB1	Log_{10} post-vaccination titer
VEEV-0005	30	M	Caucasian	Non-Hispanic/Latino	04 / 15	03 / 06	0
VEEV-0006	31	M	Caucasian	Non-Hispanic/Latino	04 / 07	02 / 03	2.51
VEEV-0011	24	M	Caucasian	Non-Hispanic/Latino	04 / 07	03 / 02	0
VEEV-0012	38	M	Caucasian	Non-Hispanic/Latino	07 / 15	02 / 06	2.51
VEEV-0013	24	M	Caucasian	Non-Hispanic/Latino	11 / 13	03 / 06	0
VEEV-0015	42	M	Caucasian	Non-Hispanic/Latino	03 / 07	02 / 02	0
VEEV-0016	25	M	Caucasian	Non-Hispanic/Latino	03 / 04	02 / 03	1.6
VEEV-0017	25	M	Caucasian	Non-Hispanic/Latino	07 / 13	02 / 03	0
VEEV-0018	46	M	Caucasian	Non-Hispanic/Latino	07 / 14	03 / 05	3.11
VEEV-0019	46	M	Caucasian	Non-Hispanic/Latino	01 / 07	05 / 03	0
VEEV-0020	24	M	Caucasian	Non-Hispanic/Latino	08 / 15	04 / 06	2.2
VEEV-0021	23	M	Caucasian	Non-Hispanic/Latino	03 / 16	02 / 05	0
VEEV-0027	25	M	African-American	Non-Hispanic/Latino	04 / 07	03 / 02	0
VEEV-0028	26	M	Caucasian	Non-Hispanic/Latino	04 / 07	03 / 02	1.9
VEEV-0029	48	M	Caucasian	Non-Hispanic/Latino	15 / 15	06 / 06	3.11
VEEV-0030	46	M	Caucasian	Non-Hispanic/Latino	01 / 07	05 / 03	0
VEEV-0031	37	M	Caucasian	Non-Hispanic/Latino	04 / 07	03 / 02	1.3
VEEV-0032	39	M	Caucasian	Non-Hispanic/Latino	04 / 07	03 / 02	0
VEEV-0033	32	M	Caucasian	Non-Hispanic/Latino	11 / 13	03 / 06	2.2
VEEV-0034	25	M	African-American	Non-Hispanic/Latino	03 / 15	04 / 06	0

The study volunteers recruited to the two original studies represented a small sampling of the general population. The studies were designed to only include males for the transcriptome analysis; females were excluded from both original studies due to concerns regarding the influence of estrogen on the immune response and the small sample size precluding adequate controls for that variable. From the results of the demographic tables, I determined that the sample population was predominantly of Caucasian race and of non-Hispanic or non-Latino ethnicity and concluded that no further demographic analysis of these factors could be made in association with vaccine outcome. The ages ranged from 24 to 48 in the Ft study and from 23 to 48 in the VEEV study; as such the age range is representative of adults, and does not include representation for either children or elderly subjects. The relatively narrow demographic characteristics of the study population may limit generalization of the study results; however, the subjects are typical for age and sex of military personnel.

The results of HLA typing were listed for each individual in each vaccine study. Typically HLA is represented by a four-digit code (e.g., 0603); however, the data I have presented have been condensed to represent only major groupings of HLA alleles due to the small overall sample size of the studies. The post-vaccination titers are presented as Log_{10} titer. Individuals who did not mount an immune response were part of the control, mock-vaccinated groups with the exception of subject VEEV-0013, who failed to produce neutralizing antibody after vaccination and was deemed to be a primary vaccine failure (all data from this individual were then removed from my research study).

Results

Descriptive Statistics for Transcriptome Data

In the Partek Genomics Suite statistical analysis program, I grouped the data by subject, time point, or treatment condition and performed descriptive statistics to assess the mean gene expression, standard deviation, median, minimum, maximum, 25th percentile, 75th percentile, and interquartile range (Q3-Q1) for the 54,675 probesets represented on each Affymetrix HG U133 2.0 array chip (Affymetrix, 2004), expressed as Log2 values. These data are represented in the following tables for the Ft study organized by subject (Table 5), the VEEV study organized by subject (Table 6), as well as the descriptive statistics for both Ft and VEEV studies combined, organized by the independent variables of time (Table 7) or treatment (Table 8).

Table 5

Descriptive Statistics for Ft Study: Log2 Transcript Expression Values by Study Subject Over All Time Points

Subject ID	Treatment	N	Mean	Std. Dev.	Median	Min	Max	Q1	Q3	IQR
Ft-0001	Ft Vaccine	9	4.6591	1.7411	4.2713	2.0692	14.8570	3.3080	5.6215	2.3134
Ft-0002	Ft Control	9	4.6584	1.6878	4.2915	2.1357	14.9013	3.3490	5.6045	2.2555
Ft-0003	Ft Vaccine	9	4.6471	1.6529	4.3097	2.0995	14.8913	3.3618	5.5861	2.2243
Ft-0004	Ft Control	9	4.6689	1.7620	4.2679	2.1054	14.8892	3.3037	5.6335	2.3299
Ft-0005	Ft Vaccine	8	4.6448	1.6662	4.2989	2.1515	14.9260	3.3508	5.5838	2.2330
Ft-0006	Ft Control	8	4.6525	1.7198	4.2810	2.1044	14.8892	3.3145	5.6048	2.2903
Ft-0007	Ft Vaccine	9	4.6468	1.6667	4.3109	2.1274	14.8938	3.3599	5.5677	2.2078
Ft-0009	Ft Vaccine	9	4.6579	1.6738	4.3119	2.1408	14.8651	3.3571	5.6027	2.2456
Ft-0010	Ft Control	9	4.6489	1.6737	4.3060	2.1188	14.8625	3.3546	5.5729	2.2183
Ft-0011	Ft Vaccine	9	4.6558	1.7136	4.2850	2.1055	14.8950	3.3246	5.6067	2.2822
Ft-0013	Ft Vaccine	9	4.6544	1.6770	4.3104	2.1238	14.9033	3.3575	5.5805	2.2229
Ft-0014	Ft Control	9	4.6544	1.6937	4.3020	2.1218	14.8902	3.3526	5.5699	2.2173
Ft-0016	Ft Vaccine	9	4.6545	1.6795	4.2963	2.1391	14.8804	3.3519	5.5906	2.2387
Ft-0017	Ft Control	9	4.6600	1.6474	4.3342	2.1218	14.9003	3.3844	5.5866	2.2023
Ft-0018	Ft Control	9	4.6645	1.6855	4.3025	2.1464	14.8855	3.3543	5.6176	2.2633
Ft-0019	Ft Control	9	4.6570	1.7414	4.2526	2.1322	14.8871	3.3095	5.6072	2.2977
Ft-0020	Ft Control	9	4.6376	1.6602	4.3016	2.1201	14.9246	3.3497	5.5772	2.2275
Ft-0021	Ft Vaccine	9	4.6414	1.6715	4.3005	2.1021	14.9275	3.3459	5.5806	2.2346
Ft-0025	Ft Control	8	4.6637	1.7274	4.2823	2.1116	14.9211	3.3261	5.6124	2.2864
Ft-0026	Ft Vaccine	8	4.6662	1.7368	4.2784	2.1194	14.9056	3.3222	5.6144	2.2922

N = number of samples; SD = standard deviation; MIN = smallest observation; MAX = largest observation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = inter-quartile range = Q3-Q1.

Table 6

Descriptive Statistics for VEEV Study: Log2 Transcript Expression Values by Study Subject

Subject ID	N	Treatment	Mean	Std. Dev.	Median	Min	Max	Q1	Q3	IQR
VEEV-0005	9	VEEV Control	4.6044	1.6364	4.2562	2.1222	14.9999	3.3325	5.5221	2.1896
VEEV-0006	9	VEEV Vaccine	4.5992	1.6360	4.2454	2.1189	15.0134	3.3321	5.5115	2.1794
VEEV-0011	10	VEEV Control	4.6307	1.6811	4.2648	2.1209	14.9804	3.3209	5.5729	2.2520
VEEV-0012	10	VEEV Vaccine	4.6866	1.7454	4.2688	2.1510	14.8982	3.3278	5.6466	2.3188
VEEV-0015	10	VEEV Control	4.6847	1.7621	4.2451	2.1130	14.8953	3.3230	5.6335	2.3104
VEEV-0016	10	VEEV Vaccine	4.6834	1.7308	4.2571	2.1180	14.8896	3.3479	5.6198	2.2718
VEEV-0017	10	VEEV Control	4.6786	1.7239	4.2777	2.1139	14.9322	3.3321	5.6447	2.3127
VEEV-0018	10	VEEV Vaccine	4.6717	1.6665	4.3132	2.1384	14.9117	3.3697	5.6196	2.2500
VEEV-0019	10	VEEV Control	4.6712	1.6727	4.3036	2.1506	14.9136	3.3689	5.6149	2.2459
VEEV-0020	10	VEEV Vaccine	4.6736	1.7331	4.2526	2.1175	14.9143	3.3297	5.6277	2.2980
VEEV-0021	7	VEEV Control	4.6797	1.7673	4.2431	2.1372	14.9128	3.3020	5.6590	2.3570
VEEV-0027	9	VEEV Control	4.6777	1.7931	4.2418	2.1276	14.9005	3.2837	5.6634	2.3797
VEEV-0028	10	VEEV Vaccine	4.6712	1.7114	4.2983	2.1271	14.9102	3.3360	5.6430	2.3070
VEEV-0029	10	VEEV Vaccine	4.6764	1.7506	4.2665	2.1242	14.8902	3.3099	5.6601	2.3502
VEEV-0030	10	VEEV Control	4.6725	1.6756	4.3145	2.1559	14.8972	3.3625	5.6280	2.2655
VEEV-0031	10	VEEV Vaccine	4.6829	1.7870	4.2457	2.1034	14.8646	3.2944	5.6650	2.3705
VEEV-0032	10	VEEV Control	4.6697	1.6958	4.2996	2.1339	14.9247	3.3420	5.6425	2.3005
VEEV-0033	9	VEEV Vaccine	4.6756	1.7345	4.2826	2.1137	14.9059	3.3162	5.6713	2.3552
VEEV-0034	9	VEEV Control	4.6744	1.7323	4.2844	2.1147	14.9069	3.3229	5.6479	2.3250

N = number of samples; SD = standard deviation; MIN = smallest observation; MAX = largest observation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = inter-quartile range = Q3-Q1.

The descriptive statistics for both the Ft and VEEV studies (Tables 5 and 6, respectively) present the Log2-transformed expression values for 54,675 probes, parsed by subject ID. All of the data were transformed as part of the importation and normalization process that I conducted, which is reflected in the largely uniform expression values between subjects. The post-transformation descriptive statistics are indicative that any potentially skewed original raw data were reduced to approximate normality by the Log2-transformation process.

Table 7

Descriptive Statistics: Log2 Expression Values by Time Point (Includes Both Ft and VEEV) for All Subjects

Time	N	Mean	Std. Dev.	Median	Min	Max	Q1	Q3	IQR
0 Hour	37	4.6549	1.6980	4.2842	2.1166	14.9101	3.3355	5.6021	2.2666
1 Hour	38	4.6600	1.7122	4.2775	2.1210	14.8936	3.3296	5.6121	2.2824
4 Hour	39	4.6555	1.6934	4.2879	2.1286	14.9135	3.3406	5.6026	2.2619
8 Hour	39	4.6621	1.7269	4.2696	2.1243	14.9069	3.3230	5.6170	2.2940
Day 1	38	4.6630	1.7043	4.2854	2.1285	14.9046	3.3382	5.6157	2.2775
Day 2	39	4.6596	1.7030	4.2852	2.1201	14.9109	3.3367	5.6079	2.2712
Day 7	20	4.6804	1.7465	4.2624	2.1247	14.8881	3.3193	5.6531	2.3339
Day 14	40	4.6645	1.7095	4.2842	2.1211	14.9008	3.3369	5.6174	2.2804
Day 21	40	4.6616	1.6997	4.2897	2.1283	14.9154	3.3398	5.6122	2.2724
Day 28	37	4.6613	1.6803	4.3010	2.1262	14.9116	3.3550	5.5997	2.2447

N = number of samples; SD = standard deviation; MIN = smallest observation; MAX = largest observation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = inter-quartile range = Q3-Q1.

Table 8

Descriptive Statistics: Log2 Expression Values by Treatment (Includes Both Ft and VEEV)

Treatment	N	Mean	Std. Dev.	Median	Min	Max	Q1	Q3	IQR
Ft Control	89	4.6566	1.6994	4.2924	2.1221	14.8949	3.3403	5.5984	2.2582
Ft Vaccine	88	4.6527	1.6876	4.2975	2.1174	14.8940	3.3441	5.5933	2.2492
VEEV Control	94	4.6643	1.7121	4.2744	2.1290	14.9265	3.3304	5.6221	2.2917
VEEV Vaccine	97	4.6709	1.7210	4.2722	2.1266	14.9088	3.3309	5.6318	2.3009

N = number of samples; SD = standard deviation; MIN = smallest observation; MAX = largest observation; Q1 = 25th percentile; Q3 = 75th percentile; IQR = inter-quartile range = Q3-Q1.

In the descriptive statistics for both the Ft and VEEV studies, I present the Log2-transformed expression values for 54,675 expression probes, parsed by time (Table 7), or by treatment (Table 8). As with the descriptive results organized by subject, the results organized by time or treatment also demonstrate means that are highly comparable.

Inferential Statistics

The hypotheses for Research Questions 1 and 2 of the research study were assessed by mixed model ANOVA followed by False Discovery Rate (FDR) test as a multiple test correction.

Analysis of Variance

Data that passed QC following importation and normalization were subjected to mixed model ANOVA to address research questions regarding association of time or treatment with level of gene expression. The ANOVA mixed model is:

$$Y_{ijklm} = \mu + \text{scan date}_{i} + \text{time point}_{j} + \text{treatment}_{k} + \text{subject (treatment)}_{kl} + \text{time point} * \text{treatment}_{jk} + \varepsilon_{ijklm}$$

Where Y_{ijklm} represents the m^{th} observation of gene expression on the i^{th} scan date j^{th} time point k^{th} treatment l^{th} subject; μ is the common effect for the whole experiment. ε_{ijklm} represents the random error present in the m^{th} observation of gene expression on the i^{th} scan date j^{th} time point k^{th} treatment l^{th} subject. The errors ε_{ijklm} are assumed to be normally and independently distributed with mean 0 and standard deviation δ for all measurements (Partek, 2016). Scan date and subject are random effects. The mixed model also considers the effect of within-subject variation. The restricted maximum likelihood (REML) approach for estimation of variance components was chosen over the alternative, Method of Moments (MoM), because the data sets were not balanced (Partek, 2016).

Following the ANOVA analysis, I subjected the data to multiple test correction using the Step-Up method to correct p-values as part of a False Discovery Report (FDR) set to limit false discoveries to 0.1.

The purpose of conducting a multiple test correction was to control Type I errors. The FDR is the proportion of false positives among all positive results. As a multiple test correction protocol, the Step-up method described by Benjamini and Höchberg (1995) is less restrictive than the Bonferroni correction method, which is again less restrictive than using a Family Wise Error Rate (FWER) multiple test correction protocol (Partek, 2016), but any of these methods is more restrictive than using an uncorrected p-value. The Step-up method was chosen for the FDR to generate a corrected p-value which was used to assess significance; the corrected p-value is the smallest overall significance level at which a particular hypothesis would be rejected. Briefly, the Step-up correction is

calculated by sorting all of the p-values in ascending order (stepping up); there are *n* p-values sorted by ascending order, where *m* represents the rank of each p-value. The Step-up calculation compares p-value*(*n*/*m*) with the specified alpha level (which in this case was set to 0.1), and the cut-off p-value is the one that generates the last product that is less than the alpha level (Partek, 2016).

Following the multiple test correction, significant transcripts were assessed for fold-change in expression. Fold change in transcript expression is calculated by Partek Genomics Suite; briefly, the program calculates the least squares mean (LS Mean) as the linear sum of the estimated means from the ANOVA model. Fold change for each contrast (e.g., Treatment at 1 hour vs. Control at 1 hour) is calculated such that:

$$\frac{\text{LS Mean (Treatment at 1 hour)}}{\text{LS Mean (Control at 1 hour)}}$$

Table 9

False Discovery Rate Report: ANOVA Analysis Contrasting Ft Vaccinated with Ft Control Subjects

Variable	*Cutoff p-value**	*# Transcripts***	*# Transcripts (± 2-Fold change in expression)*
Time	2.65E-02	14468	
Treatment	3.65E-02	19958	
Time * Treatment	4.99E-03	2730	
0 Hr Vax vs Ctrl	1.83E-06	0	
1 Hr Vax vs Ctrl	1.83E-06	0	
4 Hr Vax vs Ctrl	1.83E-06	0	
8 Hr Vax vs Ctrl	3.66E-06	2	0
Day 1 Vax vs Ctrl	1.83E-06	0	
Day 2 Vax vs Ctrl	2.93E-05	16	5
Day 14 Vax vs Ctrl	3.66E-06	2	0
Day 21 Vax vs Ctrl	1.83E-06	0	
Day 28 Vax vs Ctrl	1.83E-06	0	

*Significance Level: 0.1
**Multiple Test Correction Method: Step-Up

The combined results from my analysis using the mixed model ANOVA and FDR multiple test correction of the Ft data set provide evidence to address the first part of Research Question 1, Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft? The cut-off p-values generated by the mixed model ANOVA and subsequent FDR analysis range from 2.65E-02 to 3.66E-06. Based on cut-off p-value alone, there were 14,468 transcripts that were differentially expressed over time (for all treatments combined). In response to treatment (for all time points) there were 19,958 differentially expressed transcripts. When effect time and treatment were considered as a single variable, 2,730 transcripts were identified as being differentially expressed. The list of significant transcripts was further

constrained by fold-expression level. Fold expression is determined using Fisher's least significant difference (LSD) and is expressed as the ratio of least squares means comparing vaccinated versus control. When fold-change in expression level was also added as criteria to evaluate the hypothesis for research question 1, then the number of transcripts which were both significantly differentially expressed and which met a minimum fold-change of ±2 fold in either direction, the number of transcripts was drastically reduced for the Ft data set (Table 9). The analysis results were used to determine whether to accept the null hypothesis or reject in favor of the alternate hypothesis as follows:

H_0: There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), then assessed by fold-change in expression (< 2.0 in either direction)] in the mean gene expression level for any given transcript in the Ft vaccinated group compared to the control group, or between vaccinated and control at any specific time point.

H_a: There is significant change [first assessed by Step Up-corrected p-value ($p<0.1$), then assessed by fold-change in expression ($\geq$ 2.0 in either direction)] in the mean gene expression level for any given transcript in the Ft vaccinated group compared to the control group, or between vaccinated and control in at least one time point.

When all determining criteria are assessed, the null hypothesis must be rejected in favor of the alternate hypothesis, that there is a significant change in gene expression following Ft vaccination in comparison to the unvaccinated control subjects, due to the observation of five transcripts at Day 2 which met both criteria of $p\leq$ 2.93E-05 and ± 2-fold change in transcript expression.

Table 10

False Discovery Rate Report: ANOVA Analysis Contrasting VEEV Vaccinated with VEEV Control Subjects

Variable	*Cutoff p-value**	*# Transcripts***	*# Transcripts (± 2-Fold change in expression)*
Time	2.65E-02	14468	
Treatment	3.65E-01	19958	
Time * Treatment	4.99E-03	2730	
0 Hr Vax vs Ctrl	1.82E-06	0	
1 Hr Vax vs Ctrl	1.82E-06	0	
4 Hr Vax vs Ctrl	1.82E-06	0	
8 Hr Vax vs Ctrl	1.82E-06	0	
Day 1 Vax vs Ctrl	1.82E-06	0	
Day 2 Vax vs Ctrl	6.27E-03	3429	227
Day 7 Vax vs Ctrl	1.23E-03	673	140
Day 14 Vax vs Ctrl	2.83E-02	15482	326
Day 21 Vax vs Ctrl	1.82E-06	0	
Day 28 Vax vs Ctrl	1.82E-06	0	

*Significance Level: 0.1
**Multiple Test Correction Method: Step-Up

The resulting ANOVA and FDR analyses of the VEEV data set provide evidence to address the second part of Research Question 1, Is there a significant association between changes in the level of gene transcription and the time course of vaccination with VEEV? The cut-off p-values generated by mixed model ANOVA and subsequent FDR analysis range from 3.65E-01 to 1.82E-06. Based on cut-off p-value alone, there were 14,468 transcripts that were differentially expressed over time (for all treatments combined). In response to treatment (for all time points) there were 19,958 differentially expressed transcripts. When effect time and treatment were considered as a single variable, 2,730 transcripts were identified as being differentially expressed. When fold-

change in expression level was also added as determining criteria, then the number of transcripts which were both significantly differentially expressed and which met a minimum fold-change of ±2 fold in either direction, the number of transcripts was reduced for the VEEV data set (Table 10). The analysis results were used to determine whether to accept the null hypothesis or reject in favor of the alternate hypothesis as follows:

H_0: There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), then assessed by fold-change in expression (< 2.0 in either direction)] in the mean gene expression level for any given transcript in the VEEV vaccinated group compared to the control group, or between vaccinated and control at any specific time point.

H_a: There is significant change [first assessed by Step Up-corrected p-value ($p<0.1$), then assessed by fold-change in expression (≥ 2.0 in either direction)] in the mean gene expression level for any given transcript in the VEEV vaccinated group compared to the control group, or between vaccinated and control in at least one time point.

When all determining criteria are assessed, the null hypothesis must be rejected in favor of the alternate hypothesis, that there is a significant change in gene expression following VEEV vaccination in comparison to the unvaccinated control subjects, due to the observation of numerous transcripts which met the criteria on Day 2 of $p\leq 6.27E\text{-}03$ and ± 2-fold change in transcript expression, on Day 7 $p\leq 1.23E\text{-}03$ and ± 2-fold change

in transcript expression, and on Day 14 p≤ 2.83E-02 and ± 2-fold change in transcript expression.

Table 11

False Discovery Rate Report: ANOVA Analysis Contrasting Ft Vaccinated against VEEV Vaccinated Subjects

Variable	*Cutoff p-value**	*# Transcripts***	*# Transcripts (± 2-Fold change in expression)*
Time	2.65E-02	14468	
Treatment	3.65E-02	19958	
Time * Treatment	4.99E-03	2730	
0 Hr Ft vs VEEV	2.67E-02	14593	361
1 Hr Ft vs VEEV	4.12E-02	22527	497
4 Hr Ft vs VEEV	2.67E-02	14612	393
8 Hr Ft vs VEEV	2.16E-02	11795	310
Day 1 Ft vs VEEV	4.34E-02	23724	527
Day 2 Ft vs VEEV	3.05E-02	16655	360
Day 14 Ft vs VEEV	3.65E-02	19954	609
Day 21 Ft vs VEEV	3.32E-02	18153	355
Day 28 Ft vs VEEV	3.24E-02	17738	411

*Significance Level: 0.1
**Multiple Test Correction Method: Step-Up

The resulting ANOVA and FDR analyses of the Ft data set in comparison to the VEEV data set provide evidence to address Research Question 2, Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft in comparison to that of VEEV? The cut-off p-values generated by mixed model ANOVA and subsequent FDR analysis range from 2.16E-02 to 4.99E-03. Based on cut-off p-value alone, there were 14,468 transcripts that were differentially expressed over time (for all treatments combined). In response to treatment (for all time

points) there were 19,958 differentially expressed transcripts. When effect time and treatment were considered as a single variable, 2,730 transcripts were identified as being differentially expressed. When fold-change in expression level was also added as determining criteria, then the number of transcripts which were both significantly differentially expressed and which met a minimum fold-change of ±2 fold in either direction, the number of transcripts was reduced for the comparison of FT against VEEV data (Table 11). The analysis results were used to determine whether to accept the null hypothesis or reject in favor of the alternate hypothesis as follows:

H_0: There is no significant change [first assessed by Step Up-corrected p-value ($p>0.1$), then assessed by fold-change in expression (< 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between vaccinated and control at any specific time point.

H_a: There is significant change [first assessed by Step Up-corrected p-value ($p<0.1$), then assessed by fold-change in expression (≥ 2.0 in either direction)] in the mean gene expression level for any given transcript when comparing the Ft vaccinated group and the VEEV vaccinated group, or between vaccinated and control in at least one time point.

When all determining criteria are assessed, the null hypothesis must be rejected in favor of the alternate hypothesis, that there is a significant change in gene expression when comparing the Ft vaccinated group and the VEEV vaccinated group, due to the observation of numerous significant transcripts that met or exceeded the cut-off p-value

at each time point as well as had ± 2-fold change in expression in either direction (Table 11).

Correlational Analysis of Relationship Between Age and Neutralizing Titer

I conducted a correlation analysis to measure the strength of relationship between the age of the vaccinated subjects and the neutralizing antibody titer produced by successful primary vaccination. Due to the small sample size for each vaccination group (n=10 for Ft; n=9 for VEEV) non-parametric Spearman rank correlation was used to measure the relationship. Analyses were conducted first between the groups with participant ages reduced to age category, where age 21-30 was classified as group 1, and participants with ages 31-50 were classified as group 2. An additional ad hoc analysis was conducted with age as reported.

To address Research Question 3, I stated in the null hypothesis that there is no correlation between age and postvaccination titer. In the alternate hypothesis I stated that there is correlation between age and postvaccination titer.

Correlation of age group with post-vaccination titer: Ft study. For the comparison of age of Ft study subjects, the two-tailed test of significance indicated no relationship between age group of the Ft subjects and the Log10 post-vaccination titer ($r = 0.0395$, $p = 0.967$). The data are illustrated in Figure 7. Based on the results of the Spearman rank correlation, there is only a negligible correlation and the null hypothesis that there is no relationship between the two variables is accepted.

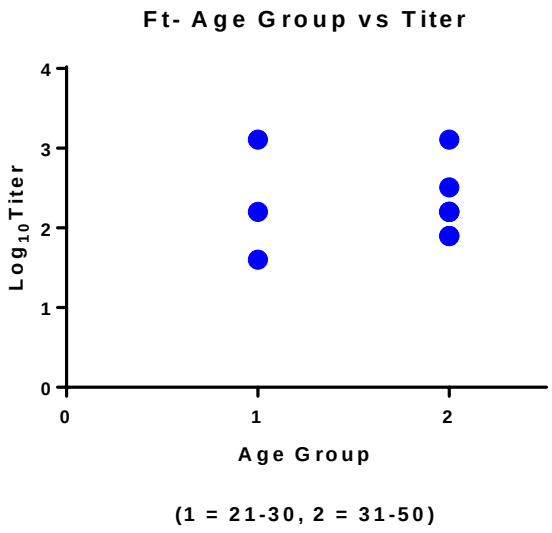


Figure 7. Correlation between age group and titer in Ft study.

For the comparison of age of Ft study subjects, the two-tailed test of significance indicated no relationship between age of the Ft subjects and the Log10 post-vaccination titer ($r = 0.0095$, $p = 0.984$). The data are illustrated in Figure 8. Based on the results of the Spearman rank correlation, there is only a negligible correlation and the null hypothesis that there is no relationship between the two is accepted.

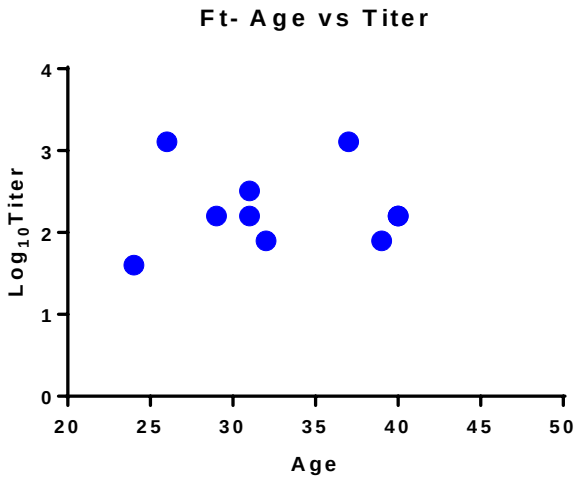


Figure 8. Correlation between Age and Titer in Ft study.

For the comparison of age of VEEV study subjects, the two-tailed test of significance indicated no relationship between age group of the VEEV subjects and the Log10 post-vaccination titer ($r = 0.5085$, $p = 0.179$). The data are illustrated in Figure 9. Based on the results of the Spearman rank correlation, there is a moderately positive relationship between age group of the VEEV vaccinated individuals and post-vaccination titer; however, the results are not statistically significant. The null hypothesis is accepted.

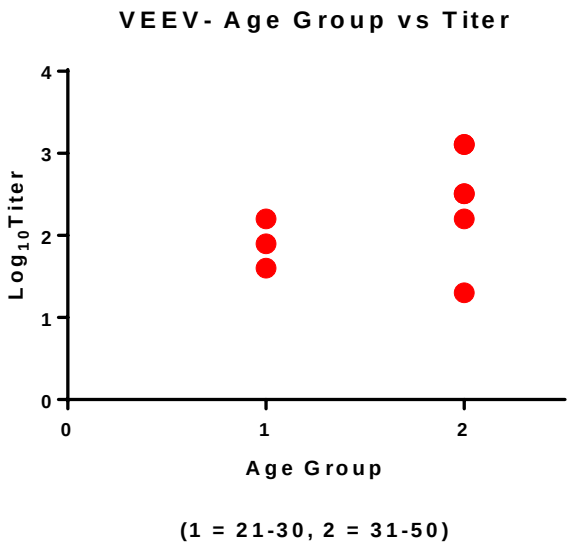


Figure 9. Correlation between age group and titer in VEEV study.

For the comparison of age of VEEV study subjects, the two-tailed test of significance indicated no relationship between age of the VEEV subjects and the Log10 post-vaccination titer ($r = 0.6245$, $p = 0.080$). The data are illustrated in Figure 10. Based on the results of the Spearman rank correlation, there is a moderately positive relationship between the age of the VEEV vaccinated individuals and post-vaccination titer; however, the results are not statistically significant. The null hypothesis is accepted.

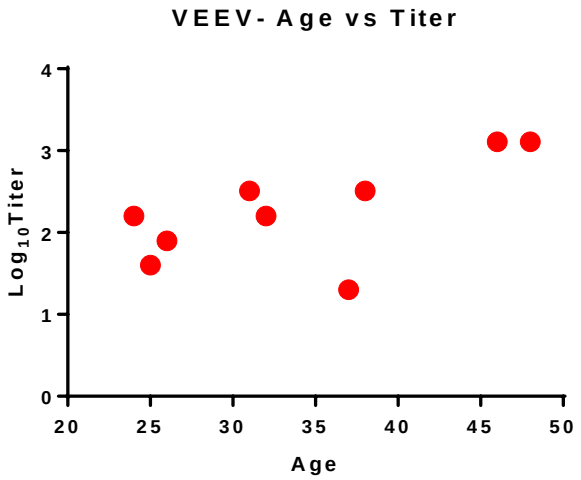


Figure 10. Correlation between Age and Titer in VEEV study.

The lack of significance prompted me to more closely examine the individual data points in the VEEV age data set; I observed that one of the data points, corresponding to an individual at 37 years of age with a Log10 post-vaccination titer of 1.3, appeared to be an outlier. When I removed that individual's data from the calculation, the two-tailed test of significance indicated a highly positive relationship between age of the VEEV subjects and the Log10 post-vaccination titer ($r = 0.8365$, $p = 0.015$). The data are illustrated in Figure 11. Based on the results of the Spearman rank correlation, there is a highly positive relationship between age of the VEEV vaccinated individuals and post-vaccination titer and the results are statistically significant at the alpha = 0.05 level. The null hypothesis is rejected and the alternative hypothesis that there is a correlation between age and vaccine outcome in the VEEV data set is accepted.

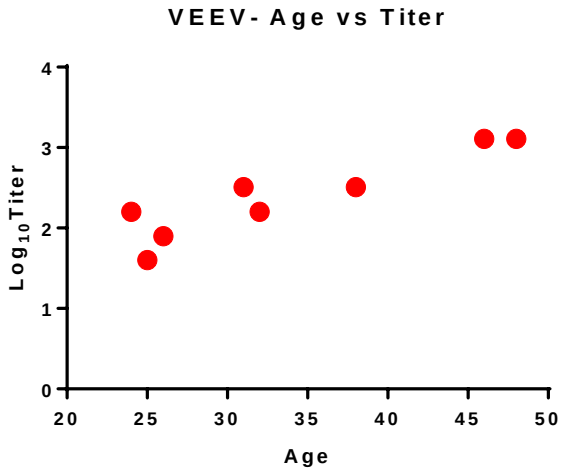


Figure 11. Correlation between Age and Titer in VEEV study (with outlier removed).

Association Between HLA Allele and Post-Vaccination Titer

To address research question 4, is there a significant association between HLA phenotype and vaccination outcome? a series of Mann-Whitney tests were conducted. Tests were not conducted between groups with only one value.

The data pairs were graphed to illustrate the spread of the data points, particularly where there was only one data point in a category; in such case no comparison could be made between that group and others (Figures 12-15).

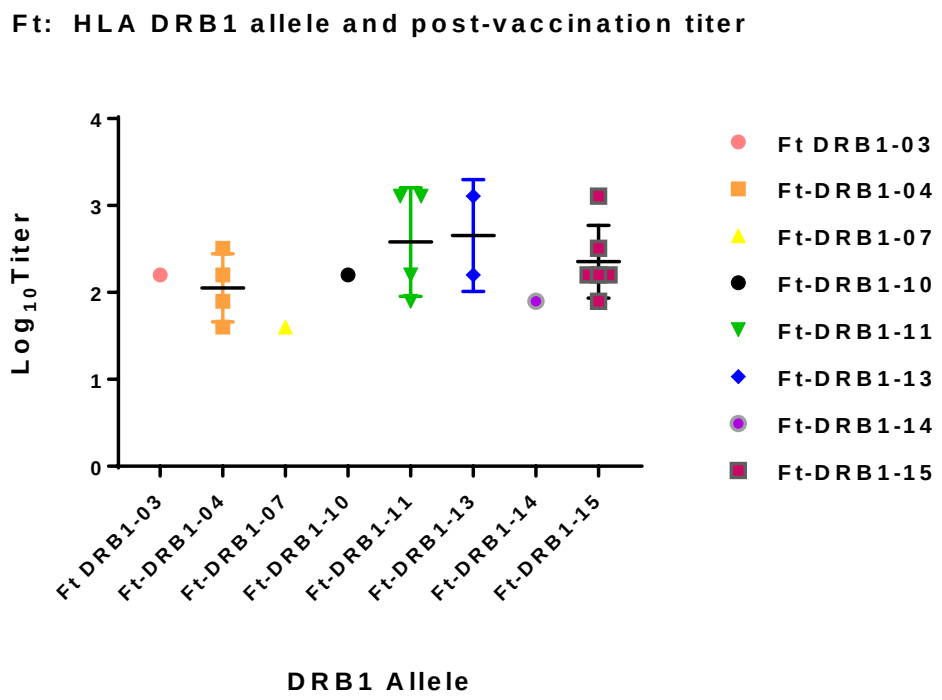


Figure12. HLA DRB1 allele and post-vaccination titer from the Ft study group.

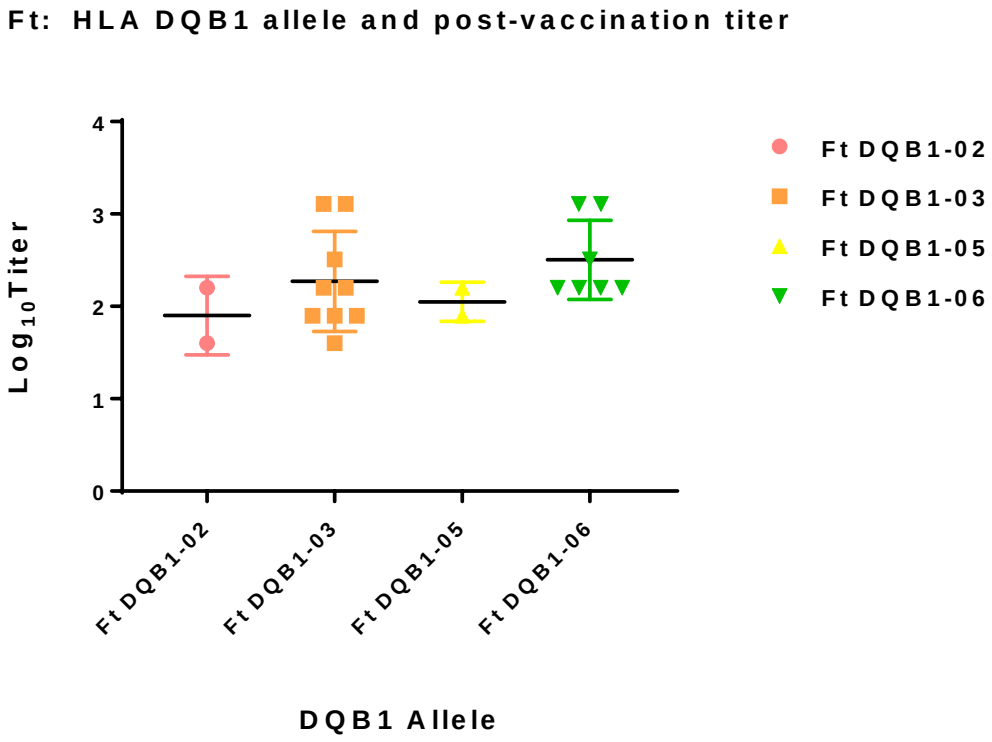


Figure 13. HLA DQB1 allele and post-vaccination titer from the Ft study group.

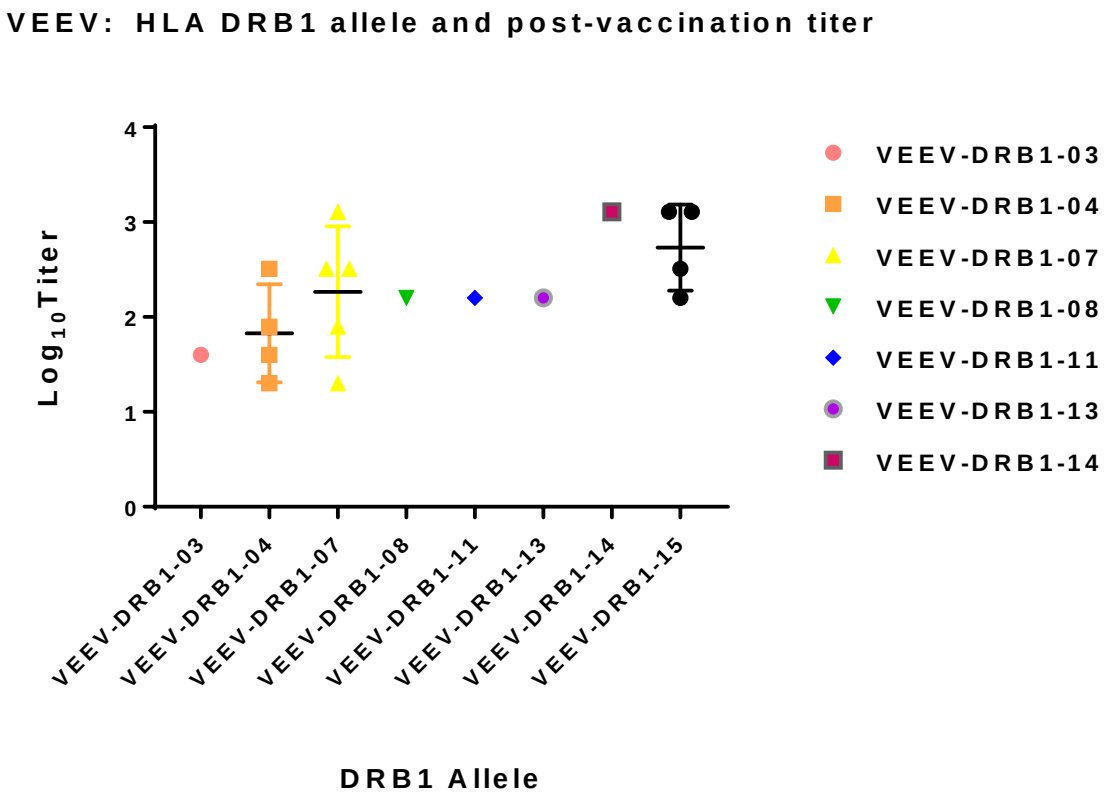


Figure 14. HLA DRB1 allele and post-vaccination titer from the VEEV study group.

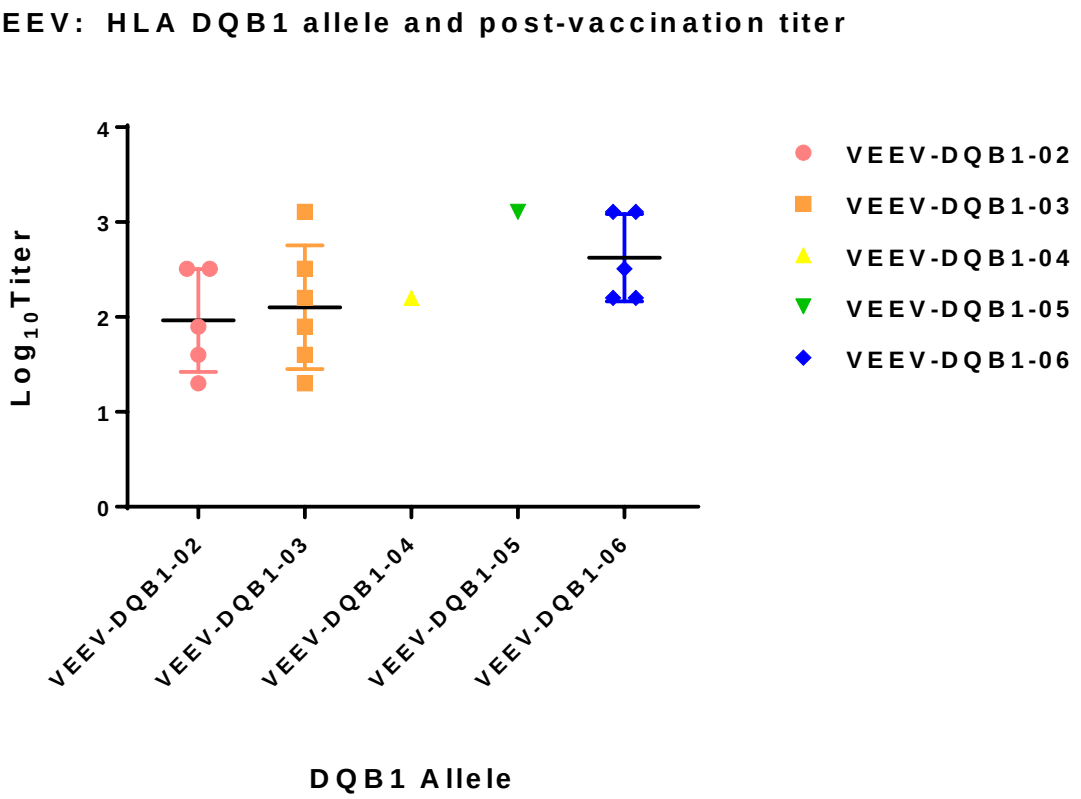


Figure 15. HLA DQB1 allele and post-vaccination titer from the VEEV study group.

Mann-Whitney test for Ft study: HLA DRB1 allele. For the Ft study group, I made statistical comparisons for the DRB1 allele between the following groups: DRB1-04, DRB1-11, DRB1-13, and DRB1-15. A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-04 (*Mdn* = 2.05) was not significantly different than the titer for study subjects with HLA allele DRB1-11 (*Mdn* = 2.655), $U = 4$, $p = 0.343$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-04 (*Mdn* = 2.05) was not significantly different than the titer for study subjects with HLA allele DRB1-13 (*Mdn* = 2.655), $U = 1.5$, $p = 0.400$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-04 (*Mdn* = 2.05) was not significantly different than the titer for study subjects with HLA allele DRB1-15 (*Mdn* = 2.2), $U = 7.5$, $p = 0.429$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-11 (*Mdn* = 2.655) was not significantly different than the titer for study subjects with HLA allele DRB1-13 (*Mdn* = 2.655), $U = 3.5$, $p > 0.999$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-11 (*Mdn* = 2.655) was not significantly different than the titer for study subjects with HLA allele DRB1-15 (*Mdn* = 2.2), $U = 10$, $p = 0.795$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-13 (*Mdn* = 2.655) was not significantly different than the titer for study subjects with HLA allele DRB1-15 (*Mdn* = 2.2), $U = 4$, $p = 0.786$.

Based on the results of the Mann-Whitney test, the null hypothesis must be accepted that there is no association between HLA DRB1 allele among the Ft study subjects and post-vaccination titer.

Mann-Whitney test for Ft study: HLA DQB1 allele. For the Ft study group, I made statistical comparisons for the DQB1 allele between the following groups: Ft DQB1-02, Ft DQB1-03, Ft DQB1-05, and Ft DQB1-06.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-02 (*Mdn* = 1.9) was not significantly different than the titer for study subjects with HLA allele DQB1-03 (*Mdn* = 2.2), $U = 5.5$, $p = 0.509$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-02 (*Mdn* = 1.9) was not significantly different than the titer for study subjects with HLA allele DQB1-05 (*Mdn* = 2.05), $U = 1.5$, $p > 0.999$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-02 (*Mdn* = 1.9) was not significantly different than the titer for study subjects with HLA allele DQB1-06 (*Mdn* = 2.2), $U = 2$, $p = 0.222$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-03 (*Mdn* = 2.2) was not significantly different than the titer for study subjects with HLA allele DQB1-05 (*Mdn* = 2.05), $U = 7.5$, $p = 0.891$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-03 (*Mdn* = 2.2) was not significantly different than the titer for study subjects with HLA allele DQB1-06 (*Mdn* = 2.2), $U = 20.5$, $p = 0.257$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-05 (*Mdn* = 2.05) was not significantly different than the titer for study subjects with HLA allele DQB1-06 (*Mdn* = 2.2), $U = 2$, $p = 0.222$.

Based on the results of the Mann-Whitney test, the null hypothesis must be accepted that there is no association between HLA DQB1 allele among the Ft study subjects and post-vaccination titer.

Mann-Whitney test for VEEV study: HLA DRB1 allele. For the Ft study group, comparisons were made for the DRB1 allele between the following groups: DRB1-04, DRB1-07, and DRB1-15. A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-04 (*Mdn* = 1.75) was not

significantly different than the titer for study subjects with HLA allele DRB1-07 (*Mdn* = 2.51), $U = 6$, $p = 0.389$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-04 (*Mdn* = 1.75) was not significantly different than the titer for study subjects with HLA allele DRB1-15 (*Mdn* = 2.81), $U = 1.5$, $p = 0.086$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DRB1-07 (*Mdn* = 2.51) was not significantly different than the titer for study subjects with HLA allele DRB1-15 (*Mdn* = 2.81), $U = 6$, $p = 0.389$.

Based on the results of the Mann-Whitney test, the null hypothesis must be accepted that there is no association between HLA DRB1 allele among the VEEV study subjects and post-vaccination titer.

Mann-Whitney test for VEEV study: HLA DQB1 allele. For the Ft study group, comparisons were made for the DQB1 allele between the following groups: DQB1-02, DQB1-03, and DQB1-06. A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-02 (*Mdn* = 1.9) was not significantly different than the titer for study subjects with HLA allele DQB1-03 (*Mdn* = 2.05), $U = 13.5$, $p = 0.864$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-02 (*Mdn* = 1.9) was not significantly different than the titer for study subjects with HLA allele DQB1-06 (*Mdn* = 2.51), $U = 5$, $p = 0.159$.

A two-tailed Mann-Whitney test indicated that the post-vaccination titer of study subjects with HLA allele DQB1-03 (*Mdn* = 2.05) was not significantly different than the titer for study subjects with HLA allele DQB1-06 (*Mdn* = 2.51), $U = 7.5$, $p = 0.193$.

Based on the results of the Mann-Whitney test, the null hypothesis must be accepted that there is no association between HLA DQB1 allele among the VEEV study subjects and post-vaccination titer.

One-way analysis of variance: Ft study. In addition to the Mann-Whitney tests, an ad hoc one-way ANOVA was also performed to measure association between HLA allele and post-vaccination outcome.

To reiterate, research question 4 asks if there is an association between the HLA allele of a study subject and the post-vaccination titer. The null hypothesis is that there is no significant difference on postvaccination titer between HLA phenotypes, while the alternative hypothesis states that there is a significant difference on postvaccination titer between HLA phenotypes.

The one-way ANOVA for the Ft study was calculated on the HLA DRB1 allele and post-vaccination titer. The analysis was not significant, $F(3,12) = 1.02$, and $p = 0.418$.

The one-way ANOVA for the Ft study was calculated on the HLA DQB1 allele and post-vaccination titer. The analysis was not significant, $F(3,16) = 1.082$, and $p = 0.385$.

One-way analysis of variance: VEEV study. The one-way ANOVA for the VEEV study was calculated on the HLA DRB1 allele and post-vaccination titer. The

analysis was not significant, $F(3,3) = 1.667$, and $p = 0.343$.

The one-way ANOVA for the VEEV study was calculated on the HLA DQB1 allele and post-vaccination titer. The analysis was not significant, $F(2,13) = 1.933$, and $p = 0.184$.

Based on the results of the one-way, the null hypothesis must be accepted that there is no association between either the HLA DRB1 or HLA DQB1 allele or post-vaccination titer in either the Ft or VEEV study.

Additional Ad Hoc Analyses: Ingenuity Pathway Analysis

Information from significantly differentially expressed transcripts, incorporating ± 2-fold changes in expression, were revealed from the mixed model ANOVA analysis conducted in Partek Genomics Suite and uploaded into QIAGEN's Ingenuity® Pathway Analysis (IPA) program (IPA, 2016). Data were analyzed through the use of IPA with a focus on the involvement of top transcripts in known canonical pathways. The significance of pathway enrichment was calculated using Fischer's Exact Test, with the cut-off for statistical significance set at $p < 0.05$. Pathway diagrams were generated using IPA.

Pathway analysis: Ft study. For the Ft study data set, only five transcripts were identified (Table 9) as meeting the criteria for differential expression and all were only differentially expressed on Day 2 post vaccination. The transcripts represented only three individual genes (each gene may have more than one transcript that represents it) and include STAT1, GBP1, and ANKRD22. All three genes have been implicated in interferon-stimulated or regulated pathways as illustrated in Figure 16, and the genes are

predominantly controlled by or affected by Type I and Type II interferon-stimulated responses.

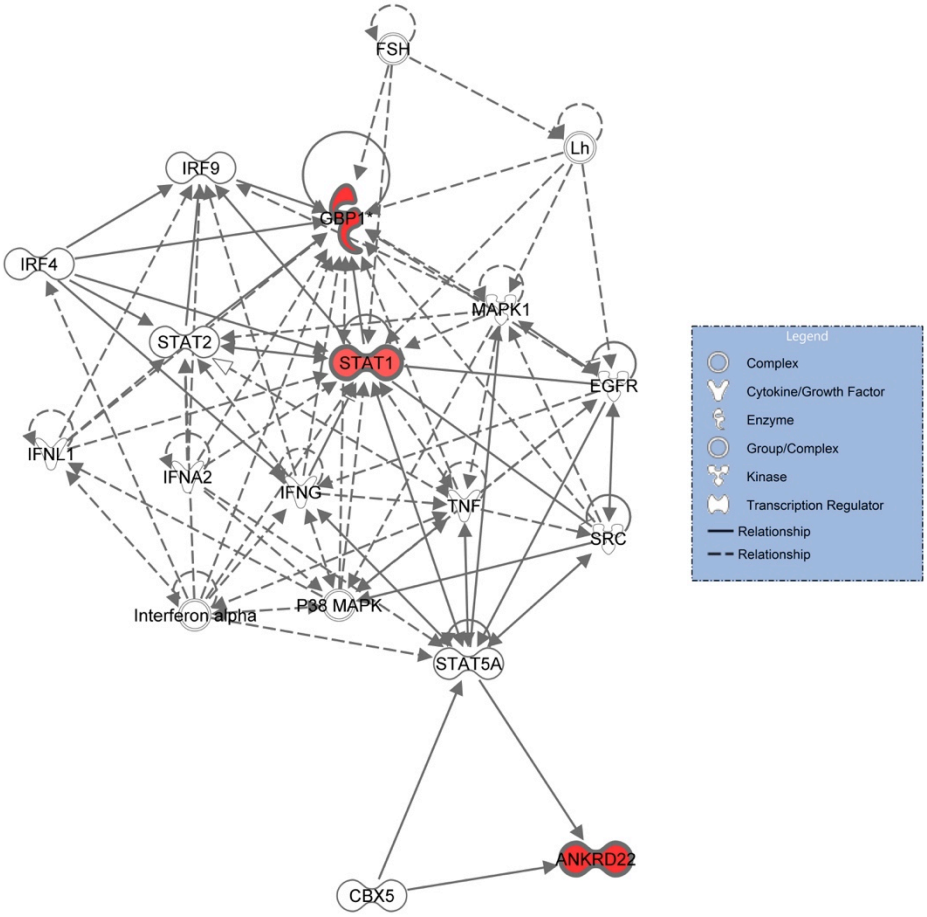


Figure 16. Pathway analysis of significant genes from Ft study.

Pathway analysis: VEEV study. For the VEEV study data set, hundreds of transcripts were identified (Table 10) as meeting the criteria for differential expression and were differentially expressed on Days 2, 7, and 14 post vaccination. The top five canonical pathways indicated for Days 2 and 7 were highly overlapping (Figure 17), whereas the top five canonical pathways revealed for Day 14 were very different (Figure 18).

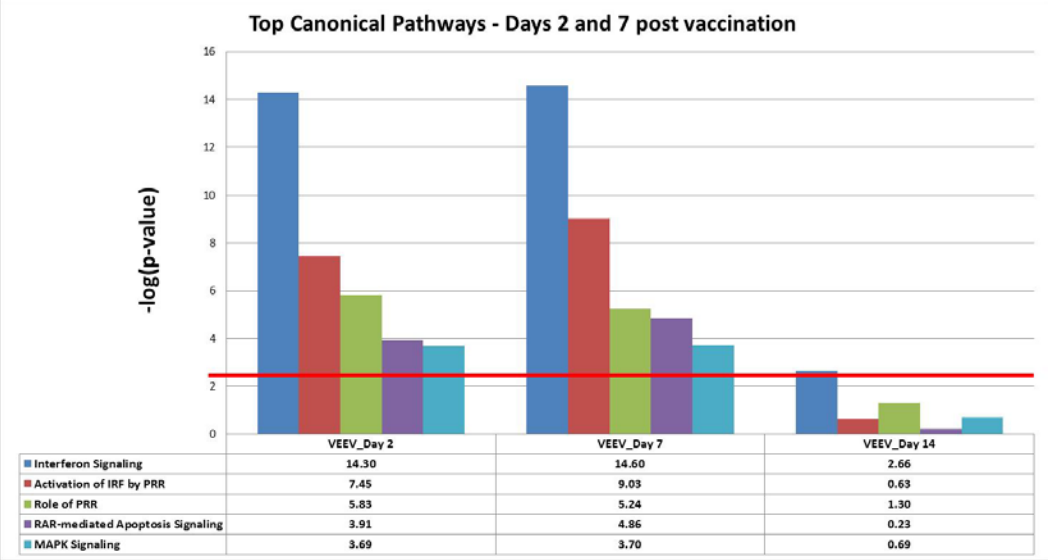


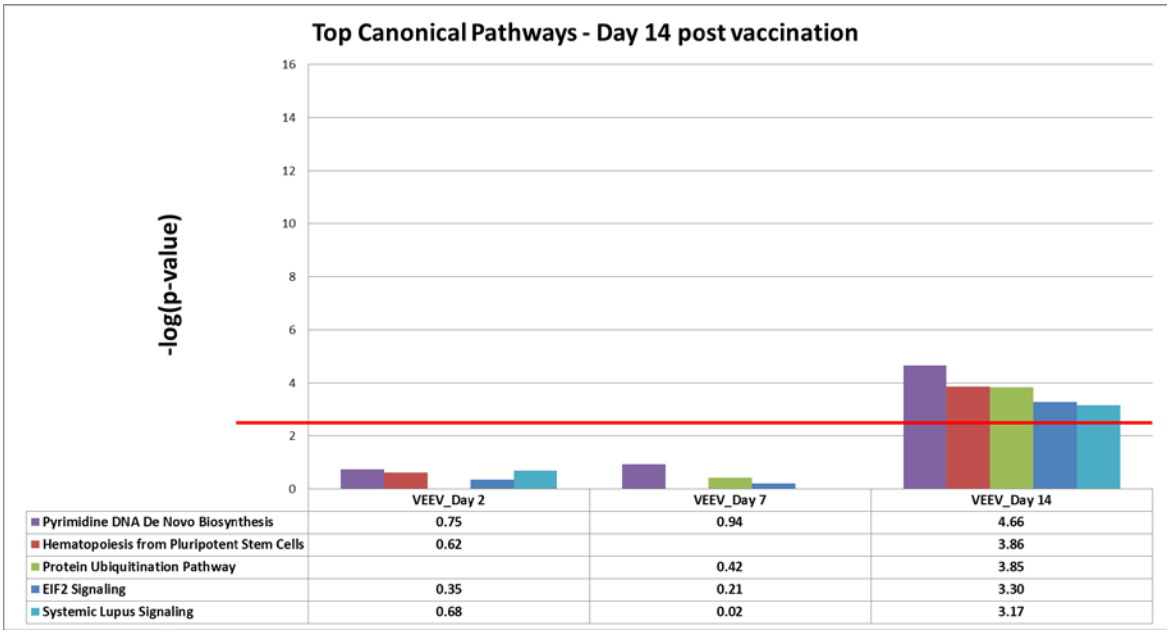
Figure 17. Top five canonical pathways for Days 2 and 7 in VEEV study.

The predominant pathway that was indicated by the analysis was interferon signaling. The minimum threshold of significance (the –log of p = 0.05) is indicated by the red line; the pathways are ordered by decreasing significance. Table 12 lists the individual genes that are implicated in each pathway illustrated.

Table 12

List of Genes in Top Canonical Pathways (Day 2-7) in VEEV Vaccine Study

Canonical pathway	*Study day*	*Differentially expressed genes in pathway*
Interferon Signaling	VEEV Day 2	OAS1,IFIT1,IFITM1,STAT1,IFIT3,STAT2,ISG15,MX1,IFI35,SOCS1,IFI6
	VEEV Day 7	OAS1,IFIT1,IFITM1,STAT1,IFIT3,STAT2,ISG15,MX1,IFI35,IFI6
	VEEV Day 14	OAS1,IFIT1,ISG15,IFI6
Activation of IRF by Cytosolic Pattern Recognition Receptors	VEEV Day 2	IRF7,STAT1,IFIH1,STAT2,ISG15,DDX58,ZBP1,IFIT2
	VEEV Day 7	IRF7,STAT1,IFIH1,STAT2,ISG15,DDX58,ZBP1,IFIT2
	VEEV Day 14	IFIH1,ISG15
Role of Pattern Recognition Receptors	VEEV Day 2	EIF2AK2,OAS1,IRF7,IRS2,IFIH1,OAS3,OAS2,DDX58 CXCL8
	VEEV Day 7	EIF2AK2,OAS1,IRF7,IFIH1,OAS3,OAS2,DDX58
	VEEV Day 14	OAS1,IFIH1,OAS3,ATM,OAS2
Retinoic acid Mediated Apoptosis Signaling	VEEV Day 2	PARP12,ZC3HAV1,TNFSF10,PARP9,PARP14
	VEEV Day 7	PARP12,ZC3HAV1,TNFSF10,PARP9,PARP14
	VEEV Day 14	CYCS
UVA-Induced MAPK Signaling	VEEV Day 2	PARP12,IRS2,ZC3HAV1,STAT1,PARP9,PARP14
	VEEV Day 7	PARP12,ZC3HAV1,STAT1,PARP9,PARP14
	VEEV Day 14	RRAS2,ATM,CYCS



	VEEV_Day 2	VEEV_Day 7	VEEV_Day 14
Pyrimidine DNA De Novo Biosynthesis	0.75	0.94	4.66
Hematopoiesis from Pluripotent Stem Cells	0.62		3.86
Protein Ubiquitination Pathway		0.42	3.85
EIF2 Signaling	0.35	0.21	3.30
Systemic Lupus Signaling	0.68	0.02	3.17

Figure 18. Top five canonical pathways for Day 14 in the VEEV study.

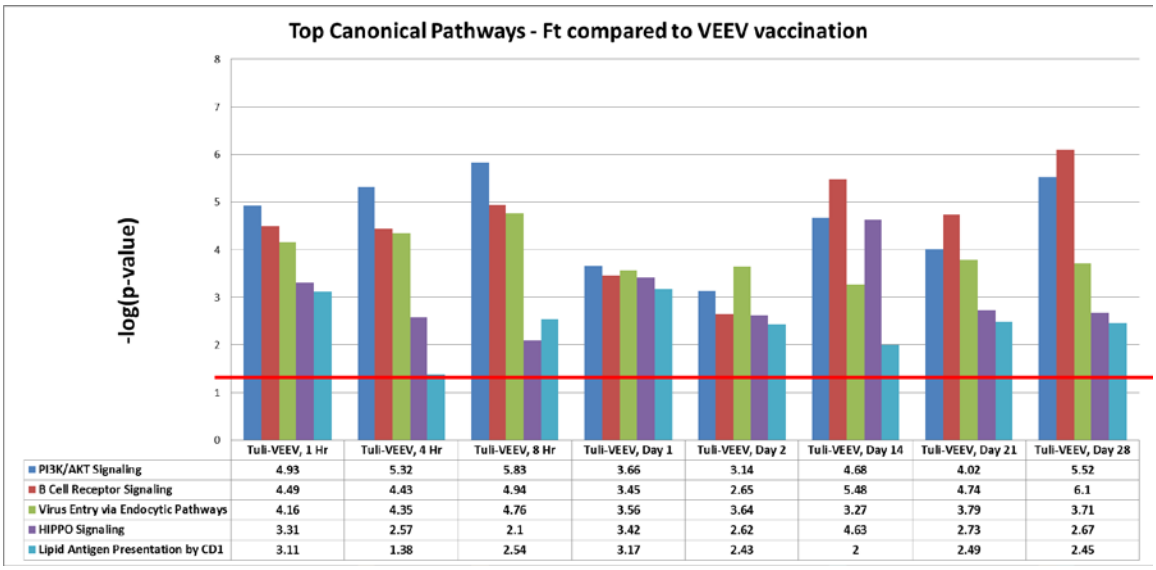
The predominant pathway that was indicated by the analysis was pyrimidine deoxyribonucleotides de novo biosynthesis. The minimum threshold of significance (the –log of p = 0.05) is indicated by the red line; the pathways are ordered by decreasing significance. Table 13 lists the individual genes that are implicated in each pathway illustrated.

Table 13

List of Genes in Top Canonical Pathways (Day 14) in VEEV Vaccine Study

Canonical pathway	*Study day*	*Differentially expressed genes in pathway*
Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	VEEV Day 2	CMPK2
	VEEV Day 7	CMPK2
	VEEV Day 14	RRM2,RRM1,TYMS,CMPK1,CMPK2
EIF2 Signaling	VEEV Day 2	EIF2AK2,IRS2
	VEEV Day 7	EIF2AK2
	VEEV Day 14	RPS27L,RPL17,RRAS2,ATM,RPL35,RPS7,RPS24,RPL36A,RPL22L1,RPL23
Hematopoiesis from Pluripotent Stem Cells	VEEV Day 2	CXCL8
	VEEV Day 7	
	VEEV Day 14	IGLC1,CD3D,IGHM,IGHG1,IGHA1
Protein Ubiquitination Pathway	VEEV Day 2	
	VEEV Day 7	UBE2L6,USP18
	VEEV Day 14	PSMA4,HSP90B1,UBR1,HSPA13,PSMA3,HSPE1,USP1,HSPH1,PSMC6,HSP90AA1, DNAJB9,DNAJB14,USP53
Systemic Lupus Erythematosus Signaling	VEEV Day 2	TNFSF13B,IL1RN,IRS2
	VEEV Day 7	IL1RN
	VEEV Day 14	SNRPD2,SNRPG,LSM5,CD28,RRAS2,CD3D,ATM,IGHM,IGHG1,PRPF39

Pathway analysis: Ft study in comparison to VEEV study. For the direct comparison between changes in the levels of gene expression from the Ft study compared with the VEEV study, hundreds of transcripts were identified (Table 11) as meeting the criteria for differential expression at every time point post vaccination. The top five canonical pathways illustrate the involvement of key canonical pathways across all days following infection with Ft in comparison to VEEV vaccination (Figure 19).



	Tuli-VEEV, 1 Hr	Tuli-VEEV, 4 Hr	Tuli-VEEV, 8 Hr	Tuli-VEEV, Day 1	Tuli-VEEV, Day 2	Tuli-VEEV, Day 14	Tuli-VEEV, Day 21	Tuli-VEEV, Day 28
PI3K/AKT Signaling	4.93	5.32	5.83	3.66	3.14	4.68	4.02	5.52
B Cell Receptor Signaling	4.49	4.43	4.94	3.45	2.65	5.48	4.74	6.1
Virus Entry via Endocytic Pathways	4.16	4.35	4.76	3.56	3.64	3.27	3.79	3.71
HIPPO Signaling	3.31	2.57	2.1	3.42	2.62	4.63	2.73	2.67
Lipid Antigen Presentation by CD1	3.11	1.38	2.54	3.17	2.43	2	2.49	2.45

Figure 19. Top five canonical pathways in the comparison of Ft to VEEV study.

The minimum threshold of significance (the –log of p = 0.05) is indicated by the red line; the pathways are ordered by decreasing significance. Table 14 lists the individual genes that are implicated in each pathway illustrated.

Table 14

List of Genes in Top Canonical Pathways in Ft Compared to VEEV Vaccine Study

Canonical pathway	*Study day*	*Differentially expressed genes in pathway*
PI3K/AKT Signaling	Tuli-VEEV, 1 Hr	PPP2R5E,PIK3CD,INPP5D,PTEN,MCL1,PTGS2,PPP2R1A,CDC37,JAK1,BCL2L1,HSP90AB1,SOS1
	Tuli-VEEV, 4 Hr	PPP2R5E,PIK3CD,PTEN,MCL1,PPP2R1A,CDC37,JAK1,BCL2L1,HSP90AB1,YWHAZ,HSP90AA1
	Tuli-VEEV, 8 Hr	PPP2R5E,PIK3CD,INPP5D,PTEN,MCL1,PPP2R1A,CDC37,JAK1,BCL2L1,HSP90AB1,YWHAZ
	Tuli-VEEV, Day 1	PPP2R5E,PIK3CD,PTEN,MCL1,PPP2R1A,JAK1,BCL2L1,HSP90AB1,YWHAB,SOS1
	Tuli-VEEV, Day 2	PPP2R5E,PIK3CD,PTEN,MCL1,PPP2R1A,JAK1,BCL2L1,HSP90AB1
	Tuli-VEEV, Day 14	PPP2R5E,PIK3CD,INPP5D,PTEN,MCL1,JAK1,BCL2L1,HSP90AB1,YWHAZ,YWHAB,PPP2R2A, SOS1
	Tuli-VEEV, Day 21	PPP2R5E,PIK3CD,INPP5D,PTEN,MCL1,PPP2R1A,JAK1,BCL2L1,HSP90AB1
	Tuli-VEEV, Day 28	PPP2R5E,PIK3CD,INPP5D,PTEN,MCL1,JAK1,BCL2L1,HSP90AB1,YWHAZ,PPP2R2A,SOS1
B Cell Receptor Signaling	Tuli-VEEV, 1 Hr	PTEN,RAC2,SOS1,RASSF5,CFL1,PIK3CD,INPP5D,BCL10,PPP3R1,APBB1IP,PTPRC,BCL2L1, NFAT5, RAP2A
	Tuli-VEEV, 4 Hr	CDC42,CFL1,PIK3CD,PTEN,PIK3AP1,IGHG1,PTPRC,RAP2B,BCL2L1,NFAT5,RAC2,RASSF5
	Tuli-VEEV, 8 Hr	CFL1,PIK3CD,INPP5D,PTEN,PIK3AP1,APBB1IP,PTPRC,RAP2B,BCL2L1,NFAT5,RAC2,RASSF5
	Tuli-VEEV, Day 1	CFL1,PIK3CD,PTEN,PPP3R1,PIK3AP1,CREB5,PTPRC,BCL2L1,NFAT5,RAC2,SOS1,RASSF5
	Tuli-VEEV, Day 2	CFL1,PIK3CD,PTEN,PIK3AP1,PTPRC,BCL2L1,NFAT5,RAC2,RASSF5
	Tuli-VEEV, Day 14	PTEN,PIK3C2A,RAC2,MAP3K13,SOS1,RASSF5,CFL1,INPP5D,PIK3CD,CREB5,PPP3R1,PIK3AP1, RAP2B, PTPRC, BCL2L1,NFAT5
	Tuli-VEEV, Day 21	CFL1,PIK3CD,INPP5D,PTEN,PPP3R1,PIK3AP1,CREB5,PTPRC,BCL2L1,NFAT5,RAC2,RASSF5
	Tuli-VEEV, Day 28	PTEN,RAC2,SOS1,RASSF5,CFL1,INPP5D,PIK3CD,CREB5,PPP3R1,PIK3AP1,RAP2B,PTPRC, BCL2L1, NFAT5

Table 14 (Con't)

List of Genes in Top Canonical Pathways in Ft Compared to VEEV Vaccine Study (Con't)

Canonical pathway	*Study day*	*Differentially expressed genes in pathway*
Virus Entry via Endocytic Pathways	Tuli-VEEV, 1 Hr	FYN,PIK3CD,FLNA,ACTB,FOLR1,AP2A1,B2M,ITGB2,AP2M1,RAC2
	Tuli-VEEV, 4 Hr	CDC42,FYN,PIK3CD,FLNA,ACTB,ITGB2,ACTA2,AP2M1,RAC2
	Tuli-VEEV, 8 Hr	PRKCD,FYN,PIK3CD,FLNA,ACTB,AP2A1,ITGB2,AP2M1,RAC2
	Tuli-VEEV, Day 1	FYN,PIK3CD,FLNA,ACTB,AP2A1,B2M,ITGB2,AP2M1,RAC2
	Tuli-VEEV, Day 2	FYN,PIK3CD,FLNA,ACTB,AP2A1,B2M,ITGB2,RAC2
	Tuli-VEEV, Day 14	FYN,PIK3CD,PIK3C2A,FLNA,ACTB,AP2A1,B2M,ITGB2,RAC2
	Tuli-VEEV, Day 21	FYN,PIK3CD,FLNA,ACTB,AP2A1,B2M,ITGB2,RAC2
	Tuli-VEEV, Day 28	FYN,PIK3CD,FLNA,ACTB,AP2A1,B2M,ITGB2,RAC2
HIPPO signaling	Tuli-VEEV, 1 Hr	PPP2R5E,DLG1,ITCH,PPP1R12A,MOB1A,PPP2R1A,CD44,STK4
	Tuli-VEEV, 4 Hr	PPP2R5E,DLG1,PPP2R1A,PPP1CB,YWHAZ,CD44
	Tuli-VEEV, 8 Hr	PPP2R5E,PPP1R12A,PPP2R1A,YWHAZ,CD44
	Tuli-VEEV, Day 1	PPP2R5E,DLG1,ITCH,PPP1R12A,PPP2R1A,YWHAB,CD44,STK4
	Tuli-VEEV, Day 2	PPP2R5E,DLG1,ITCH,PPP1R12A,PPP2R1A,CD44
	Tuli-VEEV, Day 14	PPP2R5E,DLG1,ITCH,PPP1R12A,PPP1CB,YWHAZ,YWHAB,PPP2R2A,CD44,STK4
	Tuli-VEEV, Day 21	PPP2R5E,DLG1,ITCH,PPP2R1A,CD44,STK4
	Tuli-VEEV, Day 28	PPP2R5E,PPP1CB,YWHAZ,PPP2R2A,CD44,STK4
Lipid Antigen Presentation by CD1	Tuli-VEEV, 1 Hr	AP2A1,PSAP,B2M,AP2M1
	Tuli-VEEV, 4 Hr	PSAP,AP2M1
	Tuli-VEEV, 8 Hr	AP2A1,PSAP,AP2M1
	Tuli-VEEV, Day 1	AP2A1,PSAP,B2M,AP2M1
	Tuli-VEEV, Day 2	AP2A1,PSAP,B2M
	Tuli-VEEV, Day 14	AP2A1,PSAP,B2M
	Tuli-VEEV, Day 21	AP2A1,PSAP,B2M
	Tuli-VEEV, Day 28	AP2A1,PSAP,B2M

Summary

Answers to Research Questions Based on Observed Results

I based my answers to the four research questions on the results obtained by the statistical analyses planned and executed as part of the dissertation research.

Research Question 1. Is there a significant association between changes in the level of gene transcription and the time course of vaccination with Ft or VEEV?

In the Ft study group, comparing Ft vaccinated individuals with mock-vaccinated control subjects, there were statistically significant changes in the level of expression over time as an independent variable, in response to treatment as an independent variable, in response to time*treatment considered as a single variable, for 8 hr, Day 2, and Day 14 post-vaccination. When fold-change was added to the decision criteria, only Day 2 results still retained transcripts that were both significant by p value and by fold-change in level of gene expression. The five transcripts represented only three genes, GBP1, STAT1, and ANDRD22.

In the VEEV study group, comparing VEEV vaccinated individuals with mock-vaccinated control subjects, there were statistically significant changes in the level of expression over time as an independent variable, in response to treatment as an independent variable, in response to time*treatment considered as a single variable, for Day 2, Day 7, and Day 14 post-vaccination. When fold-change was added to the decision criteria, Days 2, 7, and 14 results still retained transcripts that were both significant by p value and by fold-change in level of gene expression, although the numbers were greatly reduced. The differentially expressed transcripts were found to participate in several

canonical signaling pathways at early time points, including interferon signaling, activation of interferon response factors by pattern recognition receptors, role of pattern recognition receptors, apoptosis signaling and MAPK signaling. By Day 14, the active pathways shifted to include pyrimidine biosynthesis, hematopoiesis by pluripotent stem cells, the protein ubiquitination pathway, and EIF2 signaling.

Research Question 2. Is there a significant association between changes in the level of gene transcription over the time course of vaccination with Ft in comparison to VEEV?

Comparing Ft vaccinated individuals against VEEV vaccinated individuals; there were statistically significant changes in the level of expression over time as an independent variable and in response to treatment as an independent variable. In response to time*treatment considered as a single variable, there were significant changes in the level of gene expression at every time point post-vaccination. When a ± 2-fold-change in transcript expression was added to the decision criteria, all time point results still retained transcripts that were both significant by p value and by fold-change in level of gene expression, although the numbers were greatly reduced. The differentially expressed transcripts were found to participate in several canonical signaling pathways at various time points, including the PI3K/AKT signaling pathway, the B cell receptor signaling pathway, the HIPPO pathway, and the lipid antigen presentation by CD1 pathway.

Research Question 3. Is there a correlation between age of the vaccinated subject and vaccination outcome?

The relationship between age (or age group) and post-vaccination titer was assessed by correlation and negligible relationship was observed between the age of the Ft vaccinated study subjects and post-vaccination titer. The relationship between the age of the VEEV vaccinated study subjects and post-vaccination titer was observed to be moderately positive. In both cases, however, the resulting p values were not significant.

Research Question 4. Is there a significant association between Human Leukocyte Antigen (HLA) phenotype and vaccination outcome?

Association between HLA allele and post-vaccination titer was assessed both by nonparametric Mann-Whitney t-tests and by one-way ANOVA. There was no significant association between any HLA allele (DRB1 or DQB1) and post-vaccination titer for either Ft or VEEV study subjects.

In the next Chapter, I will discuss the results in greater detail in order to formulate conclusions. My discussion will include the limitations of the study, recommendations for future work, and implications of the research project will also be addressed.

Chapter 5: Discussion, Conclusions, and Recommendations

Overview

In this final chapter, I interpret results from the data analyses, discuss the limitations of the research study, I will make recommendations which may improve future statistical data analyses, address implications of the resulting findings on social change, and offer conclusions with possible explanations for some of the results. The discussion will begin by addressing each of the research questions and providing the findings of the research analyses.

The purpose of this study was to measure the potential relationship between gene expression over time and in response to treatment (vaccination) with either Ft or VEEV (as well as a direct comparison of Ft to VEEV); to assess the potential relationship between the age of the vaccinated individuals and vaccine outcome (successful response to primary vaccination or vaccine failure); and to assess the potential association between HLA phenotype and vaccine response.

The data analyses showed the association of a very small number of differentially expressed genes with vaccination with the live-attenuated Ft vaccine, in contrast to hundreds of differentially expressed genes that were associated with vaccination with live-attenuated VEEV vaccine. The results of the analysis in which I directly compared the levels of gene expression in Ft-vaccinated study subjects with the levels of gene expression in VEEV-vaccinated study subjects showed hundreds of differentially expressed genes in response to vaccination at every time point assessed. During analysis of the relationship between age and vaccine outcome (neutralizing antibody titer), I did

not find evidence of a relationship in either the Ft-vaccinated study group or in the VEEV-vaccinated study group. Finally, the analysis to examine the association of HLA allele and vaccine outcome (neutralizing antibody titer) failed to show a significant association between any specific DRB1 or DQB1 allele and vaccine outcome (neutralizing antibody titer) for Ft or VEEV-vaccinated study subjects.

Interpretation of the Findings

Association of Level of Gene Expression with Time and Treatment Following Vaccination

Research Questions 1 and 2 queried whether there were significant associations between changes in levels of gene transcription and the time course of vaccination with Ft or VEEV when compared against unvaccinated controls, or in direct comparisons made between Ft to VEEV vaccinated individuals.

Analysis Interpretations for Ft Vaccinated Versus Controls

In the literature review (Chapter 3), I detailed results from previous transcriptional studies involving Ft. The issue of the timing of sample collection was raised in previous studies; in the study conducted by Andersson et al. (2006) samples from seven patients presenting with naturally-occurring tularemia were collected and subjected to transcriptomic analysis (at roughly day 2-3, 6-7, 8-, 10-11, and 13 days following estimated time of infection), whereas in the Fuller et al. study (2007), PBMCs were isolated from six human volunteers at specified time points before and then after vaccination with the live-attenuated strain of Ft (before, then 18 Hr, 48 Hr, Day 8, and Day 14 following vaccination). In a third study by Paranavitana et al. (2008b), PBMCs

were collected from eight volunteers and challenged *in vitro* with the live-attenuated vaccine strain of Ft, collected at 1, 4, 8, 16, and 24 hours post-infection, and then subjected to transcriptomic analysis. In the present study, 10 vaccinated study subjects were age-matched to 10 control volunteers who received a mock vaccination. All subject pairs reported for post-vaccination sample collected within study protocol time lines (typically ± 1 hour for early time points).

The results of the mixed model ANOVA I conducted on samples from Ft vaccinated study subjects in comparison to mock-vaccinated controls only showed five transcripts which were differentially expressed, and transcript expression levels were only significant on day 2 following vaccination. The five transcripts represented three genes, STAT1, GBP1, and ANKRD22. The results for STAT1 and GBP1 confirm previously published results; however, ANKRD22 appears to be a novel transcript expressed in response to vaccination with Ft and extends the knowledge of human molecular responses to Ft vaccination. Both Andersson et al. (2006) and Paranavitana et al. (2008b) also observed differential expression of STAT1 and GBP1 genes at roughly the same time frame in their studies. None of these transcripts was identified as being differentially expressed in the study by Fuller et al. (2007). The function of ANKRD22 is not well understood, but the gene has been implicated in several disease states including as a diagnostic marker of pancreatic ductal adenocarcinoma and as an interferon-gamma stimulated gene implicated in T-cell mediated transplant rejection (Caba et al., 2014; Venner et al., 2014).

Analysis Interpretations for VEEV Vaccinated versus Controls

Differentially expressed transcripts were associated with VEEV vaccination on Days 2, 7, and 14 post-vaccination. The top canonical pathways that include the transcripts that were expressed include the interferon signaling pathway, activation of interferon-response factors by pattern recognition receptors pathway, the role of pattern recognition receptors pathway, the retinoic-acid receptor mediated apoptosis signaling pathway, and the MAPK signaling pathway. There is only a single report of human host responses to VEEV with which to compare the current results; the current results corroborate the previous publication with the observation that the role of pattern recognition receptors pathway was also indicated as a top canonical pathway in that previous report (Erwin-Cohen et al., 2012).

The study reported by Erwin-Cohen et al. (2012) was of PBMCs treated *in vitro* with the live-attenuated vaccine strain of VEEV, and samples were subsequently collected at 1, 4, 12, and 24 hours post-infection. Some of the most abundantly expressed transcripts in the VEEV human *in vivo* study which correspond to transcripts identified by the previous *in vitro* VEEV report by Erwin-Cohen et al. (2012) include OAS1, OAS2, OAS3, IFIH1, IRF7, DDX58, EIF2AK2, and STAT1. The broad involvement of interferon signaling responses was corroborated between the *in vitro* and *in vivo* studies, as well as transcripts that signal engagement of the inflammasome (Guo, Callaway, & Ting, 2015).

Many transcripts that were tentatively identified as biomarkers of a naïve or normal response by Erwin-Cohen et al. (2012) were also observed to have significantly

increased transcript expression in the current study including IFIT2, NEXN, TNFSF10, DDX58, SAMD9L, XAF1, LAMP3, and SLAMF7. Interestingly, the biomarkers that were suggested to represent or identify a molecular nonresponse to the vaccine strain of VEEV were not significantly differentially expressed in the current study, thus corroborating the data that those transcripts are only up-regulated during an alternate primary vaccine failure response.

Analysis Interpretations for Comparison of Ft to VEEV Vaccinated Responses

I structured the analysis results to compare the top expression results from the perspective of the Ft group, as well as the opposite where results are structured from the perspective of the VEEV group. The top canonical pathways included the PI3K/AKT signaling pathway, the B cell receptor signaling pathway, the pathway describing virus entry via endocytic pathways, the HIPPO signaling pathway (also known as the Salvador-Warts-Hippo pathway), and the lipid antigen presentation by CD1 pathway. Many of the top molecules in each of these pathways overlap with the other top pathways, and indicate that there is a great deal of signaling crosstalk in the pathways that transduce biological signals.

Among the hundreds of molecules that were observed to have increased transcription in response to Ft vaccination compared with VEEV vaccination, many molecules were observed to have expression that displayed a similar trend to what has previously been reported for other Ft studies (Andersson et al., 2006; Paranavitana et al., 2008b). Molecules including CASP1, GBP1, IFI16, SOD2, STAT1, NMI, PIM1, and TAP1 displayed a pattern of differential transcript expression where expression early (4-8

hours following vaccination, up to day 2) was significantly up-regulated. Other transcripts, including CD3E, CD59, and KPNA2 displayed a pattern of sustained elevated transcript expression following vaccination.

None of the transcripts identified by Fuller et al. (2007) as part of the "sustained up," "down early," "up early," or "down late" patterns of expression were observed to have significantly altered levels of expression (both by p-value and fold-change in expression) in the present study at any time point following vaccination; however, some of the transcripts displayed lower levels of altered expression (less than 2-fold in either direction).

Overall, in the present research, the molecules which I observed to have the broadest involvement in canonical signaling pathways included PPP2R5E, PIK3CD, PTEN, JAK1, BCL2L1, SOS1, NFAT5, CDC42, RAC2, RASSF5, FYN, CD44, ITCH, AP2A1, PSAP, and B2M.

PPP2R5E, PIK2CD, PTEN, JAK1, BCL2L1, and SOS1 are part of the PI3K/AKT signaling pathway. I observed a down-regulated pattern of expression for the PPP2R5E transcript in response to Ft vaccination in comparison to VEEV vaccination. The gene is a regulatory subunit which functions as part of a cascade that controls the activity of protein phosphatase PP2A (Cristóbal et al., 2015). In colorectal cancer, down regulation of PPP2R5E was suggested to be an important mechanism in the pathogenesis of disease (Cristóbal et al., 2015). The effect of Ft vaccination on PPP2R5E expression, as a regulator of PP2A, may indicate that Ft infection impinges on the PP2A pathway as a mechanism of controlling regulatory T cell (T reg) homeostasis (Apostolidis et al., 2016).

Apostolidis et al. (2016) reported that PP2A was a previously-unrecognized requirement for T reg function, and Gerriets et al. (2016) provided additional support for the idea, noting that the PI3K/AKT pathway regulates metabolism needed to support the T reg function. Novel regulators of PP2A activity, including FTY720, may facilitate proliferation of immune cells in response to infection (or vaccination) as well as caspase-dependent apoptosis (Cristóbal et al., 2015).

The involvement of PI3KCD transcripts within the PI3K/AKT pathway that I observed demonstrate the potential involvement of neutrophil activation signals in the pathogenesis of Ft disease or within the immune response generated by vaccination with the live-attenuated form of Ft; however, the mechanistic role of neutrophils in the course of tularemia infection has been suggested to be one where Ft leverages suppression of the natural turnover of neutrophils as part an innate immune suppression strategy (Schwartz et al., 2012).

Decreased expression of PTEN may be related to a vaccine-induced dampening of the innate immune responses by elicited through the activation of interferon and impinging on the expression of IL-6 (Li et al., 2016). The role of PTEN in antiviral immunity has been discussed previously (Li et al., 2016); however the current data suggest a novel function for PTEN in anti-microbial immunity as well. Indeed, decreased PTEN expression has a downstream consequence of leading to increased numbers of T reg cells; however, those T reg cells have a decreased stability (Huynh et al., 2015). JAK1 expression varied over time in response to Ft vaccination compared to VEEV vaccination; at very early time points, JAK1 expression was decreased (1, 4 hours post

vaccination) however, at 8 hours, day 2, and day 28 post-vaccination, JAK1 expression was up-regulated by Ft vaccination. The temporal effects I observed in transcript expression are in agreement with the tightly-controlled regulation of expression and activation of Janus kinases as well as the pleotropic role that JAK1 plays as a signal transducer of multiple cytokine signals (e.g., interferon signaling and various interleukins) (Roskoski, 2016).

The role of BCL2L1 in the PI3K/AKT canonical pathway is likely one of mediating apoptosis through the activation cascade of TGF-beta1 and p53 signaling/apoptosis; the transcript levels for this gene were consistently elevated in response to vaccination with Ft compared with VEEV vaccination (IPA, 2016). Finally, I showed that the expression of SOS1was consistently decreased in response to vaccination with Ft at following vaccination, which was in contrast to the expression level expected or predicted by IPA; SOS1 is a guanine nucleotide exchange factor and adaptor protein which functions via the ERK and MAPK signaling cascades to mediate the activation of RAS (Tian & Feig, 2001).

I observed that the genes encoding NFAT5, CDC42, RAC2, RASSF5 were differentially expressed in response to Ft vaccination in comparison to VEEV vaccination; these genes participate in the B cell receptor signaling pathway. NFAT5 was observed to have sustained decreased transcript expression over time following Ft vaccination which was opposite of the predicted normal response of the gene as part of the B cell receptor signaling pathway. The observation I made of differential inhibition of expression of NFAT5 suggests that the mechanism of immune suppression following

vaccination with Ft may leverage the nuclear factors of activated T cells family of transcription factors (IPA, 2016). I identified other important molecules in the B cell receptor signaling pathway that displayed increased expression in response to Ft vaccination, including CDC42, RAC2, and RASSF5. Both CDC42 and RAC2 are part of a family of small GTPases; each has also been shown to have ubiquitin protein ligase activity and may impact the B cell signaling pathway as part of a concerted immune suppression mechanism (IPA, 2016).

I found that two additional molecules of importance to the analysis were those of TLR4 and NOD2. I observed that TLR4 displayed increased expression at key time points of 4-8 hours following vaccination with Ft in comparison to VEEV as well as at Day 2; TLR4 is an important molecule involved in activation of the inflammasome. Conversely, NOD2 expression was decreased at each time point following vaccination. NOD2 plays an important role in the immune response to intracellular bacterial lipopolysaccharides (LPS) by recognizing the muramyl dipeptide (MDP) derived from them and activating the NFKB pathway (IPA, 2016). The inclusion of TLR4 as a key gene that was observed to have differential expression in response to vaccination is important to potential future vaccine constructs as TLR4has been used as a natural adjuvant to boost innate immune signals for rationally designed vaccines such as Melacine, Fenrix, and Cervarix (Reed, Hus, Carter, & Orr, 2016).

Analysis Interpretations Relative to Epidemiology Nexus Conceptual Theory

In brief terms the theoretical concept, discussed in far greater depth in chapter 3, that guided my dissertation study was a comprehensive understanding of human immune

responses to vaccination may provide key insight into the molecular epidemiology of infectious diseases and suggest innovative approaches to address global health challenges to produce new vaccines. I have termed this collection of concept ideals from leading subject matter experts the Epidemiology Nexus Concept. The results I report from my dissertation research illustrate association between changes in the levels of gene expression in response to vaccination with Ft, VEEV, and between Ft compared with VEEV, and further indicate how those differentially expressed molecules interact and overlap with key canonical signaling pathways. Collectively, I present the results as a quantitative analysis of the complex biological interactions that jointly result in immunocompetence and contribute a small piece of a knowledge foundation on the mechanisms of action for the Ft and VEEV vaccines. I formulated the Epidemiology Nexus Concept as a theory to describe the connections that exist and need to be explored between host genetic responses, demographic factors, disease responses, and the health outcomes of individuals. I conclude that the results from my research study support the Epidemiology Nexus conceptual theory formulated by subject matter experts.

Analysis Interpretations for Comparison of Age to Vaccine Responses

I conducted correlational analyses to examine the relationship between age and vaccine response in both the Ft study and VEEV studies. My interpretations of the results lead me to accept the null hypothesis for the Ft study data set. For the VEEV study, after I conducted my initial analysis for the relationship between age and production of neutralizing titer as a measure of vaccine success, I conducted an additional analysis where I removed one data point that appeared to be an outlier. Once the outlier point was

removed from the data set, I observed a highly positive relationship between age and vaccine titer in the VEEV group. These results were also statistically significant. For the Ft study, the small number of subjects in the data set was likely a contributing factor to the lack of significance in the correlation.

Analysis Interpretations for Comparison of HLA to Vaccine Responses

Based on the analyses to examine the association of HLA phenotype and vaccine response in both the Ft study and VEEV study, I accepted the null hypothesis for both the Ft and VEEV studies. For both studies, the small number of subjects in each data set was likely a contributing factor to the lack of significant association.

Limitations of the Study

The secondary analyses of the two data sets used in this study are limited by the design and quality of each of the original studies. In particular, it should be noted that each original was comprised of only male study volunteers. The study volunteers were overwhelmingly of Caucasian, not Hispanic or Latino race and ethnicity; therefore, the data are also limited in that there can be no correlation of vaccine response with any racial or ethnic group. In addition, only 10 volunteers were included in each treatment group (e.g., control and vaccinee). Sample size is not a limitation for the research questions related to the association of transcript expression with time or treatment; however, such small sample sized lacked power for the research questions of relationship between age and vaccine outcome and association between HLA phenotype and vaccine outcome.

The live attenuated strain of Ft that was used in the original study may not produce the same host responses to infection in a human that a wild-type, fully pathogenic strain of bacteria may elicit. The live attenuated strain of VEEV that was used to vaccinate volunteers belongs to the IA/B strain of VEEV; the changes in gene expression that are associated with this particular strain may or may not be the same as what would be observed if using a virus from one of the other VEEV subtypes or with fully virulent strains. There is no way to test the veracity of the live attenuated vaccine strains as surrogates of virulent infection in humans. Future studies in appropriate animal models may be able to suggest whether or not the host responses to infection are similar.

Measures were implemented during the collection of the original data to control bias, including blinding of sample collection personnel to the treatment groups and collection of all samples within a pre-defined range of acceptable collection (e.g., 4 hours post-vaccination ± 30 minutes, day 2 post-vaccination ± 2 hours, or day 21 post-vaccination ± 2 days). Some types of bias could not be controlled in the study including attrition bias (study participants had the right to terminate participation in the study at any time) and exclusion bias, in particular the exclusion of women from the study in order to control variation in transcript expression that could be due to fluctuations in female hormone signaling. An additional type of bias was revealed during the analysis and that represents a technical bias; variances in the transcriptional data were observed to coincide with the scan date that the Affymetrix chips were read following hybridization.

Recommendations

Based on both the results of the research analyses and a review of the study limitations, there are several recommendations that I suggest to facilitate future studies. First, future transcriptomic studies should plan to include women; it would be helpful to examine the association of changes in the level of gene expression in response to vaccination in each individual population. As I discussed in the literature review, it has been suggested that normal hormonal variation in adult, non-menopausal women can affect their immune responses; however, it could be highly beneficial to future vaccine studies to examine this conundrum directly. Second, in terms of testing the potential relationship between age and HLA phenotype with vaccine outcome, my results suggest that far greater numbers of volunteers will need to be examined. While 10 individuals per group was more than sufficient for the microarray portion of the study, that number fell far short for correlation of age to vaccine outcome or association of HLA with vaccine outcome. My third recommendation would be to examine the relationship between a diversified study sample and vaccine outcome. The study populations for the Ft and VEEV studies consisted predominantly of Caucasian race and non-Latino or Hispanic ethnicity.

With regard to the transcriptome studies (microarray), one of the sources of variation or noise in the data that I found was the scan date on which the Affymetrix chips were processed and read; future studies should endeavor to have all relevant chips processed within the same batch and the results read on the same day. The mixed model ANOVA that I employed to examine the effects of Time*Treatment incorporated scan

date and thus the effects of scan date as a variable were considered in the analysis. The issue of missing sample data may not be able to be resolved completely; it is impossible to prevent loss to follow-up; however, I recommend recruiting additional study participants may alleviate the number of samples falling below the minimum number for which the study is powered. As an example, if power calculations estimate that 10 volunteers are needed, then recruiting twelve with an estimated 20% loss to follow up could reduce biased results due to inadequate samples in analyses. Additionally, I observed few responses using the parameters for acceptance of Step-up p-value < 0.01 and fold-change in gene expression ± 2-fold change in expression level. For future analyses of association of changes in the levels of transcript expression, it may be beneficial to consider changes that are more modest, for example ± 1.5-fold or less. The interpretation of microarray data using lower cutoff values for fold-change elicits the question of whether biological relevance or statistical significance is more important within the context of data interpretation (Dalman, Deeter, Nimihakavi, & Duan, 2012).

Implications

Positive Social Change

The research I have described in the dissertation may have implications for social change on several levels, and as such also represents an important component of the Epidemiology Nexus concept that I believe supports the links between molecular epidemiology and health outcomes. At the human level (individual, family, community) vaccination with Ft or VEEV could alleviate mortality and morbidity due to endemic disease. Since the greater threat of disease caused by either agent is one of a biological

threat, vaccination could prevent large-scale deaths and morbidity, and in turn prevent catastrophic costs associated with the long term effects of disease, particularly in the case of VEEV infection.

At the level of policy, the research project I have described in the dissertation may serve as part of a foundation of knowledge of the molecular epidemiology of host responses to either Ft or VEEV; in turn, such knowledge may be useful in comparing the transcriptional responses to vaccination with a multitude of other infectious agents, such as traditional childhood diseases (e.g., chicken pox, measles, whooping cough, etc.), other biological warfare agents (e.g., *Bacillus anthracis*, *Yersinia pestis*, Ebola, etc.), or emerging diseases such as Zika virus and West Nile virus. Understanding the host molecular responses to these diseases through vaccination may decrease developmental costs associated with the design and approval of new vaccines which can meet the rigorous standards of FDA licensure. Indeed, the observation that I made of increased expression of a Toll-like receptor, TLR4, as induced by Ft vaccination suggests that rationally designed vaccines which incorporate TLR4 as a biological adjuvant may be able to enhance innate and adaptive immune responses to Ft vaccination.

Conclusion

The results I have reported in the dissertation study provide key informative findings regarding the association of host transcripts and treatment with either Ft or VEEV live-attenuated vaccines. From the patterns of expression in response to Ft vaccination, it is apparent that the Ft vaccine promoted a systemic immunosuppression which was observed as an overall lack of variation in transcript expression; indeed only

three genes were significantly altered and observable only at day 1 following vaccination. In contrast, I observed that vaccination with VEEV induced a robust interferon-driven response that was most noticeable at days 2, 7, and 14 post-vaccination. When levels of expression in response to Ft and VEEV vaccination were compared directly, the differences elicited in transcript expression between the two vaccines were quite different, with expression of several hundred transcripts significantly modulated at each time point following vaccination. The participation of specific genes suggested key canonical pathways which are dissimilarly involved in the innate and adaptive response to Ft versus VEEV vaccination including the PI3K/AKT pathway, B cell receptor signaling, virus entry by endocytic pathways, HIPPO signaling, and lipid antigen presentation by CD1. In addition to the suggestion that there are diverse pathways involved in the mediation of immune responses between Ft and VEEV vaccination, the data from the analyses I have made may provide a basis for natural adjuvants which can be leveraged in rational vaccine design to augment the innate and adaptive immune response, for example through the use of TLR4 as a biological adjuvant.

I reported a highly positive correlation between the age of the vaccinated volunteers and the production of neutralizing titer in the VEEV study which was statistically significant. The mechanisms which may be contribute to higher neutralizing titers in older populations are not clear and it remains to be determined if the trend is consistent in a larger group of volunteers. Future studies should be conducted to corroborate and extend the results I have observed within this research project. Greater numbers of volunteers may provide power to detect more discreet differences in gene

expression, as well as effects on vaccine outcome due to age of the individual and HLA phenotype.

References

Adams, A. P., Navarro-Lopez, R., Ramirez-Aguillar, F. J., Lopez-Gonzalez, I., Leal, G., Flores-Mayorga, J. M., … Estrada-Franco, J. G. (2012). Venezuelan equine encephalitis virus activity in the gulf coast region of Mexico, 2003-2010. *PLOS Neglected Tropical Diseases, 6*(11), e1875. doi:10.1371/journal.pntd.0001875

Affymetrix. (2004). Data sheet: GeneChip human genome arrays. Retrieved from http://media.affymetrix.com/support/technical/datasheets/human_datasheet.

Aguilar, P. V., Estrada-Franco, J. G., Navarro-Lopez, R., Ferro, C., Haddow, A. D., & Weaver, S. C. (2011). Endemic Venezuelan equine encephalitis in the Americas: hidden under the dengue umbrella. *Future Virology, 6*(6), 721-740.doi:10.2217/FVL.11.5

Alcina, A., del Mar Abad-Grau, M., Fedetz, M., Izquierdo, G., Lucas, M., Fernandez, O., … Matesanz, F. (2012) Multiple Sclerosis Risk Variant HLA-DRB1*1501 Associates with High Expression of DRB1 Gene in Different Human Populations. *PLoS ONE 7*(1), e29819. doi:10.1371/journal.pone.0029819

Andersson, H., Hartmanová, B., Bäck, E., Eliasson, H., Landfors, M., Näslund, L., … Sjöstedt, A. (2006). Transcriptional profiling of the peripheral blood response during tularemia. *Genes and Immunity, 7*(6), 503-513. doi:10.1038/sj.gene.6364321

Angel, J., Steele, A. D., & Franco, M. A. (2014). Correlates of protection for rotavirus vaccines: Possible alternative trial endpoints, opportunities, and challenges. *Human Vaccines and Immunotherapeutics, 10*(12), 3659-3671.

http://dx.doi.org/10.4161/hv.34361

Apostolidis, S. A., Rodriguez-Rodrigues, N., Suarez-Fueyo, A., Dioufa, N., Ozcan, E., Crispin, J. C., … Tsokos, G. C. (2016). Protein phosphatase 2A is requisite for the function of regulatory T cells. *Nature Immunology, 17*(5), 556-564. doi:10.1038/ni.3390

Aréchiga-Ceballos, N. & Aguilar-Setién, A. (2015). Alphaviral equine encephalomyelitis (eastern, western, and Venezuelan). *Scientific and Technical Review of the Office International des Epizooties (Paris), 34*(2), 491-501. doi: 10.20506/rst.34.2.2374

Arlt, W. & Hewison, M. (2004). Hormones and immune function: Implications of aging. *Aging Cell, 3*(4), 209-216. doi:10.1111/j.1474-9728.2004.00109.x

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

Baddal, B., Muzzi, A., Censini, S., Calogero, R., Torricelli, G., Guidotti, S., … Pezzicoli, A. (2015). Dual RNA-seq on nontypeable *Haemophilus influenza* and host cell transcriptomes reveals novel insights into host-pathogen cross talk. *mBio, 6*(6), e01765-15. doi:10.1128/mBio.01765-15

Beck, C. E. & Wyckoff, R. W. G. (1938). Venezuelan equine encephalomyelitis. *Science, 88*(2292), 530. doi:10.1126/science.88.2292.530

Benjamini, Y. & Höchberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B, 57*(1), 289-300. doi:0035-9246/95/57289

Bolstad, B. M., Irizarry, R. A., Astrand, M., & Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics, 19*(2), 185-193. doi:10.1093/bioinformatics/19.2.185

Bowen, G. S., Fashinell, T. R., Dean, P. B., & Gregg, M. B. (1976). Clinical aspects of human Venezuelan equine encephalitis in Texas. *Bulletin of the Pan-American Health Organization, 10* (1), 46-57.

Brault, A. C., Powers, A. M., Medina, G., Wang, G., Kang, W., Salas, R. A., …Weaver, S. C. (2001). Potential sources of the 1995 Venezuelan equine encephalitis subtype IC epidemic. *Journal of Virology, 75*(13), 5823-5832. doi:10.1128/JVI.75.13.5823–5832.2001

Burgner, D., Jamieson, S. E., & Blackwell, J. M. (2006). Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? *Lancet Infectious Diseases, 6*, 653-663. doi:10.1016/S1473-3099(06)70601-6

Caba, O., Prados, J., Ortiz, R., Jimenez-Luna, C., Melguizo, C., Alvarez, P. J., … Aranega, A. (2014). Transcriptional profiling of peripheral blood in pancreatic adenocarcinoma patients identifies diagnostic markers. *Digestive Diseases and Sciences, 59*(11), 2714-2720. doi:10.1007/s10620-014-3291-3

Centers for Disease Control and Prevention. (2015a). Retrieved from http://www.cdc.gov /tularemia/statistics/state.html

Centers for Disease Control and Prevention. (2015b). Retrieved from http://www.cdc.gov /tularemia/statistics/year.html

Centlivre, M., & Compadière, B. (2015). New challenges in modern vaccinology. *BMC Immunology*, *16*, 18. doi:10.1186/s12865-015-0075-2.

Cook, I. F. (2008). Sexual dimorphism of humoral immunity with human vaccines. *Vaccine*, *26*, 3551–3555. doi:10.1016/j.vaccine.2008.04.054

Cristóbal, I., Manso, R., Rincón, R., Caramés, C., Senin, C., Borrero, A., … Garcia-Foncillas, J. (2014). PP2A inhibition is a common event in colorectal cancer and its restoration using FTY720 shows promising therapeutic potential. *Molecular Cancer Therapeutics, 13*(4), 938-947. doi:10.1158/1535-7163.MCT-13-0150

Dalman, M. R., Deeter, A., Nimishakavi, G., & Duan, Z.-H. (2016). Fold change and p-value cutoffs significantly alter microarray interpretations. *BMC Bioinformatics, 13*(Suppl 2), S11. http://www.biomedcentral.com/1471-2105/13/S2/S11.

Das, R., Hammamieh, R., Neill, R., Ludwig, G. V., Eker, S., Lincoln, P., … Jett, M. (2008). Early indicators of exposure to biological threat agents using host gene profiles in peripheral blood mononuclear cells. *BMC Infectious Diseases*, *8*, 104. doi:10.1186/1471-2334-8-104

Delves, P. J. (2016). Merck Manual. Human leukocyte antigen (HLA) system. Retrieved from http://www.merckmanuals.com/professional/immunology-allergic-disorders/biology-of-the-immune-system/human-leukocyte-antigen-hla-system

Dennis, D. T., Inglesby, T. V., Henderson, D. A., Barlett, J. G., Ascher, M. S., Eitzen, E. …Tonat, K. (2001). Tularemia as a biological weapon. *Journal of the American Medical Association, 285*(21), 2763-2773. doi:10.1001/jama.285.21.2763

Di Benedetto, S., Derhovanessian, E., Steinhagen-Thiessen, E., Goldeck, D., Muller, L.,

& Pawelec, G. (2015). Impact of age, sex and CMV-infection on peripheral T cell phenotypes: results from the Berlin BASE-II study. *Biogerontology, 16,* 631-643. doi 10.1007/s10522-015-9563-2

Ellis, J., Oyston, P. C. F., Green, M., & Titball, R. W. (2002). Tularemia. *Clinical Microbiology Reviews, 15*(4), 631-646. doi: 10.1128/CMR.15.4.631–646.2002

Ehrenkranz, N. J., & Ventura, A. K. (1974). Venezuelan equine encephalitis virus infection in man. *Annual Reviews in Medicine, 25,* 9-14. doi:10.1146/annurev.me.25.020174.000301

Erwin-Cohen, R., Porter A., Pittman, P., Rossi, C., & DaSilva, D. (2012). Host responses to live-attenuated Venezuelan equine encephalitis virus (TC-83): Comparison of naïve, vaccine responder and nonresponder to TC-83 challenge in human peripheral blood mononuclear cells. *Human Vaccines and Immunotherapeutics, 8* (8), 1053 – 1065. http://dx.doi.org/10.4161/hv.20300

Eskola, J., Duclos, P. Schuster, M., MacDonald, N. E., & the SAGE Working Group on Vaccine Hesitancy. (2015). How to deal with vaccine hesitancy? *Vaccine, 33,* 4215-4217. http://dx.doi.org/10.1016/j.vaccine.2015.04.043

Foley, J. E., & Nieto, N. C. (2010). Tularemia. *Veterinary Microbiology, 140,* 332-338. doi:10.1016/j.vetmic.2009.07.017

Fourati, S. Cristescu, R., Loboda, A., Talla, A., Filali, A., Railkar, R., … Sekaly, R. P. (2016). Pre-vaccination inflammation and b-cell signaling predict age-related hyporesponsiveness to hepatitis B vaccination. *Nature Communications, 7,* 10369. doi:10.1038/ncomms10369.

Francis, E. (1919). Deer-fly fever—a disease of hitherto unknown etiology. *Public Health reports*, *92*, 1838-1839.

Franz, D. R., Jahrling, P. B., Friedlander, A. M., McClain, D. J., Hoover, D. L., Bryne, W. R., …Eitzen, E. M. (1997). Clinical recognition and management of patients exposed to biological warfare agents. *Journal of the American Medical Association, 278*(5), 399-411. doi:10.1001/jama.1997.03550050061035

Fuller, C. L., Brittingham, K. C., Hepburn, M. J., Martin, J. W., Petitt, P. L., Pittman, P. R., & Bavari, S. (2006). Dominance of human innate immunity responses in primary *Francisella tularensis* live vaccine strain vaccination. *Journal of Allergy and Clinical Immunology*, *117*(5), 1186-1188. doi:10.1016/j.jaci.2006.01.044

Fuller, C. L., Brittingham, K. C., Porter, M. W., Hepburn, M. J., Pettit, P. L., Pittman, P. R., & Bavari, S. (2007). Transcriptome analysis of human immune responses following live vaccine strain (LVS) *Francisella tularensis* vaccination. *Molecular Immunology*, *44*, 3173-3184. doi:10.1016/j.molimm.2007.01.037

Gaucher, D., Therrien, R., Kettaf, N., Angermann, B. R., Boucher, G., Filali-Mouhim, A., … Sekaly, R.-P. (2008). Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *Journal of Experimental Medicine*, *205*(13), 3119-3131. doi:10.1084/jem.20082292

Giefing-Kröll, C., Berger, P., Lepperdinbger, G., & Grubeck-Loebenstein, B. (2015). How sex and age affect immune responses, susceptibility to infections, and response to vaccination. *Aging Cell 14*, 309–321. doi:10.1111/acel.12326

Gene Ontology (GO). (2012). Biological process. Retrieved from

http://www.geneontology.org/

Gerriets, V. A., Kishton, R. J., Johnson, M. O., Cohen, S., Siska, P. J.,Nichols, A. G., … Rathmell, J. C. (2016). Foxp3 and toll-like receptor signaling balance Treg cell anabolic metabolism for suppression. *Nature Immunology, 17*(12),1459-1466. doi:10.1038/ni.3577

Ghosh, A., & Dar, L. (2015). Dengue vaccines: Challenges, development, current status, and prospects. *Indian Journal of Medical Microbiology, 33* (1), 3-15. doi:2015/33/1/3/148369

Grabenstein, J.D., Pittman, P. R., Greenwood, J. T., & Engler, R. J. (2006). Immunization to protect the US Armed Forces: heritage, current practice, and prospects. *Epidemiology Reviews, 28,* 3-26.

Gubler, D. J. (2002). The global emergence/resurgence of arboviral disease as public health problems. *Archives of Medical Research, 33,* 330-342.

Gubbles-Bupp, M. R. (2015). Sex, the aging immune system, and chronic disease. *Cellular Immunology, 294*(2), 102-10. doi:10.1016/j.cellimm.2015.02.002

Guo, H., Callaway, J. B., & Ting, J. P. Y. (2015). Inflammasomes: mechanism of action, role in disease, and therapeutics. *Nature Medicine, 21,* 677–687. doi:10.1038/nm.3893

HealthMap (2016). Retrieved from http://www.healthmap.org/en/

Hoft, D. F., Brusic, V., & Sakala, I. G. (2011). Optimizing vaccine development. *Cellular Microbiology, 13*(7), 934-942. doi:10.1111/j.1462-5822.2011.01609.x

Horzempa, J., O'Dee, D. M., Stolz, D. B., Franks, J. M., Clay, D., & Nau, G. J. (2011).

Invasion of Erythrocytes by *Francisella tularensis*. *Journal of Infectious Diseases, 204*, 51-59. doi:10.1093/infdis/jir221

Hotez, P. J., Pecoul B., Rijal, S., Boehme, C., Aksoy, S., Malecela, M., … Reeder, J. C. (2016). Eliminating the neglected tropical diseases: Translational science and new technologies. *PLOS Neglected Tropical Diseases, 10*(3), e0003895. DOI: 10.1371/journal.pntd.0003895

Huynh, A., DuPage, M., Priyadharshini, B., Sage, P. T., Quiros, J., Borges, C. M., … Turka, L. A. (2015). The phosphatase PTEN-mediated control of PI-3 kinase in Tregs cells maintains homeostasis and lineage stability. *Nature Immunology, 16*(2), 188-196. doi:10.1038/ni.3077

Ingenuity Pathway Analysis (IPA). (2016). Overview of IPA software. Retrieved from http://www.ingenuity.com/products/pathways_analysis.html

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., & Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics, 4*, 249-64. doi:10.1093/biostatistics/4.2.249

Joehanes, R., Johnson, A. D., Barb, J. J., Raghavachari, N., Liu, P., Woodhouse, K. A., … Levy, D. (2012). Gene expression analysis of whole blood, peripheral blood mononuclear cells, and lymphoblastoid cell lines from the Framingham Heart Study. *Physiological Genomics*, *44*(1), 59–75. http://doi.org/10.1152/physiolgenomics.00130.2011

Kaufmann, A. F., Meltzer, M. I., & Schmid, G. P. (1997). The economic impact of a

bioterrorist attack: Are prevention and postattack intervention programs justifiable? *Emerging Infectious Diseases, 3*(2), 83-94. doi:10.3201/eid0302.970201

Katz, J., Zhang, P., Martin, M., Vogel, S. N., & Michalek, S. M. (2006). Toll-like receptor 2 is required for inflammatory responses to *Francisella tularensis* LVS. *Infection and Immunity, 74*, 83-94. doi:10.1128/IAI.74.5.2809-2816.2006

Keim, P., Johansson, A., & Wagner, D. M. (2007). Molecular epidemiology, evolution, and ecology of Francisella. *Annals of the New York Academy of Science, 1105*, 30-66. doi:10.1196/annals.1409.011

Kellam, P. (2006). Attacking pathogens through their hosts. *Genome Biology, 7*(201), 1-8. doi:10.1186/gb-2006-7-1-201

Kellam, P. & Weiss, R. A. (2006). Infectogenomics: Insights from the host genome into infectious diseases. *Cell, 124*, 695-697. doi:10.1016/j.cell.2006.02.003

Kennedy, R. B., Ovsyannikova, I. G., Lambert, N. D., Haralambieva, I. H., & Poland, G. A. (2014). The personal touch: Strategies toward personalized vaccines and predicting immune responses to them. *Expert Reviews Vaccines, 13*(5), 657-669. doi: 10.1586/14760584.2014.905744

Klein, S. L., Jedlicka, A., & Pekosz, A. (2010). The Xs and Y of immune responses to viral vaccines. *Lancet Infectious Diseases 10*, 338-349. doi:10.1016/S1473-3099(10)70049-9

Klein, S. L., & Poland, G. A. (2013). Personalized vaccinology: One size and dose might not fit both sexes. *Vaccine, 31*, 2599-2600. doi:10.1016/j.vaccine.2013.02.070

Klein, S. L., Marriott, I., & Fish, E. N. (2015). Sex-based differences in immune function and responses to vaccination. *Transactions of the Royal Society of Tropical Medicine and Hygiene 109*, 9-15. doi:10.1093/trstmh/tru167

Krueger, C. & Tian, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biological Research for Nursing, 6*(2), 151-157. doi:10.1177/1099800404267682

Kubes, V. & Rios, F. A. (1939). The causative agent of infectious equine encephalomyelitis in Venezuela. *Science, 90*, 20-21. doi:10.1126/science.90.2323.20

Li, S., Nakaya, H. I., Kazmin, D. A., Oh, J. Z., & Pulendran, B. (2013). Systems biological approaches to measure and understand vaccine immunity in humans. *Seminars in Immunity, 25* (3), 209-218. doi:10.1016/j.smim.2013.05.003

Li, Z. K., Nie, J. J., Li J., & Zhuang, H. (2013). The effect of HLA on immunological response to hepatitis B vaccine in healthy people: A meta-analysis. *Vaccine, 31*, 4355-4361. doi:10.1016/j.vaccine.2013.06.108

Li, S., Rouphael, N., Duraisingham, S., Romero-Steiner, S., Presness, S., Davis, C., … Pulendran, B. (2014). Molecular signatures of antibody responses derived from a systems biology study of 5 human vaccines. *Nature Immunology, 15*(2), 195-204. doi: 10.1038/ni.2789

Li, S., Zhu, M., Pan, R., Fang, T., Cao, Y.-Y., Chen, S., … Guo, D. (2016). The tumor suppressor PTEN has a critical role in antiviral innate immunity. *Nature Immunology*, *17*(3), 241-251. doi: 10.1038/ni.3311

Lie, B. A., & Thorsby, E. (2005). Several genes in the extended human MHC contribute to predisposition to autoimmune diseases. *Current Opinion in Immunology, 17*, 526–531. doi:10.1016/j.coi.2005.07.001

Long, D., Deng, X., Singh, P., Loeb, M., Lauring, A. S., & Seielstad, M. (2016). Identification of genetic variants associated with susceptibility to West Nile virus neuroinvasive disease. *Genes and Immunity, 17*(5), 298-304. doi: 10.1038/gene.2016.21

M. D. Anderson Bioinformatics. (2016). Sample size for microarray experiments. Retrieved from http://bioinformatics.mdanderson.org/MicroarraySampleSize/

Mailles, A., & Vaillant, V. (2014). 10 years of surveillance of human tularemia in France. *Eurosurveillance, 19*(45), 1-11. doi:http://dx.doi.org/10.2807/1560-7917.ES2014.19.45.20956

McClain, D. J., Pittman, P. R., Ramsburg, H. H., Nelson, G. O., Rossi, C. A., Mangiafico, J. A., ... Malinoski, F. J. (1998). Immunologic interference from sequential administration of live attenuated alphavirus vaccines. *Journal of Infectious Diseases, 177*(3), 634-641. doi:10.1086/514240

de la Monte S., Castro, F., Bonilla, N. J., Gaskin de Urdaneta, A., & Hutchins, G. M. (1985). The systemic pathology of Venezuelan equine encephalitis virus infection in humans. *American Journal of Tropical Medicine and Hygiene, 34*(1), 194-202.

Mosaad, Y. M., Auf, F. A., Metwally, S. S., Elsharkawy, A. A., El-Hawary, A. K., Hassan, R. H., … El-Chennawi, F. A. (2012). HLA-DQB1* alleles and genetic susceptibility to type 1diabetes mellitus. *World Journal of Diabetes, 3*(8), 149-

155. doi:10.4239/wjd.v3.i8.149

Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71.

Nakaya, H. I., Li, S., & Pulendran, B. (2011). Systems vaccinology: learning to compute the behavior of vaccine induced immunity. *Wiley Interdisciplinary Reviews, Systems Biology and Medicine, 4*(2), 193-205. doi: 10.1002/wsbm.163

Nakazawa, Y., Williams, R., Peterson, A. T., Mead, P., Staples, E., & Gage, K. L. (2007). Climate change effects on plague and tularemia in the United States. *Vector Borne Zoonotic Diseases, 7*(4), 529-40. doi:10.1089/vbz.2007.0125

Nakazawa, Y., Williams, R., Peterson, A. T., Mead, P. S., Kugeler, K. J., & Petersen, J. M. (2010). Ecological Niche Modeling of *Francisella tularensis* Subspecies and Clades in the United States. *American Journal of Tropical Medicine and Hygiene, 82*(5), 912-918. doi:10.4269/ajtmh.2010.09-0354

Narwaney, K. J., Glanz, J. M., Norris, J. M., Fingerlin, T. E., Hokanson, J. E., Rewers, M.,& Hambidge, S. J. (2013). Association of HLA class II genes with clinical hyporesponsiveness to trivalent inactivated influenza vaccine in children. *Vaccine, 31,* 1123-1128. doi:10.1016/j.vaccine.2012.12.026

National Library of Medicine. (2016a). HLA DRB1 gene. Retrieved from https://ghr.nlm.nih.gov/gene/HLA-DRB1.

National Library of Medicine. (2016b). HLA DQB1 gene. Retrieved from https://ghr.nlm.nih.gov/gene/HLA-DQB1.

Navarro, J. C., Medina, G., Vasquez, C., Coffey, L. L., Wang, E., Suarez, A., … Weaver,

S. C. (2005). Postepizootic persistence of Venezuelan equine encephalitis virus, Venezuela. *Emerging Infectious Diseases, 11*(12), 1907-1915. doi:10.3201/eid1112.050533

Nara, P. L., Tobin, G. J., Chaudhuri, A. R., Trujillo, J. D., Lin, G., Cho, M. W., ... Wingreen, N. S. (2010). How can vaccines against influenza and other viral diseases be made more effective? *PLoS Biology, 8* (12), 1-4. doi:10.1371/journal.pbio.1000571

Nelson, C., Kugeler, K., Petersen, C., & Mead, P. (2013). Tularemia—United States, 2001-2010. *Morbidity and Mortality Weekly, 62*(47), 963-966.

National Notifiable Disease Surveillance System. (2016). Nationally Notifiable Conditions. Retrieved from http://www.cdc.gov/osels/ph_surveillance/nndss/phs/infdis.htm

Operational Biosurveillance (2010). Venezuelan equine encephalitis. Retrieved from http://biosurveillance.typepad.com/biosurveillance/venezuelan-equine-encephalitis/

Ovsyannikova, I. G., Jacobson, R. M., Ryan, J. E., Dhiman, N., Vierkant, R. A., & Poland, G. A. (2007). Relationship between HLA polymorphisms and gamma interferon and interleukin-10 cytokine production in healthy individuals after rubella vaccination. *Clinical and Vaccine Immunology, 14*(2), 115-122. doi:10.1128/CVI.00247-06

Ovsyannikova, I. G., Jacobson, R. M., Dhiman, N., Vierkant, R. A., Pankratz, V. S., & Poland, G. A. (2008). Human leukocyte antigen and cytokine receptor gene

polymorphisms associated with heterogeneous immune responses to mumps viral vaccine. *Pediatrics, 121*(5), e1091-1099. doi:10.1542/peds.2007-1575

Oyston, P. C., Sjostedt, A., & Titball, R. W. (2004). Tularaemia: bioterrorism defence renews interest in Francisella tularensis. *Nature Reviews in Microbiology, 2*(12), 967-978. doi:10.1038/nrmicro1045

Ovsyannikova, I. G., Pankratz, V. S., Vierkant, R. A., Pajewski, N. M., Quinn, C. P., Kaslow, R. A., … Poland, G. A. (2013). Human leukocyte antigens and cellular immune responses to anthrax vaccine adsorbed. *Infection and Immunity, 81*(7), 2584-2591. doi:10.1128/IAI.00269-13

Ovsyannikova, I. G., Pankratz, V. S., Salk, H. M., Kennedy, R. M., & Poland, G. A. (2014). HLA alleles associated with the adaptive immune response to smallpox vaccine: A replication study. *Human Genetics, 133*(9), 1083-1092. doi:10.1007/s00439-014-1449-x

Paranavitana, C., Pittman, P., Velauthapillai, M., & DaSilva, L. (2008a). Temporal cytokine profiling of *Francisella tularensis*-infected human peripheral blood mononuclear cells. *Journal of Microbiology, Immunology, and Infection, 41*(3), 192-199.

Paranavitana, C., Pittman, P., Velauthapillai, M., Zelazowska, E., & DaSilva, L. (2008b). Transcriptional profiling of *Francisella tularensis* infected peripheral blood mononuclear cells: a predictive tool for tularemia. *FEMS Immunology and Medical Microbiology, 54*, 98-103. doi: 10.1111/j.1574-695X.2008.00456.x

Paranavitana, C., Zelazowska, E., DaSilva, L., Pittman, P. R., & Nikolich, M. (2010).

Th17 cytokines in recall responses against *Francisella tularensis* in humans. *Journal of Interferon & Cytokine Research, 30*(7), 471-476. doi:10.1089/jir.2009.0108

Paranavitana, C., DaSilva, L., Vladmirova, A., Pittman, P. R., Velauthapillai, M., & Nikolich, M. (2014). Transcriptional profiling of recall responses *to Francisella tularensis* live vaccine strain. *Pathogens and Diseases, 70,* 141-152. doi:10.1111/2049-632X.12113

Partek, Inc. (2016). Partek Genomics Suite software. Retrieved from http://www.partek.com/pgs.

Pawelec, G., Larbi, A., & Derhovanessian, E. (2010). Senescence of the human immune system. *Journal of Comparative Pathology, 142,* S39-S44. doi:10.1016/j.jcpa.2009.09.005

Pittman, P.R., Liu, C. T., Cannon, T. L., Mangiafico, J. A., & Gibbs, P. H. (2009). Immune interference after sequential alphavirus vaccine vaccinations. *Vaccine, 27*(36), 4879-4882. doi:10.1016/j.vaccine.2009.02.090

Plotkin, S. A. (2003). Vaccines, vaccination, and vaccinology. *Journal of Infectious Diseases, 187,* 1349-1359. doi:10.1086/374419

Plotkin, S. A. (2008). Vaccines: correlates of vaccine-induced immunity. *Clinical Infectious Diseases, 47*(3), 401-409. doi:10.1086/589862

Poland, G. A., Ovsyannikova, I. G., & Jacobson, R. M. (2009). Application of pharmacogenomics to vaccines. *Pharmacogenomics, 10,* 837-852. doi:10.2217/pgs.09.25

Pulendran, B., & Ahmed, R. (2011). Immunological mechanisms of vaccination. *Nature Immunology*, *12*(6), 509-517.

Pulendran, B., Oh, J. Z., Nakaya, H., Ravindran, R., & Kazmin, D. A. (2013). Immunity to viruses: Learning from successful human vaccines. *Immunology Reviews, 255* (1), 243-255. doi:10.1111.imr.12099

Querec, T. D., Akondy, R. A., Lee, E. K., Cao, W., Nakaya, H. I., Teuwen, D., … Pulendran, B. (2008). Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature Immunology, 10*(1), 116-125. doi: 10.1038/ni.1688

Reed, S. G., Hsu, F.-C., Carter, D., & Orr, M. T. (2016). The science of vaccine adjuvants: advances in TLR4 ligand adjuvants. *Current Opinion in Immunology, 41*, 85-90. doi:10.1016/j.coi.2016.06.007

Reichert, E., Clase, A., Bacetty, A., & Larsen J. (2009). Alphavirus antiviral drug development: scientific gap analysis and prospective research areas. *Biosecurity and Bioterrorism, 7*(4), 413-427. doi:10.1089/bsp.2009.0032

Rivas, F., Diaz, L. A., Cardenas, V. M., Daza, E., Bruzon, L., Alcala, A., ... Tsai, T. (1997). Epidemic Venezuelan equine encephalitis in La Guarija, Columbia, 1995. *Journal of Infectious Disease, 175*, 828-832. doi:10.1086/513978

Roskoski, R. (2016). Janus kinase (JAK) inhibitors in the treatment of inflammatory and neoplastic diseases. *Pharmacological Research, 111*, 784-803. doi:10.1016/j.phrs.2016.07.038

Rossow, H., Forbes, K. M., Tarkka, E., Kinnunen, P. M., Hemmilä, H., Huito, O., …

Vapalahti, O. (2014). Experimental Infection of voles with Francisella tularensis indicates their amplification role in tularemia outbreaks. *PLoS One, 9*(10), e108864. doi: 10.1371/journal.pone.0108864

Ryden, P., Sjostedt, A., & Johansson, A. (2009). Effects of climate change on tularaemia disease activity in Sweden. *Global Health Action, 11*, 1-7. doi: 10.3402/gha.v2i0.2063

Sahu, S. P., Pederson, D. D., Jenny, A. J., Schmitt, B. J., & Alstad, A. D. (2003). Pathogenicity of a Venezuelan equine encephalomyelitis serotype IE virus isolate for ponies. *American Journal of Tropical Medicine and Hygiene, 68*(4), 485-494.

Scherer, C. A., Magness, C. L., Steiger, K. V., Poitinger, N. D., Caputo, C. M., Miner, D. G., …Iadonato, S. P. (2007). Distinct gene expression profiles in peripheral blood mononuclear cells from patients infected with Vaccinia virus, Yellow Fever 17D virus, or upper respiratory infections. *Vaccine, 25*(35), 6458-6473. doi:10.1016/j.vaccine.2007.06.035

Schwartz, J. T., Barker, J. H., Kaufman, J., Fayram, D. C., McCracken, J. M., & Allen, L. A. (2012). Francisella tularensis inhibits the intrinsic and extrinsic pathways to delay constitutive apoptosis and prolong human neutrophil lifespan. *Journal of Immunology*, *188*(7), 3351-3363. doi:10.4049/jimmunol.1102863

Seib, K. L., Dougan, G., & Rappuoli, R. (2009). The key role of genomics in modern vaccine and drug design for emerging infectious diseases. *PLoS Genetics, 5*(10), e1000612 1-8. doi:10.1371/journal.pgen.1000612

Staples, J. E., Kubota, K. A., Chalcraft, L. G., Mead, P. S., & Petersen, J. M. (2006).

Epidemiologic and molecular analysis of human tularemia, United States, 1964-2004. *Emerging Infectious Diseases, 12* (7), 1113-1118. doi:10.3201/eid1207.051504

Statistical Applications Software (SAS) Institute Inc. (2011). SAS/STAT® 9.3 User's Guide. Cary, NC: SAS Institute Inc.

Tamhane, A. C., & Dunlop, D. D. (2000). Statistics and Data Analysis from Elementary to Intermediate. Prentice Hall. pp 473-474.

Tärnvik, A., Priebe, H.-S., & Grunow, R. (2004). Tularemia in Europe: An ecological overview. *Scandinavian Journal of Infectious Disease, 36*, 350-355. doi:10.1080/00365540410020442

Tian, X., & Feig, L. A. (2001). Basis for signaling specificity difference between Sos and Ras-GRF guanine nucleotide exchange factors. *Journal of Biological Chemistry, 276*(14), 47248-47256. doi:10.1074/jbc.M107407200

Ulu-Kilic, A., & Doganay, M. (2014). An overview: Tularemia and travel medicine. *Travel Medicine and Infectious Disease, 12,* 609-616. doi:10.1016/j.tmaid.2014.10.007

Venner, J. M., Famulski, K.S., Badr, D., Hidalgo, L. G., Chang, J., & Halloran, P. E. (2014). Molecular landscape of T cell-mediated rejection in human kidney transplants: prominence of CTLA4 and PD ligands. *American Journal of Transplantation, 14*(11), 2565-1576. doi:10.1111/ajt.12946

Villari, P., Spielman, A., Komar, N., McDowell, M., & Timperi, R. J. (1995). The economic burden imposed by a residual case of eastern encephalitis. *American*

Journal of Tropical Medicine and Hygiene, 52(1), 8-13.

Wang, B. & Xi, Y. (2013). Challenges for MicroRNA Microarray Data Analysis. *Microarrays, 2*(2), 34-50. DOI:10.3390/microarrays2020034.

Weatherhead, J. E., Miller, V. E., Garcia, M. N., Hasbun, R., Salazar, L., Dimachkie, M. M., & Murray, K. O. (2015). Long-Term Neurological Outcomes in West Nile Virus–Infected Patients: An Observational Study. *American Journal of Tropical Medicine and Hygiene, 92*(5), 1006-1012. doi:10.4269/ajtmh.14-0616.

Weaver, S. C., & Reisen, W. K. (2010). Present and future arboviral threats. *Antiviral Research 85*, 328–345. doi:10.1016/j.antiviral.2009

Wolfe, D. N., Florence, W., & Bryant, P. (2013). Current biodefense vaccine programs and challenges. *Human Vaccines and Immunotherapeutics, 9*(7), 1-7.

Wolfe, D. N., Heppner, D. G., Gardner, S. N., Jaing, C., Dupuy, L. C.,Schmaljohn, C. S., & Carlton, K. (2014). Perspective piece: Current thinking for the development of a trivalent alphavirus vaccine for human use. *American Journal of Tropical Medicine and Hygiene, 91*(3), 442-450. doi:10.4269/ajtmh.14-0055.

Yang, Q., Modi, P., Newcomb, T., Queva, C., & Gandhi, V. (2015). Idelalisib: First-in-class PI3K delta inhibitor for the treatment of chronic lymphocytic leukemia, small lymphocytic leukemia, and follicular lymphoma. *Clinical Cancer Research, 21*(7), 1537-1542. doi:10.1158/1078-0432.CCR-14-2034

Appendix A: Memorandum for Record, Determination of Not Human Subject Research

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH INSTITUTE OF INFECTIOUS DISEASES
1425 PORTER STREET
FORT DETRICK, MARYLAND 21702-5011

REPLY TO ATTENTION OF

MCMR-UIZ-H

11 March 2016

MEMORANDUM FOR RECORD

SUBJECT: Determination of Not Human Subject Research for Project Entitled, "Genome Wide Transcriptome Analysis of Human Responses to Vaccination with Live-Attenuated Francisella tularensis or Live-Attenuated Venezuelan Equine Encephalitis Virus" (FY16-17) Submitted by Rebecca Erwin-Cohen

1. Reference Application for Research Determination, 8 March 2016, submitted by Rebecca Erwin-Cohen and memorandum to OHU&E, Re: FY-5-01 and FY06-17, 9 March 2016, submitted by Phillip Pittman, MD, MPH.

2. The USAMRIID Investigator will conduct a meta-analysis on previously generated transcriptome data conducted under human use protocols FY05-01 and FY06-17, which are now closed. The Investigator will discuss her results as part of a Ph.D. dissertation research project which will be used for publication.

3. The data sets are coded with no personal identifiers contained within the bioinformatics data files. The PI who led the human use protocols from which the data will be analyzed has granted the Investigator access to the de-identified database.

4. The Exempt Determination Official (EDO) has made the determination that the project qualifies as Not Human Subject Research, as it does not meet the definition of "human subject" under 32 CFR 219.102(f)(1) nor (f)(2).

5. Please note that changes to the PI or research, and/or unanticipated problems must be reported immediately to the Office of Human Use and Ethics by phone: 301-619-6630 or email: usarmy.detrick.medcom-usamrlid.mbx.ohue@mail.mil.

6. The point of contact is Pamela Barretto-Jones, RN, MSN, CIP at 301-619-6630 or pamela.b.barretto-jones.ctr@mail.mil@mail.mil.

MD

Encl

ARTHUR O. ANDERSON
USAMRIID EDO, OHU&E