2016

Using the host immune response to hemorrhagic fever Viruses to understand pathogenesis and improve diagnostics

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

AND

COLLEGE OF ENGINEERING

Dissertation

USING THE HOST IMMUNE RESPONSE TO HEMORRHAGIC FEVER VIRUSES TO UNDERSTAND PATHOGENESIS AND IMPROVE DIAGNOSTICS

by

IGNACIO SANCHEZ CABALLERO

M.S., Complutense University of Madrid, Spain, 2011
B.S., Pontificia University of Salamanca, Spain 2006

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader _____

John H. Connor, Ph.D.
Associate Professor of Microbiology

Second Reader _____

W. Evan Johnson, Ph.D.
Assistant Professor of Medicine

USING HOST IMMUNE RESPONSES TO HEMORRHAGIC FEVER VIRUSES TO IMPROVE DIAGNOSTICS AND UNDERSTAND PATHOGENESIS

(Order No.)

IGNACIO SANCHEZ CABALLERO

Boston University Graduate School of Arts and Sciences

and

College of Engineering, 2016

Major Professor: John H. Connor, Associate Professor of Microbiology

ABSTRACT

Hemorrhagic fever viruses cause severe infections characterized by a hyperactive immune response that often leads to multiorgan failure and death. Current diagnostic tests are based on detecting viral proteins and nucleic acids in the blood. These are late-stage events during infection, which makes it impossible to perform a diagnosis before they are present in the collected sample. In this thesis, I explore an alternative approach using the transcriptional changes of circulating immune cells during the early stages of infection to identify unique markers of viral infection. The main advantage of this method is that it can be used to identify highly pathogenic viruses before standard detection methods become effective.

I initially used RNA sequencing data to compare the host patterns of expression of macaques infected with either Lassa virus or Marburg virus, two related hemorrhagic fever viruses. I identified a set of genes that quickly become upregulated after a viral infection and remain highly expressed throughout the entire disease course, irrespective

of the specific virus that caused the infection. I was also able to identify a set of biomarker genes that follow unique patterns of expression depending on the type of infection. I used an independent dataset to validate the potential of these genes to be used as biomarkers of infection. Additionally, I compared these results to the patterns of expression of macaques infected with Ebola virus, looking at multiple experimental conditions, tissues and routes of infection. Finally, I validated the host patterns of expression using two independently generated datasets corresponding to infection by different strains of arenaviruses and filoviruses.

Studying the host immune response has the potential to improve the diagnosis of viral hemorrhagic fevers and other diseases. It can also accelerate our efforts to understand the underlying molecular mechanisms that lead to pathogenesis and severe disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BDBV ..... Bundibugyo virus

BLAST ..... Basic Local Alignment Search Tool

BSL ..... BioSafety Level

CDC ..... Centers for Disease Control

CHPV ..... Chapare virus

CPE ..... Cytopathic Effect

CPM ..... Counts Per Million

DRC ..... Democratic Republic of Congo

EBOV ..... Ebola virus

EHF ..... Ebola Hemorrhagic Fever

ELISA ..... Enzyme-Linked Immunosorbent Assay

ENCODE ..... Encyclopedia of DNA Elements

EVD ..... Ebola Virus Disease

FDA ..... Food and Drug Administration

GEO ..... Gene Expression Omnibus

IFN ..... Interferon

IFNAR ..... Interferon Alpha/Beta receptor

ISG ..... Interferon Stimulated Gene

LASV ..... Lassa virus

LCMV ..... Lymphocytic choriomeningitis virus

LLOV ..... Lloviu virus

LUJV ..... Lujo Virus

MARV ..... Marburg virus

MOI ..... Multiplicity of Infection

MPXV ..... Monkeypox Virus

MVD ..... Marburg virus

NHP ..... Nonhuman Primate

PBMC ..... Peripheral Blood Mononuclear Cells

PCA ..... Primary Component Analysis

PCR ..... Polymerase Chain Reaction

PFU ..... Plaque Forming Units

RAVV ..... Ravn virus

RESTV ..... Reston virus

RT-PCR ..... Reverse Transcription Polymerase Chain Reaction

SUDV ..... Sudan virus

TAFV ..... Taï Forest virus

USAMRIID ..... United States Army Medical Research Institute for Infectious Diseases

VACV ..... Vaccinia Virus

VLP ..... Virus-like protein

VSV ..... Vesicular Stomatitis Virus

WHO ..... World Health Organization

CHAPTER 1

INTRODUCTION

Two opposing forces determine the severity of disease associated with an infection: the ability of a pathogen to replicate and spread inside the host, and the ability of the host's immune system to detect and stop the infection. Pathogenesis—the molecular mechanism that leads to a disease state—can vary widely depending on the causative agent. Some infections are asymptomatic or cause only minor disease, such as those caused by human rhinoviruses and adenoviruses. Other infections are associated with severe clinical symptoms and high fatality rates, such as those caused by hemorrhagic fever viruses. Even among hemorrhagic fever viruses, severity rates vary significantly, and can range from less than 1% to 90% case fatality rates. The reasons for these variations are not well understood, but it has been proposed that severe disease is caused by an overactive host response that is ineffective against infection and leads to aggravated symptoms (Günther & Lenz, 2004; Mohamadzadeh, Chen, & Schmaljohn, 2007).

Ebola virus, Marburg virus and Lassa virus are among the most severe pathogens that are thought to be associated with a dysregulated immune response. They cause severe febrile illnesses that often lead to bleeding disorders and death, they have no associated FDA-approved treatments or vaccines, and they pose a critical risk to public health. The effect that these viruses have on the immune response has been partially studied in cell culture or animal models, but the majority of studies were end-point examinations or they were limited to a handful of genes. In this dissertation, I follow a

multiple-timepoint comprehensive approach to describe the transcriptional changes that circulating immune cells undergo after the host becomes infected by either Lassa virus, Marburg virus or Ebola virus, and how these changes vary depending on the infecting pathogen and the stage of disease progression.

This chapter summarizes the virology, epidemiology, and pathogenesis of these viruses, as well as the clinical manifestations, and host immune responses that take place after infection. The two filoviruses, Ebola virus and Marburg virus, are presented first, followed by the arenavirus, Lassa virus.

1.1 Filoviruses

Ebola virus and Marburg virus are members of the *Filoviridae* family. They have single-stranded negative-sense RNA genomes that encode seven structural proteins (nucleoprotein, NP; VP35; VP40; glycoprotein, GP; VP30; VP24; and RNA-dependent RNA polymerase, L) as well as two nonstructural proteins (soluble glycoprotein, sGP; and small soluble glycoprotein, ssGP). Eight different filoviruses have been identified (Kuhn et al., 2012): five in the *Ebolavirus* genus (Ebola virus, EBOV; Sudan virus, SUDV; Taï Forest virus, TAFV; Bundibugyo virus, BDBV; and Reston virus, RESTV), two in the *Marburgvirus* genus (Marburg virus, MARV; and Ravn virus, RAVV), and one in the recently discovered *Cuevavirus* genus (Lloviu virus, LLOV).

Before 2014, Ebola virus outbreaks have been documented in Gabon, Congo, and the DRC (Figure 1.1A). The largest outbreaks of this virus species were Yambuku (1976, 318 cases, 88% case fatality rate) and Kikwit (1995, 315 cases, 79% case fatality rate),

both in the DRC. Sudan virus outbreaks have only occurred in South Sudan (Nzara, 1976, 284 cases, 53% CFR) and Uganda (Gulu, 2000, 425 cases, 53% CFR). Bundibugyo virus outbreaks have taken place in the border between Uganda and the DRC, notably in 2007 (149 cases, 25% CFR). As seen in Figure 1.1A, all pre-2014 Ebola cases occurred in the eastern equatorial region of Africa, with the exception of a single case reported in South Africa, as well as the only known case of Tai Forest virus case (Taï Forest National Park, 1994), which took place in Côte d'Ivoire and was not lethal.

After 2014, the geographical distribution of Ebola virus disease changed dramatically with a major epidemic that began in Guinea and spread through several West African countries, including Liberia and Sierra Leone, and which resulted in more than 27,000 cases and at least 11,000 deaths (41% CFR) (Figure 1.1B) (Centers for Disease Control and Prevention, 2015a). This was the largest disease outbreak ever recorded, and the threat of a possible spread beyond West Africa has highlighted the importance of studies that focus on understanding the pathogenesis and the development of potential treatment and prevention strategies.

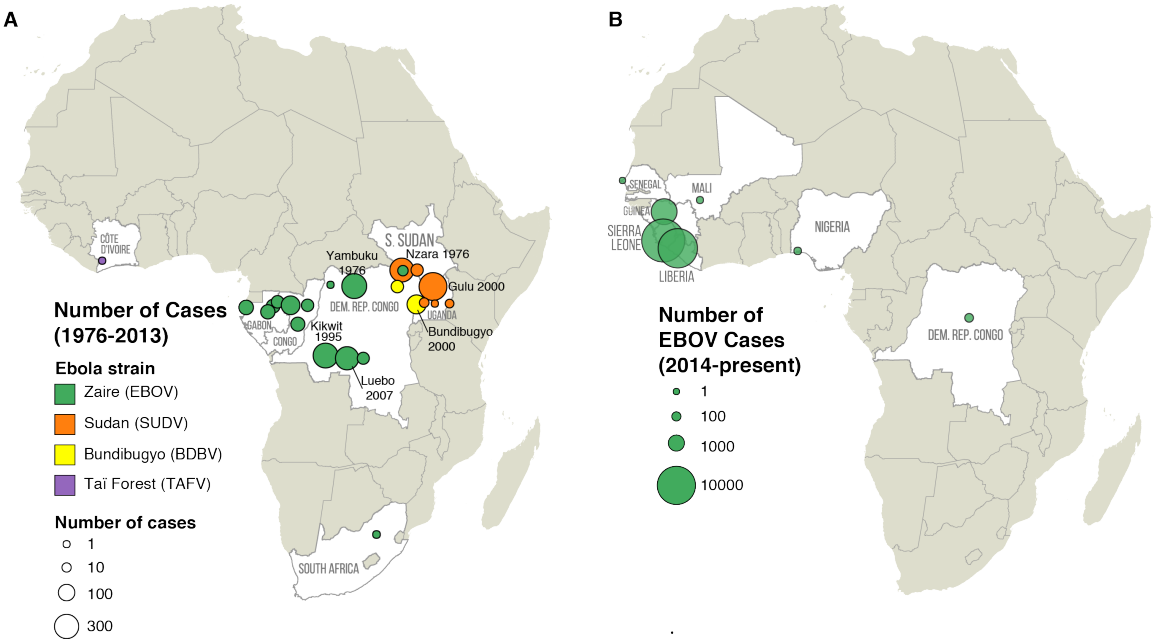


Figure 1.1. Distribution map of Ebola virus disease in Africa, before and after 2014. Circles represent individual outbreaks. Circle radius is proportional to the number of reported cases. Circle color represents different *Ebolavirus* genera. (A) Ebola virus outbreaks before 2014 took place in Gabon, Congo, and the former Zaire (now DRC). Sudan virus outbreaks have only taken place in South Sudan and Uganda. Bundibugyo virus outbreaks have taken place in the border between Uganda and the DRC. The only reported case of Tai Forest virus took place in the Taï Forest National Park in the Ivory Coast. (B) In 2014, a large Ebola virus epidemic spread throughout the West African countries of Guinea, Liberia and Sierra Leone.

Similar to the pre-2014 pattern of Ebola disease, the majority of Marburg outbreaks have taken place in equatorial Africa. The two largest were reported in the DRC (Durba, 1998, 154 cases, 83% case fatality rate) and in Angola (Uige province, 2004, 252 cases, 90% case fatality rate) (Figure 1.2) (Centers for Disease Control and Prevention, 2015b). No Marburg outbreaks have yet been reported in Western Africa, but species distribution frameworks used to model the potential geographic spread of Marburg virus into this region have concluded that this is a realistic scenario (Pigott et al., 2015).

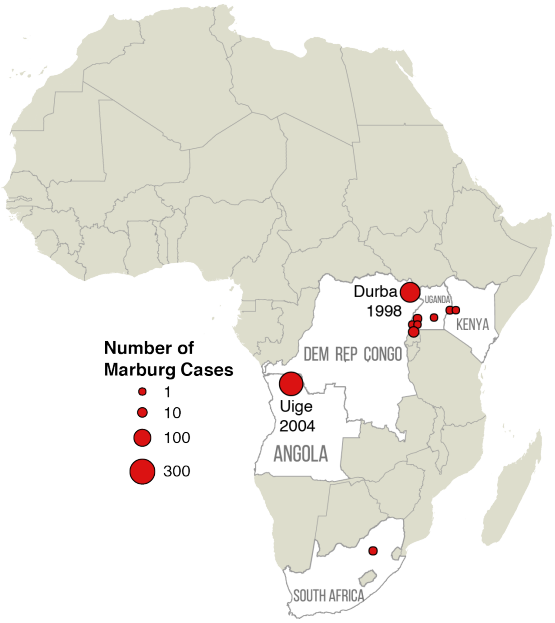


Figure 1.2. Distribution map of Marburg virus disease.
The majority of Marburg outbreaks have taken place in equatorial Africa. Circles represent individual outbreaks. Circle radius is proportional to the number of reported cases.

Bats are believed to be the reservoir species for Ebola virus. Three species of African fruit bats (*Hypsignathus monstrosus, Epomops franqueti, and Myonycteris torquata*) were found to harbor Ebola viral RNA and anti-Ebola antibodies (Eric M Leroy et al., 2005), but isolating live virus from these animals proved unsuccessful. In contrast, ecological investigations in 2007 led to the identification of Egyptian fruit bats (*Rousettus aegyptiacus*) as the likely reservoir species for Marburg virus. This was reported after successfully isolating live virus from bats captured in a Ugandan cave where miners had contracted the disease (Towner et al., 2009).

1.2 Clinical manifestations of Ebola Virus Disease and Marburg Virus Disease

The clinical manifestations and pathogeneses of Ebola Virus Disease (EVD) and Marburg Virus Disease (MVD) are very similar. Infection is thought to occur through close contact with body fluids from infected animals or humans (Rougeron, Feldmann, Grard, Becker, & Leroy, 2015). The mean incubation period—the time between infection and the onset of symptoms—ranges between 2 and 21 days, but is typically closer to 5-10 days. Patients become ill abruptly, with common symptoms including fever, vomiting, diarrhea, and headache (Brauburger, Hume, Mühlberger, & Olejnik, 2012; Feldmann & Geisbert, 2011; Kortepeter, Bausch, & Bray, 2011; Siddhartha Mahanty & Bray, 2004; WHO Ebola Response Team, 2014). Hemorrhagic symptoms vary in severity and are present on one third of patients during the peak of the illness; they include petechiae, a maculopapular rash, ecchymoses and mucosal bleeding (Bwaka et al., 1999; Feldmann & Geisbert, 2011; Rougeron et al., 2015). Laboratory markers of advanced infection include a marked fall in blood lymphocyte and platelet count, and a rise in neutrophils and serum aminotransferase concentrations (Bradfute et al., 2010; Bwaka et al., 1999; Feldmann & Geisbert, 2011; Wauquier, Becquart, Padilla, Baize, & Leroy, 2010). During the late stage of infection, patients with fatal disease experience neurological symptoms (confusion, delirium, seizures), multi-organ failure (pancreas, liver, kidney), hypovolemic shock and diffuse intravascular coagulopathy (Feldmann et al., 2007; C J Peters & LeDuc, 1999). Patients that survive the disease improve around days 7-11 after symptom

onset, as the humoral response develops and the virus is cleared from the blood (Ksiazek et al., 1999; Rougeron et al., 2015).

1.3 Filovirus pathogenesis

The exact mechanism of filovirus pathogenesis is unknown, but there is significant information about how these viruses infect cells. Ebola virus and Marburg virus enter the body through lymphatic or blood vessels and initially target monocytes, macrophages, and dendritic cells (Alves et al., 2010; Thomas W Geisbert, Hensley, et al., 2003; Schnittler & Feldmann, 1998; Ströher et al., 2001). It is thought that these infected cells release large quantities of proinflammatory cytokines and chemokines, like TNF-alpha and IL1-beta, which leads to the recruitment of additional macrophages to the infected area and is thought to result in vascular leakage (Bosio et al., 2003). It also causes neutrophils to invade nearby tissue and cause lesions (Gerszten et al., 1999). Furthermore, macrophages start overexpressing tissue factor, which contributes to the diffuse intravascular coagulopathy that is characteristic of late-stage filovirus infection (S Baize et al., 1999; Schnittler & Feldmann, 1998; Wauquier et al., 2010). An additional mechanism contributing to pathogenesis is a significant decrease in the number of lymphocytes, especially T cells and NK cells (Thomas W Geisbert, Hensley, et al., 2003). These cells are not permissive to filovirus infection, but they undergo massive bystander apoptosis through a mechanism that is not entirely understood (T W Geisbert et al., 2000; Hensley, Alves, et al., 2011; Wauquier et al., 2010). It is thought that these two processes—increased levels of circulating proinflammatory cytokines and lymphocyte

apoptosis—are caused by a dysregulated immune response, which explains the inability of fatal patients to control viral replication (S. Baize, Leroy, Mavoungou, & Fisher-Hoch, 2000; S Baize et al., 1999; Feldmann & Geisbert, 2011; Thomas W Geisbert, Hensley, et al., 2003; Misasi & Sullivan, 2014).

1.4 Host immune response to filovirus infection

Nonfatal filovirus infections are associated with an antibody response directed against NP and VP40, cell-mediated responses to GP, and virus clearance driven by the cytotoxic T-cell response (Paessler & Walker, 2013). In contrast, fatal filovirus infections are characterized by a dysregulated innate immune response and an ineffective adaptive immune response. *In vitro* and *in vivo* studies have shown that filovirus-infected dendritic cells and macrophages produce chemokines, such as CCL3 (MIP-1) and CCL2 (MCP-1), that lead to the recruitment of additional macrophages (Thomas W Geisbert, Hensley, et al., 2003; Gupta, Mahanty, Ahmed, & Rollin, 2001; Hensley, Alves, et al., 2011; Wauquier et al., 2010). Additionally, these cells also secrete inflammatory mediators, such as TNF-alpha, CXCL8, CXCL10, IL2, and IL6 and IL10, in much higher concentrations in fatal than in nonfatal cases (S Baize et al., 1999, 2002; Thomas W Geisbert, Hensley, et al., 2003; Hensley, Young, Jahrling, & Geisbert, 2002; Villinger et al., 1999; Wauquier et al., 2010). T cell-priming by dendritic cells is likely impaired by this uncontrolled release of cytokines and, together with the associated massive bystander lymphocyte apoptosis, is thought to contribute to the lack of an effective adaptive immune response during filovirus infection (Bradfute & Bavari, 2011).

A central player in the innate immune response is the family of cytokines known as Type I interferons. Filoviruses have evolved powerful mechanisms to inhibit their production, as well as to prevent the activation of the interferon signaling pathways. Once secreted, interferon binds to the receptors in neighboring cells (as paracrine signaling), or to those in the same cell (as autocrine signaling) (Christopher F. Basler, 2015), and initiate the JAK-STAT signaling cascade, which results in the upregulation of hundreds of genes (known as interferon-stimulated genes, or ISGs). The proteins that these genes encode collectively improve the cell's ability to inhibit virus replication. Interferon expression is induced by the activation of pattern recognition receptors, like RIG-I (DDX58) or MDA5 (IFIH1). These receptors become activated after sensing RNA products of virus replication, such as 5′ triphosphates and dsRNAs (Kawai & Akira, 2010). A fact that highlights the importance of the interferon pathway is that in wild-type mice, Ebola virus infection doesn't result in disease. However, knockout mice lacking the type I interferon receptor (IFNAR), or the main interferon signaling protein (STAT1), are uniformly susceptible to fatal Ebola infection (Bray, 2001). Ebola virus protein VP35 and VP24, as well as Marburg VP40, are reported to be the main inhibitors of interferon production and signaling. As represented in Figure 1.3A, Ebola VP35 prevents the phosphorylation of IRF3 and IRF7, which are transcription factors that dimerize after becoming phosphorylated, translocate to the nucleus and initiate the synthesis of interferon (C. Basler & Mikulasova, 2003; Cárdenas et al., 2006; T. H. Chang et al., 2009; Hartman, Towner, & Nichol, 2004; Prins, Cárdenas, & Basler, 2009). Additionally, as seen in Figure 1.3B, Ebola VP24 prevents phosphorylated STAT1 from translocating

to the nucleus and inducing the production of ISGs (Reid et al., 2006), while Marburg VP40 prevents the phosphorylation of the interferon receptor-associated kinases JAK1, TYK2, STAT1 and STAT2 (Valmas et al., 2010).

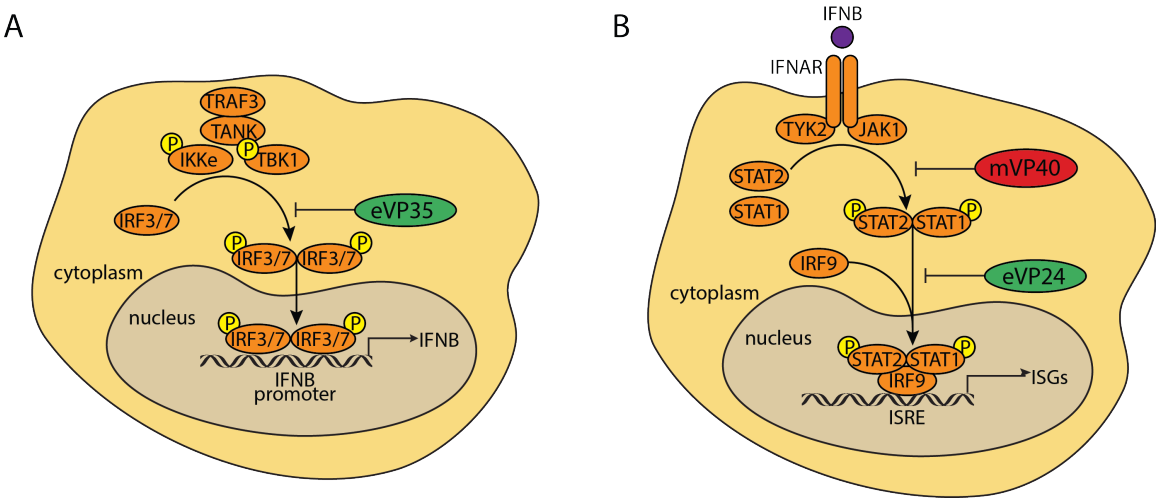


Figure 1.3: Filovirus inhibition of IFN signaling.
(A) Type I interferons are regulated by the transcriptional factors IRF3 and IRF7. When these factors become phosphorylated, they are able to dimerize and translocate to the nucleus to bind the promoter sequence of type I interferon genes (IFNB, in this case). Ebola VP35 inhibits the phosphorylation event that makes this possible. (B) When type I interferons bind the interferon receptor (IFNAR), they cause the phosphorylation of STAT1 and STAT2, which are then able to form a heterodimer along with the IRF9 transcription factor, translocate to the nucleus, and bind the promoter regions (Interferon-Stimulated Response Elements) that control the transcription of ISGs. Marburg VP40 blocks the phosphorylation event, and Ebola VP24 prevents phosphorylated STAT1 from translocating into the nucleus.

Numerous studies have investigated how filovirus infection affects interferon production, with contrasting results (summarized in
Table 1.1). Some *in vitro* studies report that filoviruses completely limit or block interferon production (Bosio et al., 2003; Gupta et al., 2001; Harcourt, Sanchez, & Margaret, 1999; Kash, Mühlberger, & Carter, 2006), while others report that it increases (Thomas W Geisbert, Young, Jahrling, Davis, Larsen, et al., 2003; Hensley et al., 2002).

Similarly, several *in vivo* studies report that filovirus infection is associated with increasing concentrations of interferon in humans, macaques and mice (Thomas W Geisbert, Hensley, et al., 2003; Gupta et al., 2001; Hensley et al., 2002; Siddhartha Mahanty, Gupta, et al., 2003; Villinger et al., 1999; Warfield et al., 2009), while one study reports that such increases were not observed (Wauquier et al., 2010).

Experiment type	**Virus**	**Infected host or cell type**	**Production of Type I interferon**	**Reference**
In vitro	EBOV (check)	Endothelial cells (HUVEC)	NO	Harcourt1999
In vitro	EBOV	Macrophages (human)	NO	Gupta2001
In vitro	EBOV and MARV	Dendritic cells (human)	NO	Bosio2003
In vitro	EBOV	Macrophages and dendritic cells	YES	Hensley2002
In vitro	EBOV	Endothelial cells (HUVEC)	YES	Geisbert2003a
In vitro	Mouse-adapted RAVV	BALB/c mice	YES (only 3 dpi)	Warfield2009
In vivo	EBOV	Human	YES	Villinger1999
In vivo	EBOV	Human	YES (serum)	Gupta2001
In vivo	EBOV	Cynomolgus and rhesus macaques	YES (by 5 dpi, serum)	Hensley2002
In vivo	EBOV	Mice (BALB/c, subcutaneous injection)	YES (by 2 dpi, in liver spleen and serum)	Mahanty2003a
In vivo	EBOV	Cynomolgus macaques	YES (by 5-6 dpi, serum)	Geisbert2003c
In vivo	MARV (Ci67	Cynomolgus macaque (plasma,	YES (by 6 dpi in plasma, 3	Fritz2008

	strain)	spleen)	dpi in spleen)	
In vivo	EBOV	Human	NO	Wauquier2010

Table 1.1 *In vitro* and *in vivo* studies that investigated the presence of interferon after filovirus infection.

These inconsistent results likely have multiple explanations. One may be cell variability: in vitro studies typically derive interferon-producing cells (macrophages and dendritic cells) from monocytes, but these cells may not accurately reproduce the *in vivo* infection scenario; for example, by not producing interferon at natural levels, or doing so intermittently. Another explanation could be that different studies measured different interferon isoforms, and this may lead to different rates of ISG activation. These contradictory findings suggest that a more comprehensive effort to understand the interferon response is necessary.

To this end, in Chapter 3, I investigate the transcriptional patterns of circulating immune cells in a macaque model of Ebola infection. My results show that approximately 50 interferon-stimulated genes show strong increases in expression early after infection, which supports the view that the innate immune response becomes activated during filovirus infection. I also compared two different routes of infection (aerosol and intramuscular) and saw similar patterns of expression. The aerosol model also showed that these innate immune responses take place in primary and secondary target tissues. Lastly, I contrasted these patterns of expression to those of Ebola-vaccinated macaques and found that this treatment rendered the innate immune response almost non-existent. This runs counter to the prevailing hypothesis that innate immune signaling is suppressed during infection.

1.5 Arenaviruses

The geographic region where the 2015 Ebola outbreak took place is also the location where most cases of Lassa fever take place every year. Lassa virus is a member of the *Arenaviridae* family. It has a single-stranded RNA genome composed of two segments. Each RNA segment is ambisense, encoding one gene in the positive sense, and another in the negative sense. The large segment encodes the RNA-dependent RNA polymerase (L), and a zinc-binding protein (Z). The small segment encodes the structural proteins nucleoprotein (NP), and the glycoprotein precursor (GPC); which is later cleaved into the envelope proteins (G1 and G2). Arenaviruses are classified into New World and Old World arenaviruses. New World arenaviruses include viruses common in the Americas, and Old World arenaviruses are found in Africa. The exception is Lymphocytic Choriomeningitis Virus (LCMV), which has a worldwide distribution but is classified as an Old World arenavirus. Hemorrhagic fever viruses in the New World arenaviruses include Junin virus (JUNV), Machupo virus (MACV), Guanarito virus (GTOV), Chapare virus (CHPV), and Sabia virus (SABV). The only two Old World arenaviruses that are known to cause hemorrhagic fevers are Lujo virus (LUJV) and Lassa virus (LASV).

Lassa virus is the causative agent of Lassa fever, a major endemic disease in West Africa, mainly in Guinea, Sierra Leone, Liberia and Nigeria (Figure 1.4). Based on the high rates of Lassa virus seroconversion among Sierra Leonean villagers (5%-20% per year), it was estimated that the virus infects 100,000-300,000 people every year with a case fatality rate of 1%-2% (J. McCormick & Webb, 1987). This rate has been shown to

be much higher among hospital-admitted patients (up to 69%) (S P Fisher-Hoch et al., 1995; Shaffer et al., 2014). The true rate of Lassa Fever infections in West Africa is likely obscured by the lack of systematic surveillance and convenient diagnostics (Juan C Zapata & Salvato, 2015).

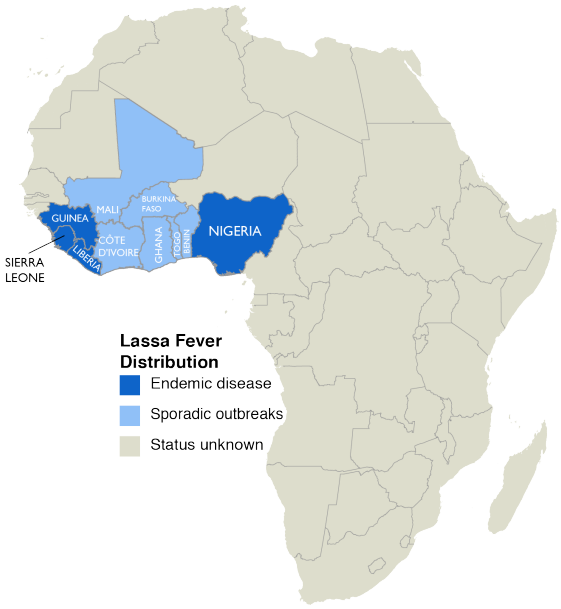


Figure 1.4. Distribution map of Lassa Fever.
Guinea, Sierra Leone, Liberia and Nigeria are associated with large recurring epidemics (dark blue). Areas in light blue report sporadic infections. Elsewhere, the frequency of Lassa fever is unknown (gray).

The reservoir species of Lassa virus is the African rodent *Mastomys natalensis* (Lecompte et al., 2006; Monath, Newhouse, Kemp, Setzer, & Cacciapuoti, 1974), which is ubiquitous throughout sub-Saharan Africa. These rodents live in close proximity to humans. In some provinces in Sierra Leone, up to 90% of rodents caught in houses are *Mastomys*; a large proportion of which carry Lassa virus (Günther & Lenz, 2004). Humans likely become infected by breathing air or consuming food that is contaminated with *Mastomys* urine or other excreta (Monath et al., 1974).

1.6 Clinical manifestations of Lassa Fever

Lassa fever has a longer incubation period than Ebola virus and Marburg virus, and it ranges from 7 to 18 days. Initial symptoms are nonspecific and similar to those of flu. They include fever, weakness, malaise, severe headache, and a sore throat (J. McCormick & Webb, 1987). This lack of distinctive initial symptoms makes the early diagnosis difficult and delays the initiation of treatment (Juan C Zapata & Salvato, 2015). Around 50% of patients also experience nausea and diarrhea. There is no characteristic skin rash and, in contrast with filovirus infection, no coagulation disorders (S P Fisher-Hoch et al., 1985). The liver is the most consistently affected organ in humans, and fatal disease is correlated with increases in the concentration of aspartate transaminase (AST) and alanine transaminase (ALT) in serum (Günther & Lenz, 2004). Bleeding is only seen in 15% to 20% of patients and is mainly limited to mucosal surfaces. In contrast to other hemorrhagic fevers, diffuse intravascular coagulation is rarely seen in patients with Lassa fever (Cummins et al., 1989; Susan P. Fisher-Hoch, McCormick, Sasso, & Craven, 1988).

1.7 Lassa virus pathogenesis

The primary routes of virus entry are thought to be cuts or abrasions on the skin, as well as close contact with secretions or blood from infected patients (J. B. McCormick & Fisher-Hoch, 2002). Macrophages and dendritic cells become infected at the site of viral entry at an early stage and facilitate the dissemination of the virus via blood and

lymph, infecting all the major organs (Walker et al., 1982). Infected macrophages have two detrimental effects: they lead to the formation of blood clots due to overexpressed tissue factor (Thomas W Geisbert, Young, Jahrling, Davis, Kagan, et al., 2003), and they increase vascular leakage due to the uncontrolled release of the inflammatory proteins (Zampieri, Sullivan, & Nabel, 2007).

The broad tropism of the Lassa virus is explained by the ubiquity of its cellular receptors (alpha-dystroglycan (Cao et al., 1998; Kunz et al., 2005) and high-mannose receptors (Shimojima, Stroher, Ebihara, Feldmann, & Kawaoka, 2012)). Similar to filovirus infection, there is strong bystander apoptotic lymphopenia and neutrophilia (J. B. McCormick & Fisher-Hoch, 2002). The late stage of disease is characterized by an uncontrolled rise in virus load in multiple organs.

In survivors, deafness is one of the most common symptoms during the convalescent stage; 30% of hospitalized patients suffer acute loss of hearing in one or both ears (Cummins, 1990). It remains unknown if this damage is a side-effect of viral replication or of the immune response.

1.8 Host immune response to Lassa virus infection

Fatal Lassa fever cases are associated with a dysregulated host innate immune response that is unable to stop viral replication (Johnson et al., 1987; Clarence J Peters & Zaki, 2002). This is believed to take place because dendritic cells, which are primary targets of Lassa virus, become infected and are only able to produce a limited panel of cytokines, their maturation becomes impaired and they are less capable of activating T

cells (Bosio et al., 2003; Lubaki et al., 2013; Siddhartha Mahanty, Hutchinson, et al., 2003; B. Yen, Mulder, Martinez, & Basler, 2014). Additional human studies about the concentrations of cytokines during infection have revealed that low serum levels of the pro-inflammatory chemokines CXCL8 (IL8) and CXCL10 (IP-10) and high levels of the anti-inflammatory cytokine IL6 are strongly correlated with fatal outcome (S Mahanty et al., 2001). In macaques, high levels of type I interferon can be found in the blood of both fatal and nonfatal cases, but while the concentration of interferon decreases quickly in nonfatal cases, it remains high in fatal cases (Sylvain Baize et al., 2009). This high level of interferon is probably unable to mitigate the massive viral replication observed at the terminal stages of disease, and likely contributes to pathogenesis.

Neutralizing antibodies to Lassa GP appear in survivors months after the infection has been resolved (Peter B. Jahrling, Frame, Rhoderick, & Monson, 1985), and it is believed that cell-mediated immunity plays a more important role in Lassa virus clearance than humoral immunity. In macaques, nonfatal cases have high numbers of activated T cells while fatal cases are associated with lower levels of T cell activation and uncontrolled viral replication (Sylvain Baize et al., 2009; Yun & Walker, 2012). In humans, strong memory $CD4^+$ T cell responses against LASV NP and GP have been detected in LASV-seropositive individuals living in West Africa, where Lassa is endemic. This suggests that nonfatal infections are associated with the activation of CD4 T cells (J ter Meulen et al., 2000; Jan Ter Meulen et al., 2004). The failure of the innate immune system to activate a rapid T cell response capable of clearing the infection may lead these cells to contribute to further pathogenesis.

In summary, hemorrhagic fever virus infections are associated with a complex pathogenesis and an immune response that remains only partially understood. The study of individual components of these complex processes can provide useful information, but the main limitations are the number of genes that can be studied simultaneously, or the frequency with which they can be measured throughout infection. Transcriptional profiling is a high-throughput method that makes it possible to accurately quantify the amount of expression of every gene at multiple times post-infection and provides an unbiased view of the gene expression patterns that take place in a given sample.

1.9 Transcriptional profiling

Identifying the transcriptional patterns of host genes during infection has made it possible to increase our understanding of pathogenesis, help us distinguish between severe and asymptomatic patients, and differentiate between bacterial and viral infections (Barrenas et al., 2015; Huang et al., 2011; Kash et al., 2006; a. K. Zaas et al., 2013; Juan C Zapata & Salvato, 2015). The first studies that measured host gene expression changes after a hemorrhagic fever virus infection used northern blot or RT-PCR assays, which limited the number of observations to a handful of genes (S Baize et al., 1999; C F Basler et al., 2000; Harcourt et al., 1999; E M Leroy, Baize, Debre, Lansoud-Soukate, & Mavoungou, 2001; E. Leroy, Baize, & Volchkov, 2000; Ströher et al., 2001). With the arrival of microarrays and high-throughput sequencing technology, the number of genes that could be observed in a single experiment increased dramatically.

Several studies have applied this approach to understanding the host patterns of transcription during hemorrhagic fever virus infection. One study looked at the gene expression changes of liver cells and found that immune response genes become upregulated after Ebola infection, but that the expression of interferon-stimulated genes does not change after infection, presumably due to viral inhibition of the interferon response (Kash et al., 2006). This was confirmed by another study that showed that liver cells infected with a VP35-mutated Ebola virus strain resulted in a strong upregulation of ISGs (Hartman, Ling, Nichol, & Hibberd, 2008). Additional studies looked at the patterns of Ebola infection during live infection in macaques (Rubins et al., 2007) and mice (Cilloniz et al., 2011), and identified strong increases in the expression of cytokine, chemokines and apoptotic genes, which can help explain the clinical symptoms observed throughout infection. These studies illustrate the usefulness of applying a whole-transcriptome approach to understand the activation of host responses, since it makes it easy to identify which genes may be playing a critical role throughout infection and can help guide the development of new drugs aimed at modulating this response.

In Chapter 2, I use transcriptional sequencing data to characterize the immune responses of macaques infected with different hemorrhagic fever viruses, Lassa virus (Malhotra et al., 2013) and Marburg virus (Connor et al., 2015, in press). I describe the changes in gene expression that make up the common immune response in both types of infection and I discuss their function. Additionally, I identify a subset of genes whose differences in expression during the early stages of infection are strong enough to distinguish Lassa infection from Marburg infection. I also explain how these results were

validated using two methods: by recapitulating the expression levels of a subset of genes using RT-PCR, and by applying multidimensional scaling on the microarray expression values of a partially independent set of samples. These results serve as evidence that host transcriptional signatures have the potential to differentiate between specific viral infections during the early stages of disease.

In Chapter 3, I apply RNA sequencing to determine the major transcriptional changes that take place during Ebola virus infection in a macaque model of infection. I compare the strong response of interferon-stimulated genes, which is constant across all tissues, with the gene expression changes of genes that take place only in specific tissues. I also contrast the serum concentrations of a number of cytokines reported by two Ebola pathogenesis studies and correlate them with the expression levels that we observe in samples from liver, spleen and peripheral blood mononuclear cells.

In Chapter 4, I use gene expression data from a recent study that looked at the host response of macaques to different arenaviruses (Rasmussen et al., 2014), and compare it to the patterns of expression of Lassa-infected macaques described in Chapter 2. I show that the expression patterns I found to be important in distinguishing Lassa and Marburg infection are also seen in an independent dataset, which serves as an indicator that the signal we identified can be reliably detected under different experimental conditions. Additionally, I compare the patterns of gene expression from Chapter 3, to those obtained in an independent study that infected macaques with Ebola and quantified the mRNA expression using microarrays (Rubins et al., 2007).

1.10 Complementary side projects

In addition to my main research project, I have taken part in a number of side projects that benefited from applying a bioinformatics approach. Here, I will describe two of the most substantial research questions that I have worked on:

1) the study of host expression patterns during vaccinia virus infection, and

2) the metagenomic analysis of the viral biodiversity of rodents in French Guiana.

1.10.1 Expression of host factors during vaccinia virus infection.

Vaccinia virus is a DNA virus that requires the presence of host proteins to complete its life cycle. The aim of this project was to identify host genes that increased in expression during vaccinia infection and to understand their role in facilitating vaccinia replication. I was able to identify a number of differentially regulated genes and found that a majority of them were being targeted by a common transcription factor, the heat-shock master transcriptional regulator HSF1. This project is described in Chapter 5.

1.10.2 Viral biodiversity of rodent communities in French Guiana.

Metagenomics makes it possible to collect biological samples from an environment and determine the composition of the microbial communities they contain. The goal of this project was to identify the viruses that were present in different tissue samples from rodent and bat species in French Guiana. I was able to characterize a number of previously unreported viral species, including a high coverage sequence of a novel strain of arenavirus (Lymphocytic choriomeningitis virus). This project is described in Chapter 6.

CHAPTER 2

LASSA AND MARBURG VIRUSES ELICIT DISTINCT HOST TRANSCRIPTIONAL RESPONSES EARLY AFTER INFECTION

2.1 Preamble

Lassa virus and Marburg virus are two causative agents of viral hemorrhagic fever. Their diagnosis is difficult because patients infected with either pathogen present similar nonspecific symptoms early after infection. Current diagnostic tests are based on detecting viral proteins or nucleic acids in the blood, but these cannot be found during the early stages of disease, before the virus starts replicating in the blood. Using the transcriptional response of the host during infection can lead to earlier diagnoses compared to those of traditional methods. In this study, we use RNA sequencing to obtain a high-resolution view of the in vivo transcriptional dynamics of peripheral blood mononuclear cells (PBMCs) throughout both types of infection. We report a subset of host mRNAs, including heat-shock proteins like HSPA1B, immunoglobulins like IGJ, and cell adhesion molecules like SIGLEC1, whose differences in expression are strong enough to distinguish Lassa infection from Marburg infection in non-human primates. We have validated these infection-specific expression differences by using microarrays on a larger set of samples, and by quantifying the expression of individual genes using RT-PCR. These results suggest that host transcriptional signatures are correlated with specific viral infections, and that they can be used to identify highly pathogenic viruses during the early stages of disease, before standard detection methods become effective.

The viral infections and the RNA isolation were performed at the United States Army Medical Research Institute of Infectious Diseases (USAMRIID). These samples were sent to Boston University, where they were prepared for sequencing, and they were sequenced at Tufts University. I analyzed the results of the RNA sequencing samples, as well as the microarray samples that were used for validation. I also performed the RT-PCR validation of HSPA1B and SIGLEC1.

Note: The majority of the work presented in this chapter has been previously published in (Caballero et al., 2014).

2.2 Introduction

Lassa fever and Marburg fever are acute viral hemorrhagic illnesses. Lassa Fever is caused by the Lassa virus and it is endemic to a large region of West Africa, where 100,000-300,000 cases have been estimated to occur each year (J. B. McCormick et al., 1987). More than 20 cases have been imported to the United States, Canada, United Kingdom and Japan since 1969, when the disease was first recognized (Buckley & Casals, 1970; Macher & Wolfe, 2006). The case fatality rate in hospitalized patients has recently been reported to be 69% (Shaffer et al., 2014). Marburg fever is a viral hemorrhagic illness caused by the Marburg virus. Several outbreaks have been reported in the Democratic Republic of the Congo (Daniel G Bausch et al., 2006) and Angola (Towner et al., 2006) with case fatality rates approaching 90%. Travelers visiting Uganda have also imported the virus into the Netherlands (Timen, 2009). Both diseases lack effective therapeutics and early diagnostics.

Early clinical symptoms of infection are shared between different hemorrhagic fevers, and they resemble those of a flu-like illness: fever, headache, myalgia, sore throat, vomiting, diarrhea and dry cough (Drosten, Kümmerer, Schmitz, & Günther, 2003). During the 2014 West Africa Ebola outbreak, samples from suspected patients were routinely tested to rule out infection by Lassa virus, since both pathogens were present in the region and distinguishing between them was an important part of the differential diagnosis (Sylvain Baize et al., 2014; Green, 2014).

Current diagnostic tests for Lassa and Marburg fevers include enzyme-linked immunosorbent assays (ELISA) and reverse transcriptase polymerase chain reaction (RT-PCR) (D G Bausch et al., 2000; Grolla, Lucht, Dick, Strong, & Feldmann, 2005; Panning et al., 2010). Because these methods rely on the direct detection of viral proteins or nucleic acids in the blood, they are not useful during the incubation period of the virus, which happens during the early stages of infection (D G Bausch et al., 2000; Drosten et al., 2003). Additionally, most potential therapies for these pathogens—ribavirin for Lassa virus (P B Jahrling, Peters, & Stephen, 1984), post-infection vaccination for filoviruses (Thomas W Geisbert & Feldmann, 2011), and stable nucleic acid-lipid particle (SNALP) treatment for Marburg virus (Ursic-Bedoya et al., 2014)—work best when administered early. Developing tests that diagnose these infections during the early stages of disease would positively impact survival.

One approach for developing more rapid diagnostics involves incorporating the host response to infection as part of the diagnostic assay. During an infection, circulating immune cells are present in high numbers and they respond to distress signals from many

different tissues by altering the levels of expression of specific genes. These changes in expression often take place before the virus starts replicating in the blood (viremia) (Baas et al., 2006; Malhotra et al., 2013), which means that they could be used to perform earlier diagnoses than those allowed by traditional methods. To ensure that the predictive power of these changes in expression can distinguish between different causative agents, the host gene expression profiles associated with each type of infection must be determined in advance. Two recent reports have used a host response signature to discriminate symptomatic patients with viral infections from those infected with bacteria (Hu, Yu, Crosby, & Storch, 2013; a. K. Zaas et al., 2013).

I analyzed the transcriptome of PBMCs from non-human primates (cynomolgus macaques) exposed to Lassa virus and to Marburg virus (Malhotra et al., 2013), and found that the circulating immune cells develop a robust antiviral response days before clinical symptoms. In this article, I show that the expression differences in the genes that make up this response can be used to distinguish between Lassa and Marburg infections as early as 3 days after exposure.

2.3 Methods

2.3.1 LASV sequential sampling study

The samples used in this study came from a larger sequential sampling study designed to characterize the disease progression of viral hemorrhagic fevers after aerosol exposure. Seven adult male and female cynomolgus macaques (*Macaca fascicularis*), ranging in weight from 5.5 to 8.7 kg were obtained from licensed and approved vendors.

During the acclimation to the biosafety level 4 (BSL4) containment suite, a pre-challenge blood sample (8 days before exposure) was collected to provide a baseline for further analyses. (These samples are referred to in this manuscript as *pre-infection* or *0 days post-infection* samples.) On day 0, NHPs were anesthetized and minute volumes calculated by performing whole-body plethysmography (Buxco Research Systems, Wilmington, NC) just prior to the exposure. NHPs were exposed individually at a target dose of 1000 Plaque Forming Unit (PFU) of LASV Josiah in a head-only chamber in a class III biological safety cabinet maintained under negative pressure within a BSL-4 laboratory. Aerosols were created by a 3-jet Collison nebulizer (BGI Inc., Waltham, MA) and controlled by the USAMRIID Automated Bioaerosol Exposure System (ABES-II). The virus preparation used to infect NHPs was free of contamination of endotoxin and mycoplasma. After challenge, blood samples were collected at various days post-infection (dpi), based on approved collection allowances, as well as at euthanasia. Groups of two to three NHP were euthanized at 3, 6, 8 and 10 dpi, and peripheral blood mononuclear cells (PBMCs) were prepared from collected blood.

2.3.2 MARV sequential sampling study

The Marburg study was performed under the same conditions as the Lassa study, with three differences: 1) The virus strain used was MARV Angola, 2) the number of adult male and female macaques was nine, and 3) macaques were euthanized and PBMCs collected from blood at 1, 3, 5, 7 and 9 dpi.

2.3.3 RNA processing

PBMCs were isolated from blood pre-diluted with saline using ACCUSPIN System-Histopaque-1077 tubes as per manufacturer's recommendations, and subsequently lysed in TRI Reagent LS (Sigma-Aldrich, St. Louis, MO) at USAMRIID.

2.3.4 Sequencing library preparation

A subset of the total RNA samples from both datasets was processed for RNA sequencing (Figure 2.1A). Sequencing libraries were prepared for mRNA sequencing using TruSeq v2 RNA Sample Preparation kits (Illumina, San Diego, CA). Briefly, per the manufacturer protocol, mRNA was purified from approximately 500ng to 1μg of total RNA per sample using poly-T oligo-attached magnetic beads, then fragmented and primed with random hexamers for cDNA synthesis. Multiple indexing adapters were ligated to the ends of the double-stranded cDNA, and the DNA fragments with adapter molecules on both ends were then selectively PCR enriched to create the sequencing libraries. Libraries were validated for size and purity on an Agilent 2100 Bioanalyzer using the DNA-1000 chip (Agilent Technologies, Santa Clara, CA) before they were normalized and pooled for multiplex sequencing. Cluster generation and sequencing were performed by the Analytical Core Facility at Tufts University (Department of Physiology).

2.3.5 RNA sequencing analysis

After sequencing the *Macaca fascicularis* PBMC RNA samples from both types of infection (Lassa and Marburg), the generated reads were trimmed and the adapters

were removed using Trimmomatic (Bolger, Lohse, & Usadel, 2014). The processed reads were aligned to the *Macaca mulatta* genome (Ensembl release 75) using TopHat 2.0.6 (Trapnell, Pachter, & Salzberg, 2009) with default parameters (segment length of 25, allowing up to 2 segment mismatches) and using the RhesusBase2 (Zhang et al., 2013) annotated *M. mulatta* transcriptome as reference. Gene counts were obtained using HTSeq (Anders, 2010) to count reads that aligned uniquely to each gene. Counts were normalized to compensate for differences in library size using the trimmed mean of M-values normalization method (Robinson & Oshlack, 2010) included in the edgeR BioConductor package (Robinson, McCarthy, & Smyth, 2010). The normalized sequenced reads were also aligned to the Lassa virus (NCBI BioProject accession PRJNA14864) and Marburg virus (NCBI BioProject accession PRJNA15199) genomes using BWA (Li & Durbin, 2009) to quantify the number of viral reads in each sample.

A gene was deemed to show statistically significant changes in expression at a specific time after infection if the moderated t-test (Robinson et al., 2010) between the infected and pre-infection samples resulted in a multiple testing-corrected p-value lower than 0.05, with an absolute fold change in expression greater or equal to 3, and an average number of reads across all samples greater than 4 counts per million (CPM). When calculating relative changes to the pre-infection samples, infected samples were not subtracted from their individual uninfected controls (since not every infected sample had a pre-infected control), but from an average of all the pre-infection samples. Statistical significance, however, was calculated using the individual pre-infection samples, not their average.

2.3.6 Microarray sample preparation

PBMCs were processed for microarray analysis as described previously (J. Y. Yen et al., 2011). Briefly, total RNA was extracted from the TRI Reagent LS samples, then amplified using the Low-Input Quick Amp Labeling kit (Agilent Technologies, Santa Clara, CA) and hybridized to Whole Human Genome Oligo Microarrays (Agilent Technologies, Santa Clara, CA) in a 2-color comparative format along with a reference pool of messenger RNA. Images were scanned using the Agilent High-Resolution Microarray Scanner and raw microarray images were processed using Agilent's Feature Extraction software.

2.3.7 Microarray Expression Analysis

Agilent two-color human gene expression microarrays were processed using limma (Smyth, 2004). Expression values were obtained by calculating the log-ratio between the intensities of the red channel (which corresponds to experimental samples) and the green channel (which corresponds to Human Universal Reference RNA (Novoradovskaya et al., 2004)). These were later background-corrected and normalized using LOESS (Smyth & Speed, 2003).

2.3.8 Multidimensional Scaling

The limma (Smyth, 2004) implementation of Multidimensional Scaling was used to reduce the dimensionality of a matrix composed of the expression values of the following 12 genes across the 66 RNA samples: MX1, OAS1, IFI44L, CXCL10, HSPA1B, IGJ, HSPA1L, BIRC3, SIGLEC1, TNK2, TNFSF10, NR4A2.

2.3.9 Reverse Transcriptase Polymerase Chain Reaction

RT-PCR reactions were carried out in 50 µL reactions using Power SYBR® Green RNA-to-CT™ 1-Step Kit from Life Technologies: 25 µL SYBR® Green RT-PCR Mix, 2 µL of each primer at 100 µM, and .4 µL RT Enzyme Mix. Thermal cycling was conducted as follows: 30 minutes at 48 ºC, 10 minutes at 95 ºC, 15 seconds at 95 ºC (40 cycles), and 1 minute at 60 ºC (40 cycles).

2.4 Results

2.4.1 Clinical Symptoms

Cynomolgus macaques were exposed to an aerosol dose of 1000 plaque forming units of Lassa virus or Marburg virus, as previously described (Malhotra et al., 2013; Connor et al., 2015, in press) (see Figure 2.1A). In these models, animals began to develop a fever around day 3 after exposure to Lassa virus and day 4 after exposure to Marburg virus (Figure 2.1B, Table 2.1). The onset of viremia took place around day 8 for Lassa and day 4 for Marburg (confirmed by both plaque assay and RT-PCR). The liver enzymes aspartate transaminase (AST) and alanine transaminase (ALT) showed increased levels of concentration in both diseases around day 6. D-dimers started to increase around day 5 for Lassa and day 6 for Marburg. In both types of infection, all of these markers peaked during the late stages of infection (days 9-10). These data confirmed that clinical markers of infection that could be used for diagnosis (e.g. viremia) did not appear in these models until day 4 post-infection.

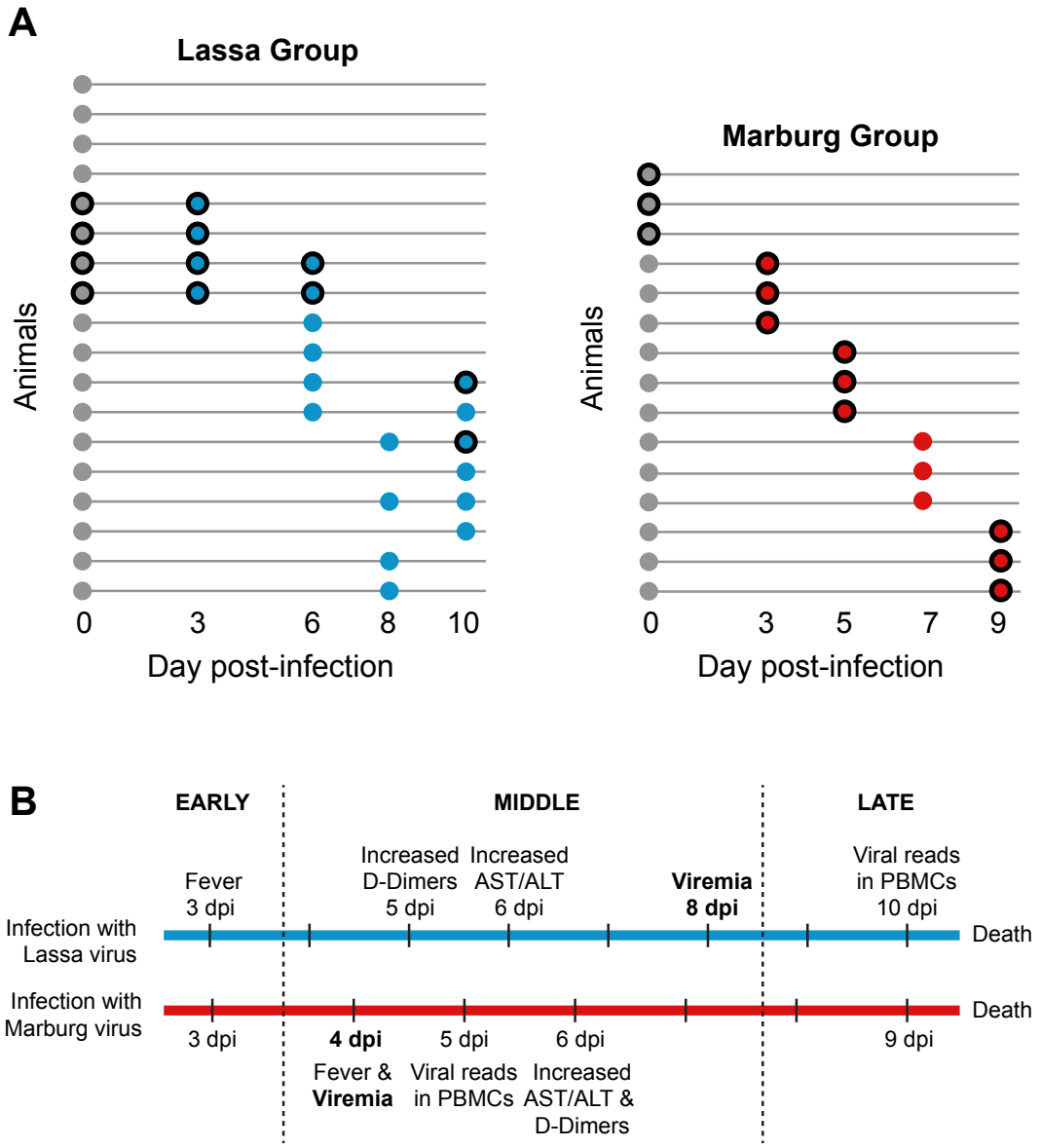


Figure 2.1 Quantified samples in each virus group.
(A) Each circle represents an RNA sample isolated from the peripheral blood mononuclear cells (PBMCs) of a cynomolgus macaque. Grey circles represent pre-infection samples from different macaques. Blue circles represent samples infected with Lassa virus. Red circles represent samples infected with Marburg virus. All samples were quantified using microarrays. Samples with a black border were also quantified using RNA sequencing. (B) Diagram of the disease timecourse highlighting the most relevant clinical events, from the appearance of fever 3 days post-infection to the late stage of disease and death.

	Lassa	Marburg
Fever	Day 3	Day 4
Viremia	Day 8	Day 4
Increased AST/ALT	Day 6	Day 6
Increased D-dimers	Day 5	Day 6

Table 2.1: Times at which clinical markers are detected in each type of infection. Viremia was measured using RT-PCR in the Lassa group, and RT-PCR and plaque assay in the Marburg group.

2.4.2 RNA sequencing

To determine the transcriptional changes that take place in the circulating immune system in response to a hemorrhagic fever virus infection, we quantified the abundance of mRNAs present in peripheral blood mononuclear cells (PBMCs) of cynomolgus macaques infected with either Lassa or Marburg virus. The samples were obtained before the animals were infected, and at multiple times after infection (see Figure 2.1A). We selected for polyadenylated transcripts and sequenced 12 RNA samples from each group using Illumina technology. This process generated an average of 46.27 ± 5.92 million reads per sample; 72% of which (33.46 ± 4.33 million) aligned to the *M. mulatta* (rhesus macaque) genome—the closest organism to *M. fascicularis* with an annotated genome. Multiple factors may explain why 28% of reads remained unaligned: an incomplete assembly of the *M. mulatta* genome and transcriptome, genetic differences between *M. mulatta* and *M. fascicularis*, or the appearance of amplification artifacts during the sequencing process. We assessed the expression level of individual mRNAs by counting the number of reads that aligned uniquely to each gene, and then normalizing this number to allow comparisons across all samples.

To determine the amount of sequencing reads that came from each virus, we aligned the reads to the Lassa virus and Marburg virus genomes. In the Lassa infection group, only the samples taken 10 days post-infection (dpi) contained viral reads (less than 20 reads). In the Marburg infection group, around 2000 viral reads were seen 5 dpi, and around 80,000 reads 9 dpi. These numbers suggest that peripheral blood mononuclear cells are early targets for Marburg virus but not for Lassa virus. The number of viral reads found in PBMCs correlates well with the level of viremia, as reported in (Malhotra et al., 2013; Connor et al., 2015, in press).

2.4.3 The common transcriptional response of PBMCs during Lassa and Marburg infection is dominated by type I interferon-stimulated genes

To identify what transcriptional changes are shared by different hemorrhagic fever virus infections, we identified genes whose levels of expression 1) closely matched in both types of viral challenge, 2) strongly increased or decreased 3 days post-infection (dpi), and 3) increased or sustained this level of expression throughout the remaining time points (5-10 dpi). We labeled this group of genes the *common response* (Figure 2.2B). Some of its representative members include transcription factors IRF7 and STAT1, which serve as master regulators of host immunity; pattern recognition receptors DDX58 (RIG-I), IFIH1 (MDA-5) and DHX58 (LGP2), which activate different signaling cascades in the innate immune system; type I interferon-stimulated genes ISG15, ISG20, OAS1, OAS2, OASL, MX1, IFIT1, IFIT2, IFIT3, HERC5, HERC6, IFI6, IFI35, IFI44 and IFI44L, which play a variety of antiviral roles (Schoggins et al., 2011); and the cytokine CXCL10, which attracts activated T cells (Dufour et al., 2002).

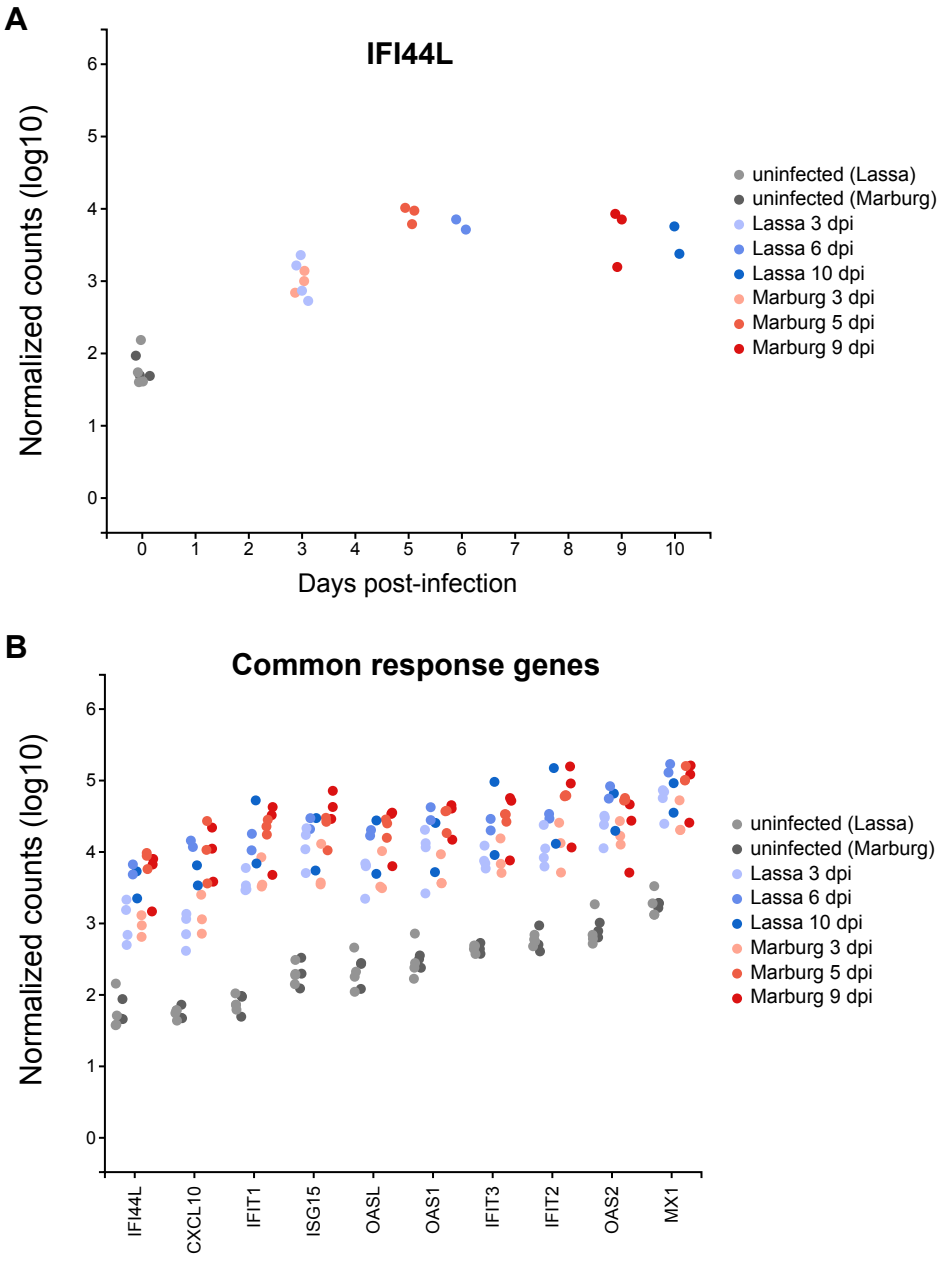


Figure 2.2 Genes showing similar patterns of expression during Lassa virus and Marburg virus infection.
Each point represents the level of expression (measured in log10-normalized read counts) of a gene that behaves similarly in both types of infection (increasingly darker shades of blue for Lassa, and increasingly darker shades of red for Marburg). Panel (A) represents the expression of gene IFI44L in the y axis and the time post-infection in the x-axis. Panel (B) condenses the information in (A) and shows additional genes along the x-axis.

2.4.4 A subset of genes shows unique patterns of expression in each type of infection

After identifying the subset of genes that made up the common transcriptional response, we looked for genes that showed unique transcriptional patterns. To do this, we considered genes that 1) showed a statistically significant increase or decrease in their 3 dpi levels of expression for one type of viral infection but not the other, and 2) increased or sustained this difference in expression throughout the later stages of the infection. In the Lassa infection group, we detected several genes with distinctly upregulated expression (Figure 2.3), including SIGLEC1, a cell adhesion molecule expressed by dendritic cells and macrophages (Crocker, Paulson, & Varki, 2007; Izquierdo-Useros et al., 2012); TNFSF10 (TRAIL), a death receptor ligand implicated in apoptosis and immune response signaling (Johnstone, Frew, & Smyth, 2008); and TNK2, a tyrosine kinase implicated in cell spreading and migration (Howlin, Rosenkvist, & Andersson, 2008). Other genes, like CENPF, ENPP4, and GZMA, a protease responsible for the release of cytotoxic T-cell granules (Hink-Schauer, Estébanez-Perpiñá, Kurschus, Bode, & Jenne, 2003), were significantly downregulated during Lassa infection in comparison to Marburg infection.

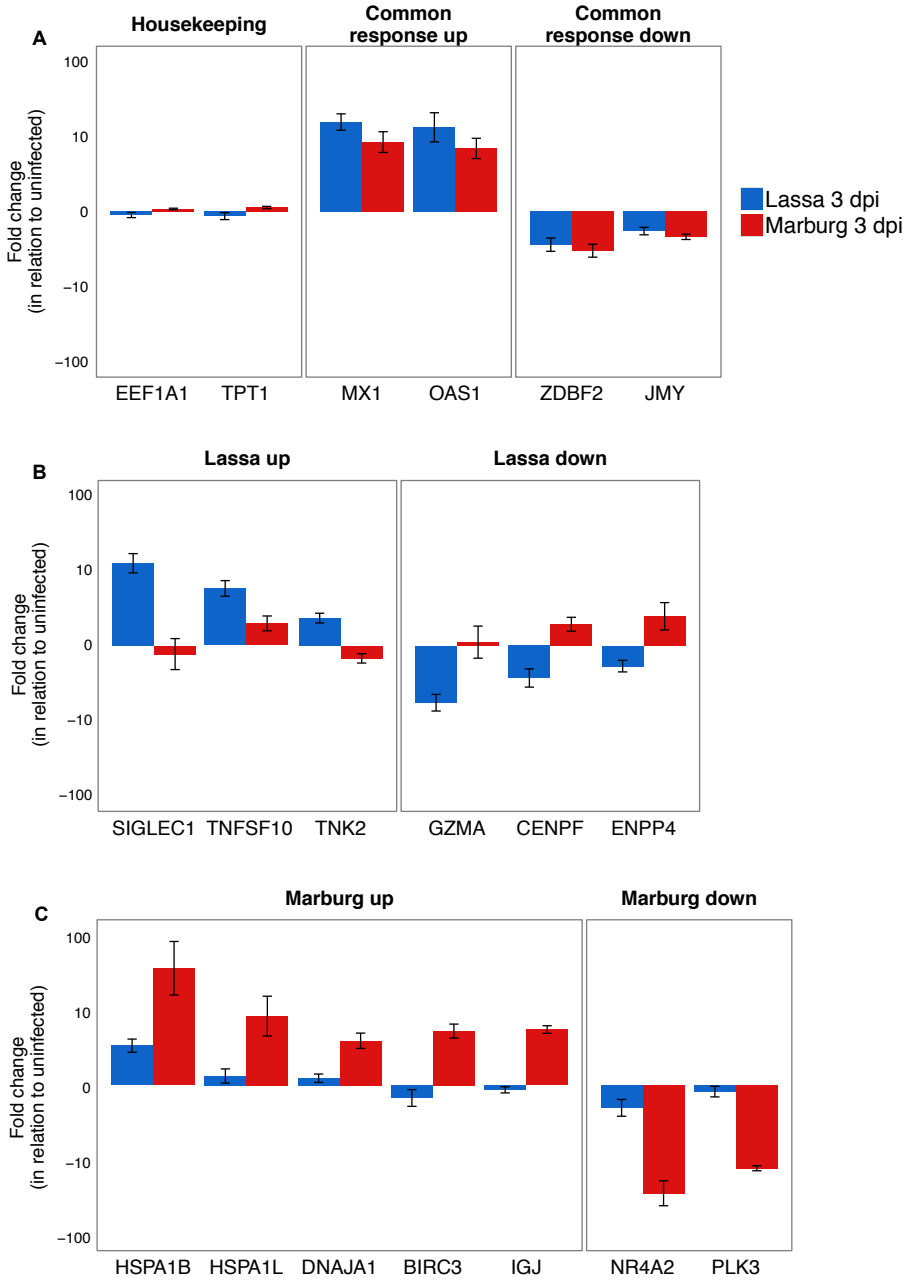


Figure 2.3 Gene expression patterns during Lassa or Marburg virus infection. A few representative genes are shown in each panel representing each category: (A) housekeeping genes, genes that uniformly increase or decrease their expression 3 days after both infections, (B) genes that show unique patterns of expression during Lassa infection, and (C) during Marburg infection. The y-axis represents the average fold change in expression when comparing infected and uninfected samples.

In the Marburg infection group, we identified a distinct upregulation in the expression of genes encoding heat shock proteins (HSPA1B, HSPA1L and DNAJA1), the antiapoptotic gene BIRC3, and the immunoglobulin genes IGJ, IGLV10-54, among others. We also found genes that showed lower levels of expression during Marburg infection when compared to Lassa infection. Examples of these include PLK3, a kinase associated with stress response and apoptosis (Bahassi et al., 2002; Xie et al., 2001); and NR4A2, a nuclear receptor which has been shown to induce apoptosis when downregulated (Ke et al., 2004).

2.4.5 Validation of infection-specific patterns of expression using microarrays and RT-PCR

We wanted to know if the genes that showed significant differences in each type of infection could be used to classify unknown samples. Although we were not able to obtain an independent test dataset, we used additional samples from two other studies to characterize the host transcriptional response to Lassa and Marburg infection (Malhotra et al., 2013; Connor et al. 2015, in press). This dataset included the 24 samples that we measured using RNA sequencing, as well as 26 additional uninfected samples, 12 Lassa-infected samples and 3 Marburg-infected samples (Figure 2.1A), for a total of 65 samples, all of which were quantified using Agilent two-color human microarrays. We looked at the genes that we had previously identified (using RNA sequencing) as being differentially expressed during one of the two infections, and found that the measurements from both platforms were highly correlated (0.81 Pearson's correlation).

To understand in more detail the differences in expression reported by RNA sequencing and microarrays, we classified a subset of genes into four categories: 1) **Housekeeping genes** were those that showed high levels of expression across all samples. 2) **Early common response genes** were those where infected samples showed higher levels of expression than uninfected samples, and had similar magnitudes in both types of infection 3 days post-infection. 3) **Early Marburg-specific response genes** were those where Marburg-infected samples showed significantly higher (or lower) patterns of expression, compared to uninfected or Lassa-infected samples 3 days post-infection. 4) **Early Lassa-specific response genes** were those where Lassa-infected samples showed significantly higher patterns of expression, compared to uninfected or Marburg-infected samples 3 days post-infection.

To determine if we could use the common and specific genes to distinguish between different types of infected samples, we use Primary Component Analysis (PCA) reduce the dimensionality of the microarray gene expression values to two dimensions. Instead of using every gene in the arrays, we selected a subset of 12 genes that included early common response genes, early Marburg-specific response genes and early Lassa-specific response genes (Figure 2.4). Reducing the dimensionality of all samples using these genes resulted in three clusters: uninfected samples, infected with Lassa virus, and infected with Marburg virus. Each of the infected clusters included not only the early-infected samples, but also those samples taken at later stages of infection. This indicates that the expression patterns of these genes are useful indicators throughout both early and late stages of infection.

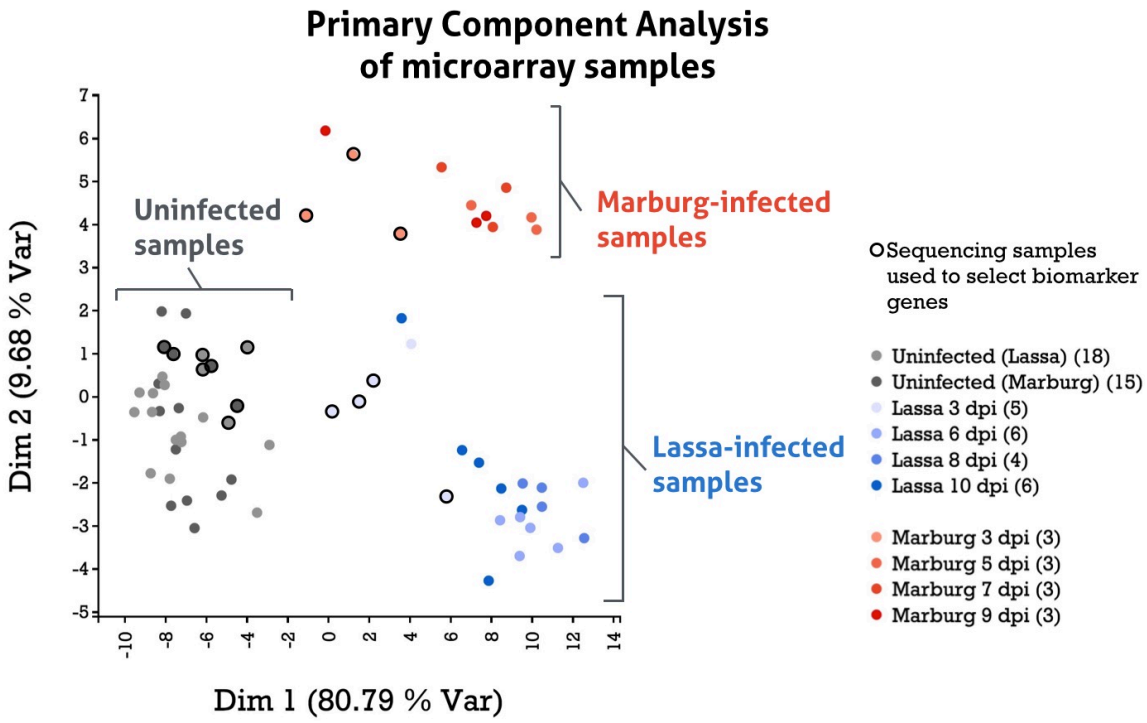


Figure 2.4 Primary Component Analysis of Microarray Samples.
Each point corresponds to an RNA sample. Lassa-infected and Marburg-infected samples are shown in increasingly darker shades of blue and red, respectively. Distances between points represent the similarity among samples based on the application of Primary Component Analysis (PCA) to the microarray expression values of 12 biomarker genes: 4 common response genes (MX1, OAS1, IFI44L and CXCL10), 4 Marburg-specific response genes (HSPA1B, IGJ, HSPA1L and BIRC3) and 4 Lassa-specific response genes (SIGLEC1, TNK2, TNFSF10 and NR4A2). Uninfected from each batch are shown in two shades of grey. Microarray samples that correspond to the sequencing samples that were used to identify the biomarkers genes are shown with a black border.

We then looked at the reported expression levels for these genes in both platforms (Figure 2.5). We only used samples that were quantified in both sequencing and arrays to ensure that differences in variability would only be related to the platform, not to the number of samples. The majority of genes showed similar levels of expression in both platforms. The biggest differences came from genes like HSPA1L and IGJ, which showed much smaller fold changes in arrays than in sequencing. On average, both highly

upregulated and downregulated genes showed fold changes with a 1.5-2 times smaller magnitude in arrays when compared to sequencing.

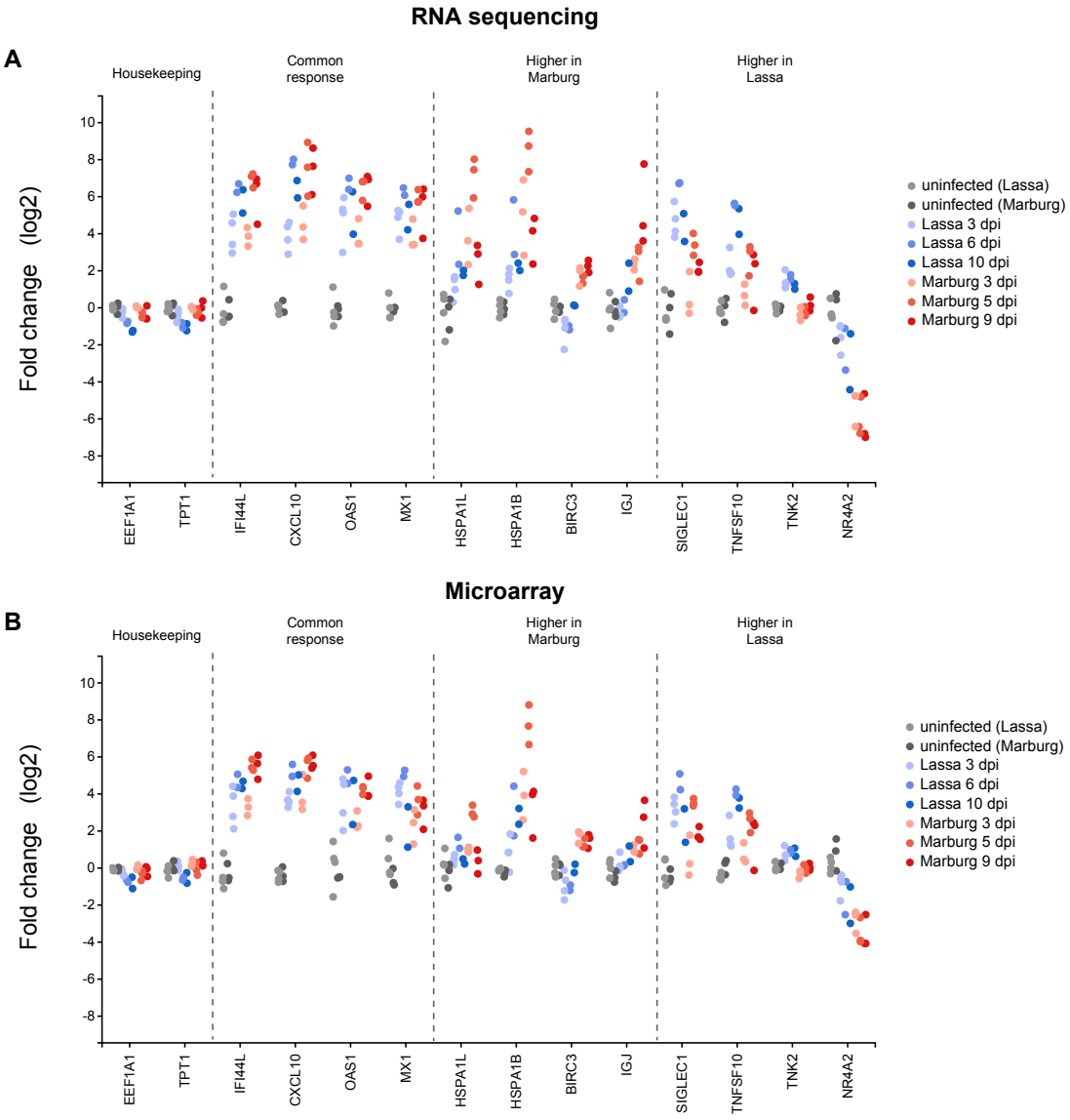


Figure 2.5 Biomarker genes quantified using RNA sequencing and microarrays. For each gene, the y-axis represents its amount of expression in each RNA sample, and the x-axis represents these samples ordered by time of infection, colored in increasingly darker shades of blue for Lassa, and red for Marburg. The samples in (A) were quantified using RNA sequencing and the fold change represents the log2 difference between the average infected and uninfected normalized read counts. The samples in (B) were measured using two-color microarrays and the fold change represents the log2 ratio between the intensity of the red channel and the green channel (see Methods).

This comparison confirmed that both RNA sequencing and microarrays show unique changes in mRNA levels early after infection. This led us to test the hypothesis that a more direct RT-PCR analysis of a small number of these sentinel genes would show similar patterns of unique regulation, with the advantage that the format of this assay could be applied more quickly and cheaply in the field. We chose to analyze the expression of two genes, SIGLEC1 and HSPA1B, which showed unique upregulation at early stages of Lassa or Marburg virus infection, respectively. RT-PCR assays carried out on uninfected samples, and on samples collected 3 days post-infection, showed that there was no statistically significant change in SIGLEC1 expression between these two time points in Marburg-infected animals, but that there was a 70-fold change in expression in Lassa-infected animals. Similarly, HSPA1B showed a minor change in Lassa-infected animals, but a 36-fold change in Marburg-infected animals. This confirmed that SIGLEC1 becomes upregulated during the early stages of Lassa infection but not Marburg infection, while HSPA1B is only upregulated during the early stages of Marburg infection (Figure 2.6).

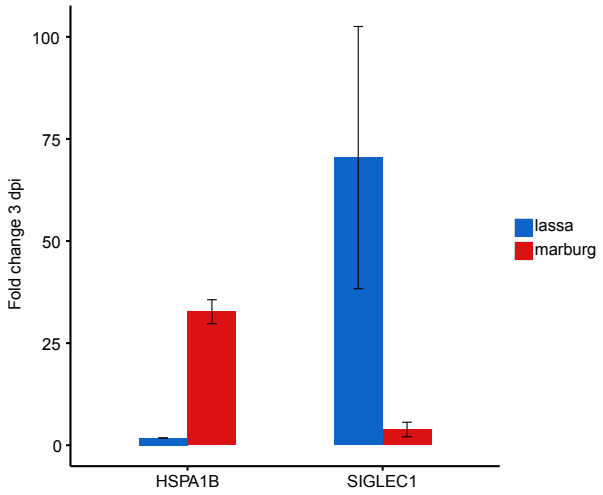


Figure 2.6 RT-PCR validation of 3 day post-infection expression values. SIGLEC1 and HSPA1B are the genes with the highest fold change differences between each virus group 3 days post-infection. The y-axis represents the difference in expression between each gene and the constitutively expressed gene BIN2. The RT-PCR reactions for each gene were performed in duplicate (n=2).

2.5 Discussion

In the work described in this manuscript, we investigated the hypothesis that the circulating immune response can be used to aid early diagnosis of viral infections that are associated with high levels of mortality. When we analyzed the transcriptional patterns of the circulating immune system in response to two different hemorrhagic fever viruses, we identified a subset of genes that uniquely increased their rate of transcription 3 days after infection and that sustained these differences throughout infection (Figure 2.3).

An important aspect of this study was that the uniquely regulated genes were not the only genes that changed their patterns of expression during the early times of infection (before the detection of clinical symptoms), but that many other genes—pattern recognition receptors, interferon-stimulated genes, cytokines, and immune and antiviral

transcriptional regulators—also became rapidly upregulated. This suggests that early after infection, there is a strong activation of the type I interferon innate immune response. This early innate response has also been reported for infections with other viruses, such as Ebola (Gupta et al., 2001), influenza (Huang et al., 2011) and respiratory syncytial virus (RSV) (Chen et al., 2011), and suggests that the activation of the innate immune system is being driven by a different mechanism. Surprisingly, we did not detect a significant upregulation of type I interferon genes, although we did see it for their downstream effectors. Protein studies indicate that type I interferon is present during Marburg virus infection, suggesting that interferon transcription might be localized to plasmacytoid dendritic cells, which are not strongly represented in the PBMC population (Fritz, Geisbert, Geisbert, Hensley, & Reed, 2008).

In addition to the common transcriptional changes that we observed at early times of infection, we also identified a number of genes with unique patterns of expression in each type of infection. The unique markers of Marburg infection were generally classifiable into functional categories, including heat-shock proteins (HSPA1B, HSPA1L, DNAJA1) and immunoglobulin-associated genes (IGJ, IGLV10-54). These results suggest that the response to Marburg infection leads to the expansion of both stress adaptive and immunological maturation pathways at very early times of infection.

Some of the most uniquely expressed genes during Lassa infection are SIGLEC1 (also known as CD169 or sialoadhesin) and TNFSF10 (also known as TRAIL). The upregulation of these genes suggests an increase in SIGLEC1+ TRAIL+ macrophages early in Lassa infection (Crocker et al., 2007; Russier, Reynard, Tordo, & Baize, 2012;

van der Kuyl et al., 2007). A more detailed understanding of the mechanisms leading to the unique signatures identified will require further investigation. While both Lassa virus and Marburg virus are known to evade antiviral responses by inhibiting the expression of specific host genes (Christopher F Basler & Amarasinghe, 2009; Hayes & Salvato, 2012), it is unlikely that these mechanisms are directly responsible for the expression differences that we found in PBMCs, since they take place during the pre-viremic stages of infection. The specific differences observed take place even in the absence of any viral genetic material in the PBMCs, which suggests that these circulating cells are responding to an external signal. An alternative explanation for the existence of unique signatures might be related to each pathogen's cell tropism: Marburg virus preferentially targets monocytes, macrophages, dendritic cells and Kupffer cells (Hensley, Alves, et al., 2011), and leads to significant lymphocyte apoptosis (T W Geisbert et al., 2000). Lassa virus targets dendritic cells, Kupffer cells, hepatocytes, adrenal cortical cells and endothelial cells (Hensley, Smith, et al., 2011), and is not known to induce lymphocyte decreases. It is possible that a combination of which cells become initially infected and the magnitude of the early response combine to ultimately produce the infection-specific signatures that we observe.

We have shown that microarrays can be used to validate the changes in expression that take place in Lassa- and Marburg-infected PBMCs in addition to RNA sequencing. Although the majority of genes that undergo significant early changes in expression were highly correlated in both platforms, microarrays showed a smaller range of gene expression differences between both types of infected samples and greater

variability. Three previous studies that analyzed the immune response to Lassa virus infection (Malhotra et al., 2013; Juan Carlos Zapata et al., 2013) and LCMV virus infection (Djavani et al., 2007) were in general agreement with the major trends that we identified using RNA sequencing: mainly the upregulation of interferon-stimulated genes, pattern recognition receptors and pro-inflammatory cytokines.

The infection-specific transcriptional changes that we have found suggest a method for discriminating between diseases with similar early clinical symptoms, as is the case for Lassa and Marburg infection. In infected animals, these transcriptional changes take place prior to the appearance of viremia and other overt symptoms (Malhotra et al., 2013), and they remain sustained throughout the late stages of infection, thereby providing a way to diagnose which virus is most likely causing the infection. This diagnosis can be carried out during the early time window when treatment is most effective, and before the pathogen starts replicating in the bloodstream on a massive scale. To ensure that an RT-PCR-based diagnostic would be able to distinguish between diseases other than Lassa and Marburg disease, it would be important to determine the host transcriptional response to additional pathogens like influenza virus, dengue virus or the malaria parasite, which cause similar initial symptoms to those of viral hemorrhagic fevers.

In conclusion, the expression of a unique set of host genes increases or decreases preferentially in Lassa- or Marburg-infected immune cells. During Marburg virus infection, heat-shock protein and immunoglobulin genes show significantly greater levels of expression than during Lassa virus infection, which is characterized by the

upregulation of the cell adhesion molecule SIGLEC1, the death receptor ligand TNFSF10, and the tyrosine kinase TNK2. These changes in expression appear early after infection and remain sustained throughout the entire course of disease, which provides a way to diagnose which virus is most likely to be causing the infection without having to wait until the appearance of viremia or other overt symptoms. The fact that these changes were identified using multiple platforms (RNA sequencing, microarrays and RT-PCR) suggests a path for translating these gene expression profiles into a diagnostic test with direct clinical applications.

CHAPTER 3

IN VIVO EBOLA VIRUS INFECTION LEADS TO A STRONG INNATE RESPONSE IN CIRCULATING IMMUNE CELLS

3.1 Preamble

Ebola virus is the causative agent of a severe syndrome in humans with a fatality rate that can approach 90%. During infection, the host immune response is thought to become dysregulated, but the mechanisms through which this happens are not entirely understood. It would be helpful to have a more detailed view about the transcriptional changes that take place during infection, especially during the early stage, before the virus can be detected in the blood. In this study, we use RNA sequencing to identify the gene expression patterns of circulating immune cells from macaques infected with Ebola virus. Approximately half of the 100 genes with the strongest early increases in expression were interferon-stimulated genes, such as ISG15, OAS1, IFIT2, HERC5, MX1 and DHX58. Other highly upregulated genes included cytokines CXCL11, CCL7, CCL8, IL2RA, IL2R1, IL15RA, FAS, and CSF2RB, which have not been previously reported to change during Ebola infection. Comparing this response in two different models of exposure (intramuscular vs. aerosol) revealed a similar signature of infection with a time delay in the aerosol model. The strong innate response in the aerosol model was seen not only in circulating cells, but also in primary and secondary target tissues. Conversely, the innate immune response of vaccinated macaques was almost non-

existent. This suggests that the innate response is a major aspect of the cellular response to infection.

The viral infections, RNA isolation and sequencing for the group infected via intramuscular injection were performed at the University of Washington. The viral infections and the RNA isolation for the group infected via aerosol exposure were performed at the United States Army Medical Research Institute of Infectious Diseases (USAMRIID). The aerosol samples were sent to Boston University, where they were prepared for sequencing, and they were sequenced at the Broad Institute. I analyzed the results of the RNA sequencing samples for both groups.

3.2 Introduction

Infection with Ebola causes Ebola hemorrhagic fever, a disease associated with mortality rates between 25-90% (Feldmann & Geisbert, 2011). The earliest clinical symptoms are non-specific and flu-like, such as high fever, headache and myalgia (Feldmann & Geisbert, 2011; WHO Ebola Response Team, 2014; Wong, Kobinger, & Qiu, 2014). During the large outbreak in Western Africa (2014-present), the symptoms also included diarrhea and vomiting (WHO Ebola Response Team, 2014). The late stage of disease is associated with immune cell imbalances such as neutrophilia and coagulation disorders like diffuse intravascular coagulopathy (Bradfute et al., 2010; Paessler & Walker, 2013; Reed, Hensley, Geisbert, Jahrling, & Geisbert, 2004).

Several studies have reported that during infections caused by respiratory and hemorrhagic fever viruses, circulating immune cells undergo major gene regulatory

changes that result in the upregulation of many innate immune system genes (Caballero, Bonilla, Yen, & Connor, 2012; Mejias et al., 2013; a. K. Zaas et al., 2013; A. K. Zaas et al., 2009). One of the most interesting aspects about this surge in transcriptional activity is that it constitutes an early measurable indicator of infection. In the present study, we use multiple RNA sequencing datasets to understand the gene expression patterns in peripheral blood mononuclear cells (PBMCs) of cynomolgus macaques infected with Ebola virus. One dataset was generated using macaques infected with Ebola virus that had previously been either vaccinated against Ebola or against a different virus. The advantage of this dataset is that it provides a unique comparison between Ebola-naïve and Ebola-vaccinated animals, and it makes it possible to determine which genes become differentially regulated as a direct result of Ebola infection. The other dataset was useful to determine the differences in gene expression that are dependent on specific routes of infection. Some of the animals were infected with Ebola virus via aerosol, while others received an intramuscular injection. A common feature of all of these datasets was that in addition to the infected samples, they contained pre-infection controls, which made it possible to compare across different studies.

3.3 Materials and Methods

3.3.1 Macaque model of intramuscular exposure to Ebola virus

As described in (Marzi et al., 2013), this study used cynomolgus macaques divided into two groups: vaccinated and Ebola-naïve. The vaccinated group received an intramuscular injection of rVSV/EBOV-GP and, several days later, a lethal dose of

EBOV strain Kikwit. The Ebola-naïve group was treated with a non-protective dose of rVSV/MARV-GP and, several days later, received the same lethal dose of EBOV administered to the vaccinated group. Blood samples were taken at 0, 4 and 7 days post-infection (n=3 per timepoint/group). PBMCs were isolated from the 18 blood samples and RNA sequencing was performed as specified in (Barrenas et al., 2015). The sequencing data is available in GEO (Barrett et al., 2013) under accession GSE64538.

3.3.2 Macaque model of aerosol exposure to Ebola virus

As described in (Twenhafel et al., 2013), this study used rhesus macaques exposed to EBOV via aerosol. 12 PBMC samples were collected at 0, 3, 6 and 8 days post-infection (approximately n=2 per timepoint). 20 spleen samples were obtained at 0, 1, 3, 4, 5, 6, 7 and 8 dpi (approx. n=2 per timepoint). 12 adrenal gland samples were obtained at 1, 3, 4, 6, 7 and 8 dpi (approx. n=2 per timepoint). 8 axillary lymph node samples were obtained at 0, 1, 3, 4, 5 and 6 dpi (approx. n=1 per timepoint). 11 brain samples were obtained at 1, 3, 4, 6 and 7 dpi (approx. n=2 per timepoint). 4 liver samples were obtained at 0, 3, 5 and 8 dpi (approx. n=1 per timepoint). 3 pancreas samples were obtained at 0, 4 and 8 dpi (approx. n=1 per timepoint). RNA was extracted from all tissues (including PBMCs) and sequenced. The PBMC samples referenced above, as well as 24 additional samples taken 1, 4, 5 and 7 dpi (approx. n=4 per timepoint) were also processed using microarrays as specified in (J. Y. Yen et al., 2011). The microarray expression data is available in GEO under accession GSE68809.

3.3.3 Mouse model of exposure to Ebola virus

As described in (Cilloniz et al., 2011), this study used BALB/c mice divided into two groups. One group was inoculated intraperitoneally with wild-type EBOV. The other group was inoculated using the same method, but with mouse-adapted EBOV instead. Spleen samples were collected and homogenized at 0, 12, 24, 48 and 72 hours post-infection from three individuals (n=3) per timepoint. RNA was extracted from the 15 spleen samples and Agilent microarray processing was performed as specified in (Cilloniz et al., 2011). The sequencing data is available in the University of Washington Viromics server http://bit.ly/Cilloniz_2011_microarray.

3.3.4 Macaque model of aerosol exposure to Lassa and Marburg virus

As described in (Caballero et al., 2014), this study used cynomolgus macaques divided into two groups. One group was exposed to LASV Josiah and the other to MARV Angola, both via aerosol. RNA was extracted from 24 PBMC samples (n=12 per group) and RNA sequencing was performed as specified in (Caballero et al., 2014). The sequencing data is available in the Sequence Read Archive (Kodama, Shumway, & Leinonen, 2012) under accession numbers PRJNA222891 (Lassa) and PRJNA222892 (Marburg).

3.3.5 Fold change analysis of sequencing data

After the raw sequencing reads of each study were generated, they were trimmed and the adapters were removed using Trimmomatic (Bolger et al., 2014). The processed

reads were aligned to the Macaca mulatta genome (Ensembl release 77) (Cunningham et al., 2014) using TopHat 2.0.11 (Trapnell et al., 2009) with default parameters (segment length of 25, allowing up to 2 segment mismatches). Gene counts were obtained using HTSeq (Anders, 2010) to count reads that aligned uniquely to each gene. Counts were normalized to compensate for differences in library size using the trimmed mean of M-values normalization method included in the edgeR BioConductor package (Robinson et al., 2010).

A gene was deemed to show statistically significant changes in expression at a specific time after infection if the moderated t-test between the infected and pre-infection samples resulted in a multiple testing-corrected p-value lower than 0.05, with an absolute fold change in expression greater or equal to 3, and an average number of reads across all samples greater than 4 counts per million (CPM). When calculating relative changes to the pre-infection samples, infected samples were not subtracted from their individual uninfected controls (since not every infected sample had a pre-infected control), but from an average of all the pre-infection samples. Statistical significance, however, was calculated using the individual pre-infection samples, not their average. The Benjamini-Hochberg algorithm was used to correct the false discovery rate of multiple testing.

3.3.6 Fold change analysis of microarray data

Agilent two-color human gene expression microarrays were processed using limma (Smyth, 2004). Fold changes were obtained by calculating the log-ratio between the intensities of the red channel (corresponding to experimental samples) and the green channel (corresponding to Human Universal Reference RNA (Novoradovskaya et al.,

2004)). These were later background-corrected and normalized using the LOESS algorithm. Agilent one-color mouse gene expression microarrays were processed in a similar manner as the two-color arrays, but using the reported expression value for each probe instead of the log-ratio between the two channels. Benjamini-Hochberg was used to correct the false discovery rate of multiple testing.

3.4 Results

To determine the circulating immune system response to an Ebola virus infection, we used sequencing data from PBMCs of cynomolgus macaques infected with Ebola via intramuscular injection. Three samples were used from different animals before they were infected, 4 days and 7 days after infection, as described in (Barrenas et al., 2015).

3.4.1 A strong and sustained early host transcriptional response after Ebola infection contains many genes associated with the innate immune response.

We observed 128 genes with an 8-fold or higher upregulation by 4 days post-infection in comparison to the pre-infected samples (with an adjusted p-value of less than 0.05 and an average of at least 4 read counts per million). By 7 days post-infection, only 32 of these genes had a downregulated expression. The levels of expression of the other 96 genes remained constant or increased. Of these, 49 genes are known to be responsive to innate immunity transcription factors such as IRF3, IRF7 and STAT1 (also known as interferon-stimulated genes, or ISGs) (Schoggins et al., 2011; Uchiumi, Miyazaki, &

Tanuma, 2011) (Figure 3.1A) and are herein referred to as canonical ISGs. Figure 3.1A shows the log2-fold change of these interferon-stimulated genes at 0, 4 and 7 days post-infection, highlighting a few representative examples (MX1, ISG15, OAS1, DHX58 (RIG-I), IFIT2, and HERC5). The other 47 strong and sustained genes show similar responses to the ISG group but they have not been previously identified as downstream effectors of the interferon response (Figure 3.1B). This population includes several neutrophil-associated genes (OLFM4, CD177, SERPINB1, S100P, PTX3 and MMP8), suggesting that there is an infiltration of neutrophils in PBCMs during the course of infection.

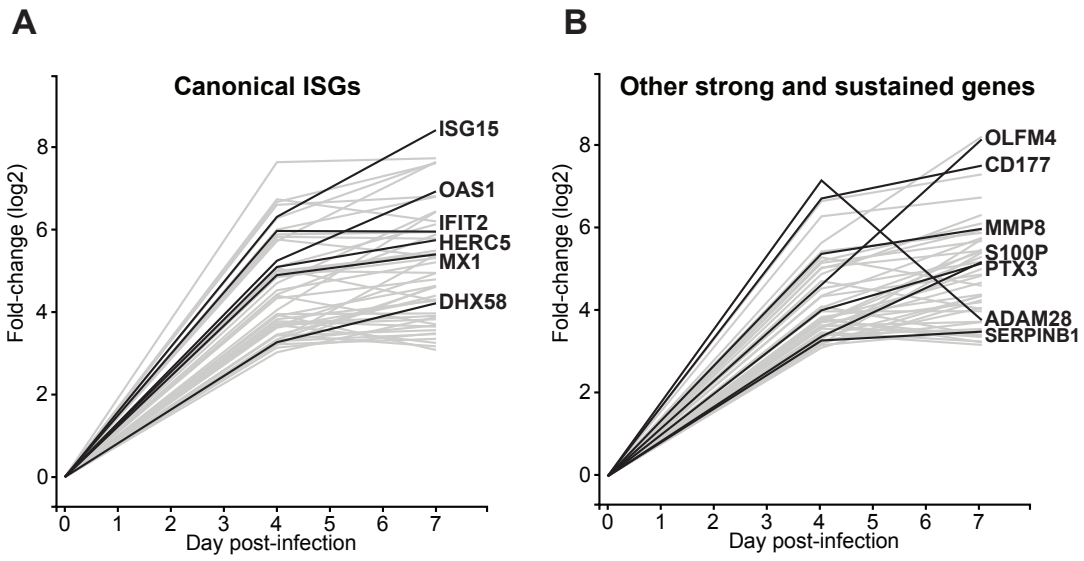


Figure 3.1 Host genes showing strong and sustained changes in expression early after Ebola virus infection.
Each line represents the average fold-change in expression (log2, y-axis) of each gene at different times (x-axis) after Ebola infection. Panel A) shows canonical interferon-stimulated genes while panel B) shows other strong and sustained genes. A representative sample of genes is shown in black to illustrate the individual patterns of expression.

To determine if this strong innate immune response was observed in other animal models of Ebola infection, we also analyzed the ISG response in mice that had been

infected with wild type EBOV (which is not pathogenic in mice) and mouse-adapted EBOV (which is pathogenic) (Cilloniz et al., 2011). We found that a similar strong ISG response was apparent 3 days after infection in both strains, irrespective of pathogenesis.

We validated these results using expression data from two previous studies that infected rhesus macaques with Ebola virus (intramuscular injection) and treated them with an anticoagulant drug (recombinant human activated protein C, rhAPC, and recombinant nematode anticoagulant protein c2, rNAPc2 (Hensley et al., 2007; J. Y. Yen et al., 2011)). We found significant similarity for a majority of early response genes even though the expression data was quantified using microarrays. For example, in the rhAPC samples, the expression of IFIT2 increased 4 log2-fold and 7.4 log2-fold at 3 and 6 days post-infection, respectively; in the rNAPc2 samples the increases were 4.6 log2-fold and 6.9 log2-fold at 3 and 6 days post-infection, respectively.

3.4.2 The sustained transcription of early responsive genes is not present in a non-productive infection

To determine if the appearance of an early innate immune response following a hemorrhagic fever virus infection was a response to the local damage caused by the initial intramuscular injection of a negative strand RNA virus, we compared the innate immune response to Ebola virus in animals that had been previously vaccinated against Ebola to those that had been immunized against a different hemorrhagic fever virus and were therefore susceptible to Ebola virus (Ebola-naïve group). Figure 3.2 shows the changes in expression of four canonical ISGs. In the Ebola-naïve group, for example, SIGLEC1 undergoes a 5.72 log2-fold change 4 days post-infection, while in the vaccinated group

the change is only 2.53 log2-fold change. By 7 days post-infection the difference is even greater: 7.64 in the Ebola-naïve group, and 3.36 in the vaccinated group. We observe that every gene making up the early transcriptional response shows significantly lower levels of expression in vaccinated animals in comparison to Ebola-naïve animals.

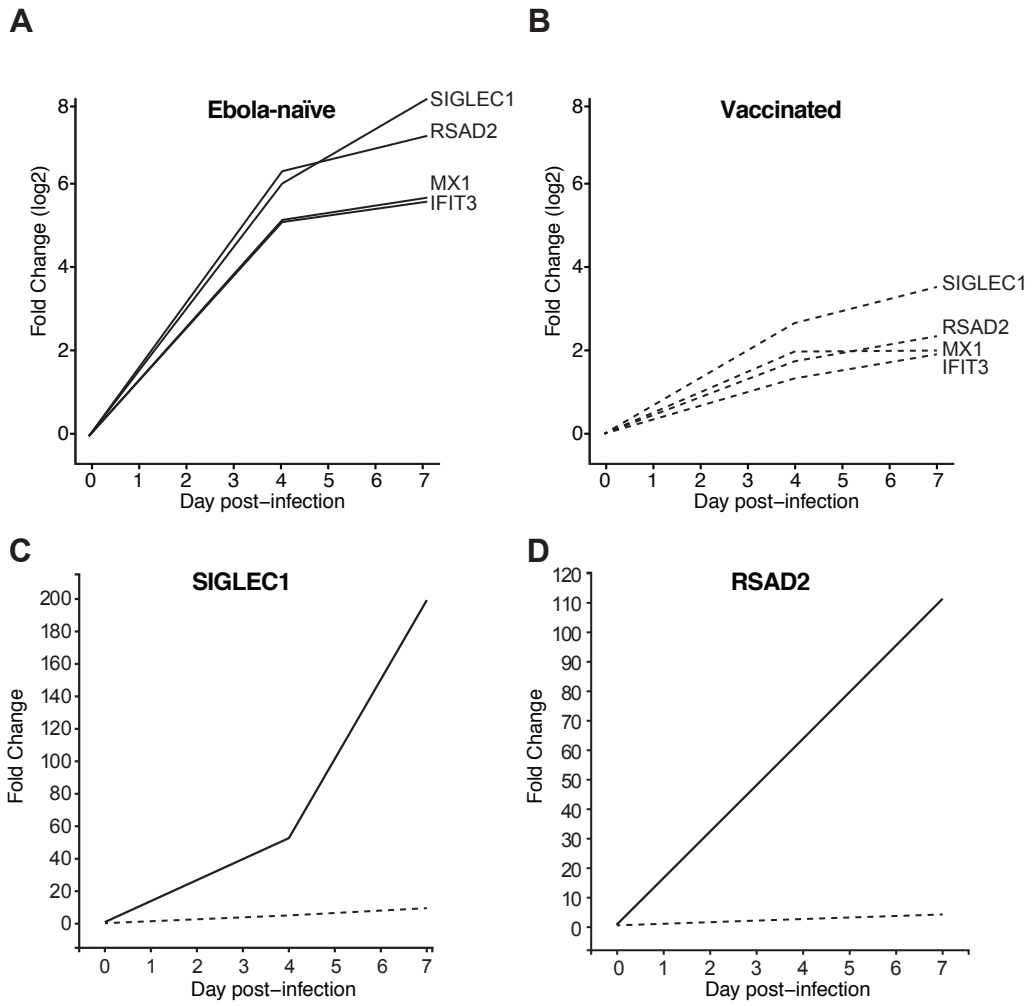


Figure 3.2: Ebola-naïve and vaccinated responses to Ebola infection.
Each line represents the average fold-change in expression (log2, y-axis) of MX1, RSAD2, SIGLEC1, and IFIT3 at different times (x-axis) after Ebola infection. Panel A) shows genes from Ebola-naïve macaques, and Panel B) shows genes from vaccinated animals. Panels C) and D) show the fold-change of SIGLEC1 and RSAD2 in vaccinated (dashed lines) and Ebola-naïve macaques (solid lines).

The expression levels of most cytokines during Ebola infection in vaccinated animals remained constant, with the exception of three cytokines that were significantly upregulated (CCL7, CXCL10 and CXCL11). These genes mirrored the expression patterns seen in the Ebola-naïve animals, but at a lower magnitude (Figure 3.3). For example, by 4 days post-infection the expression level of CXCL11 went up 4.5 log2-fold in the Ebola-naïve group, but only 2.8 log2-fold in the vaccinated group.

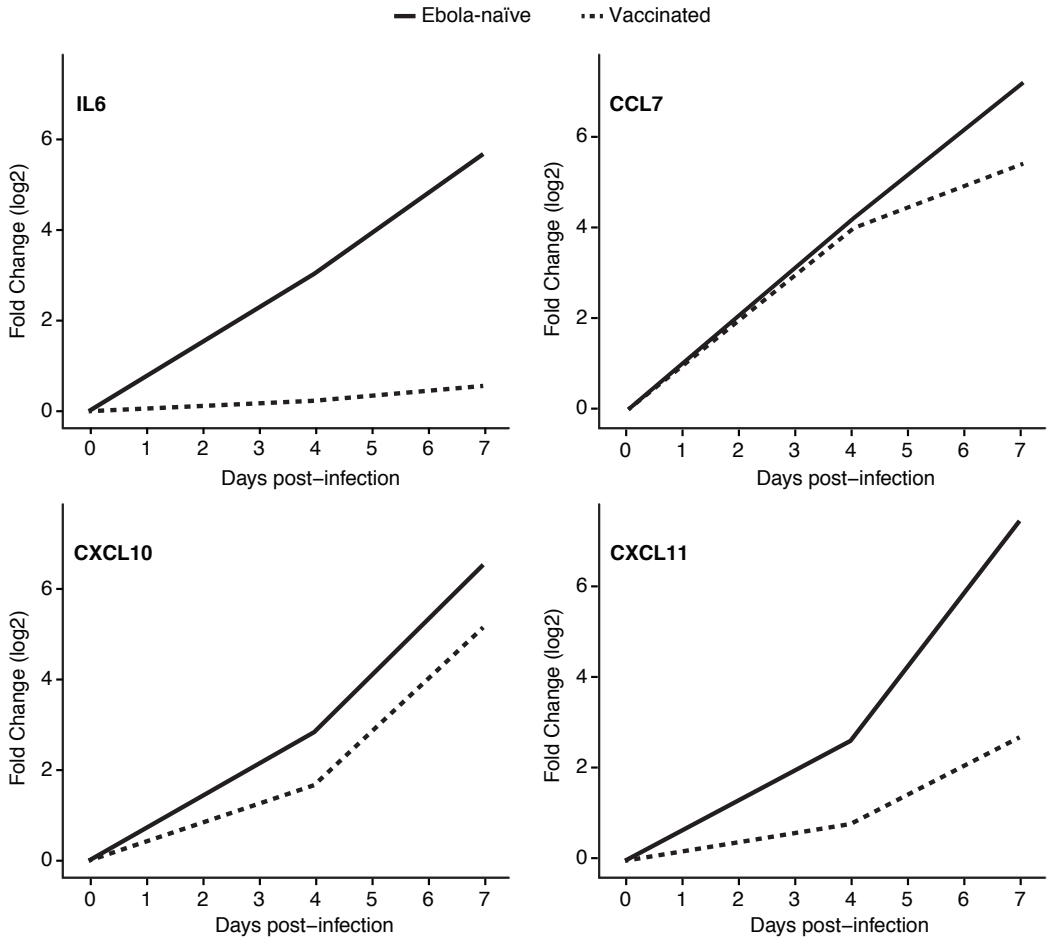


Figure 3.3: Effects of vaccination in the expression of cytokines after Ebola infection.
Each line represents the average expression (log10 counts per million, y-axis) of three cytokines that become upregulated in vaccinated animals: CCL7, CXCL10, and CXCL11 at different times (x-axis) after Ebola infection. IL6 is shown to illustrate the behavior of

a cytokine that only becomes expressed in Ebola-naïve animals. Genes from Ebola-naïve macaques are shown in solid lines while those from the vaccinated animals appear in dashed lines.

3.4.3 Cytokine responses during Ebola infection

In addition to our analysis of early-expressed genes, we were also interested in understanding the cytokines expressed during infection. Previous studies have focused on a protein-level analysis in the serum and reported an increase in the concentration of a number of cytokines after Ebola infection (Ebihara et al., 2011; McElroy & Erickson, 2014; Wauquier et al., 2010). We observe that a subset of these cytokines reported to be present in the serum also show transcriptional upregulations in PBMCs by 4 days post-infection (IL6, IL1B, IL1RN, CXCL10) and by 7 days post-infection (CCL2, CCL3, CSF1, TNF, CXCL1, and CXCL8). Figure 3.4A shows the average fold-change (log2) for each of these cytokines on the y-axis, as well as the time after infection when the samples were taken on the x-axis. These results suggest a biphasic cytokine response to infection. IL1B is the only cytokine with an unusual transcription pattern, peaking at day 4 and returning to pre-infection levels by day 7.

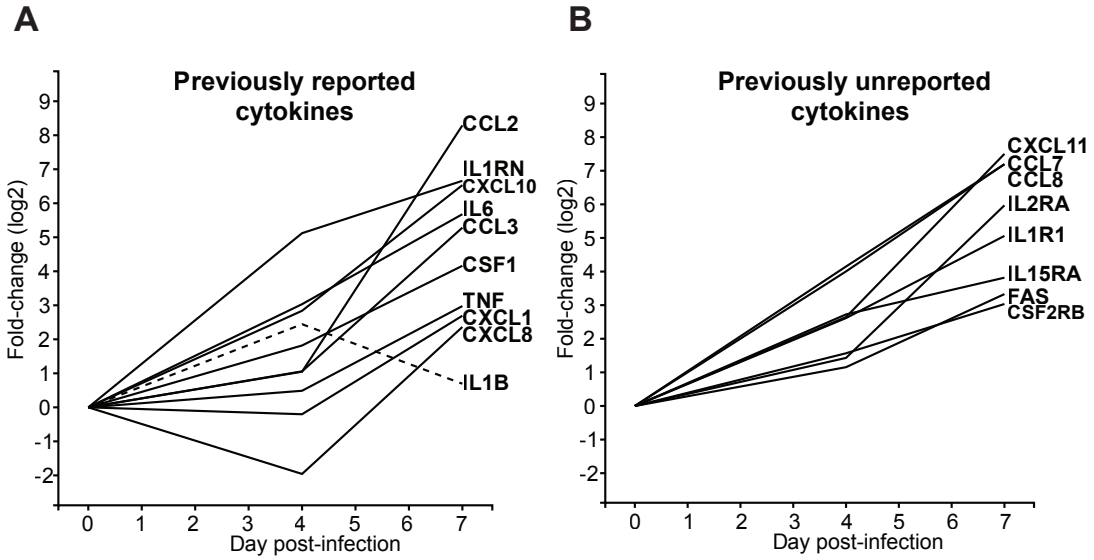


Figure 3.4: Changes in expression of cytokines during Ebola infection.
Each line represents the average fold-change in expression (log2, y-axis) of each cytokine gene at different times (x-axis) after Ebola infection. Panel A) shows ten cytokines that undergo strong transcriptional changes and that were previously reported to experience changes in serum concentration during Ebola infection. IL1B (dashed stroke) shows a transient level of expression, peaking 4 dpi. Panel B) shows seven cytokines that follow similar patterns of expressions to those in Panel A) but that have not been previously reported to play a role during Ebola infection.

Additionally, we identified several cytokines that were not previously known to play a role during Ebola infection. Among these, some became upregulated during the early stage of infection (day 4) CCL7, CCL8, CXCL11, IL1R1 and IL15RA, and the ones that do so during the late stage of infection (day 7) include the receptors FAS, IL2RA, and CSF2RB (Figure 3.4B).

3.4.4 The early transcriptional response is common to multiple hemorrhagic fevers

We wanted to compare the early host transcriptional response to Ebola infection to the early response that was previously noted in other hemorrhagic fevers (Caballero et al., 2014). To do this, we used gene expression data from cynomolgus macaques exposed via aerosol to either Lassa virus (Caballero et al., 2014; Malhotra et al., 2013) or Marburg virus (25, J. H. Connor, submitted), and found that the innate response to both infections is highly similar to that of Ebola infection. This is illustrated in Figure 3.5, which shows the changes in expression of four canonical ISGs (MX1, ISG15, DHX58 and OAS1) during Ebola, Lassa, and Marburg infection. All four genes are significantly upregulated at the earliest infected timepoint, and they remain sustained throughout the late stage of disease. For example, MX1 goes up 4.7 log2-fold 3 days after Lassa infection, 4.1 log2-fold 3 days after Marburg infection, and 4.9 log2-fold 4 days after Ebola infection. By 6-7 days post-infection, the log2-fold changes are 6.1, 6, and 5.4, respectively. By 9-10 days post-infection, the expression of MX1 seems to decrease mildly in Lassa and Marburg infection, but it remains at very high levels compared to the pre-infection baseline. During Lassa infection, ISG15 and OAS1 undergo similar expression changes to those of MX1: a strong increase, followed by a slight decrease in expression. For Marburg and Ebola infection, the early levels of expression of these genes increase more rapidly than during Lassa infection, and this trend continues during the late stage of disease. These data supports the hypothesis that the immune system responds to different viral infections via a common, early, sustained, and strong innate immune response.

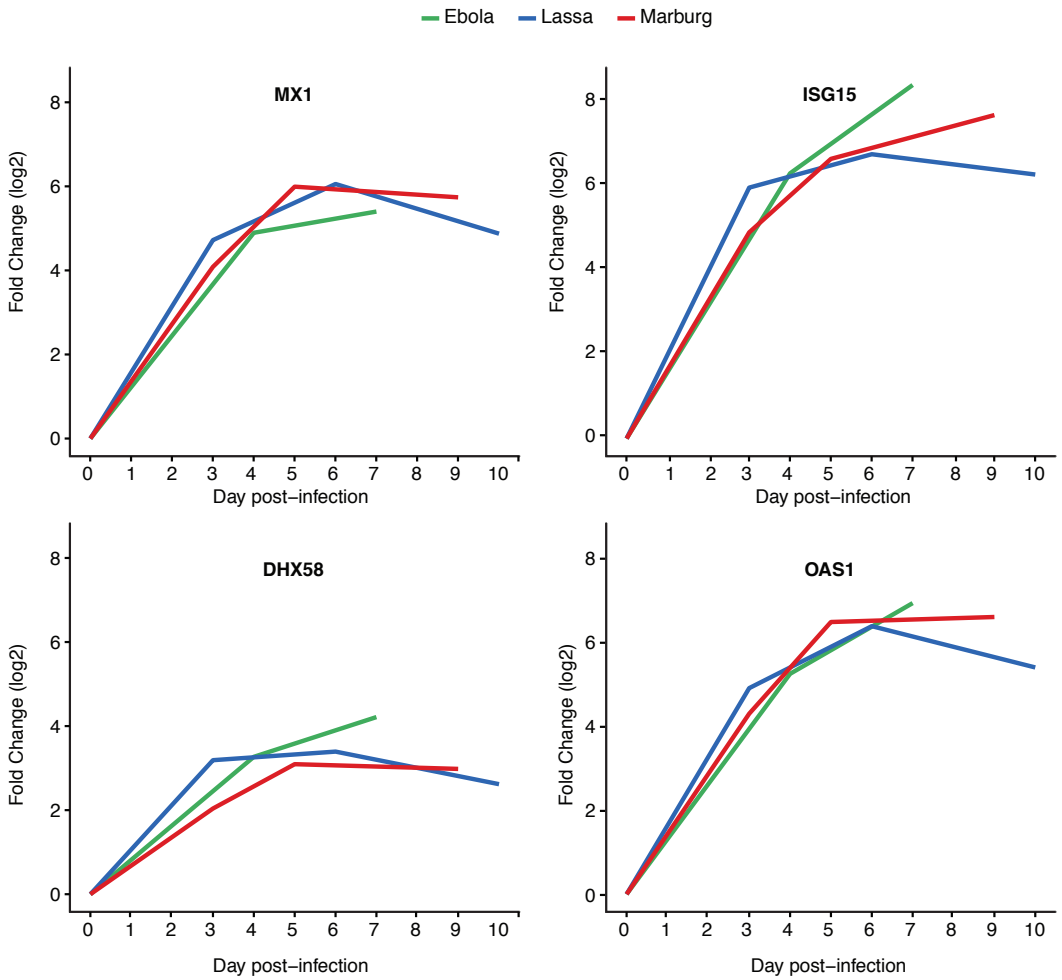


Figure 3.5: Changes in expression of interferon-stimulated genes during different hemorrhagic fever infections.
Each line represents the average fold-change in expression (log2, y-axis) of MX1, ISG15, DHX58 or OAS1 at different times (x-axis) after Ebola (green), Lassa (blue) or Marburg (red) infection.

3.4.5 Different routes of infection lead to a similar early transcriptional response

To determine if the route of infection could alter the early transcriptional response to Ebola infection, we looked at gene expression data from a study that exposed

macaques to Ebola virus via aerosol and we compared it to the intramuscular injection data. Figure 3.6 shows the expression of twenty genes that have the strongest changes 4 days post-infection in the intramuscular group (Figure 3.6A), as well as the pattern of expression of those same genes under aerosol exposure (Figure 3.6B). RNA sequencing was used to quantify the expression in the intramuscular group, while microarrays were used for the aerosol group. We find that both groups show a strong increase in expression starting at 3-4 days post-infection and increasing during the course of disease.

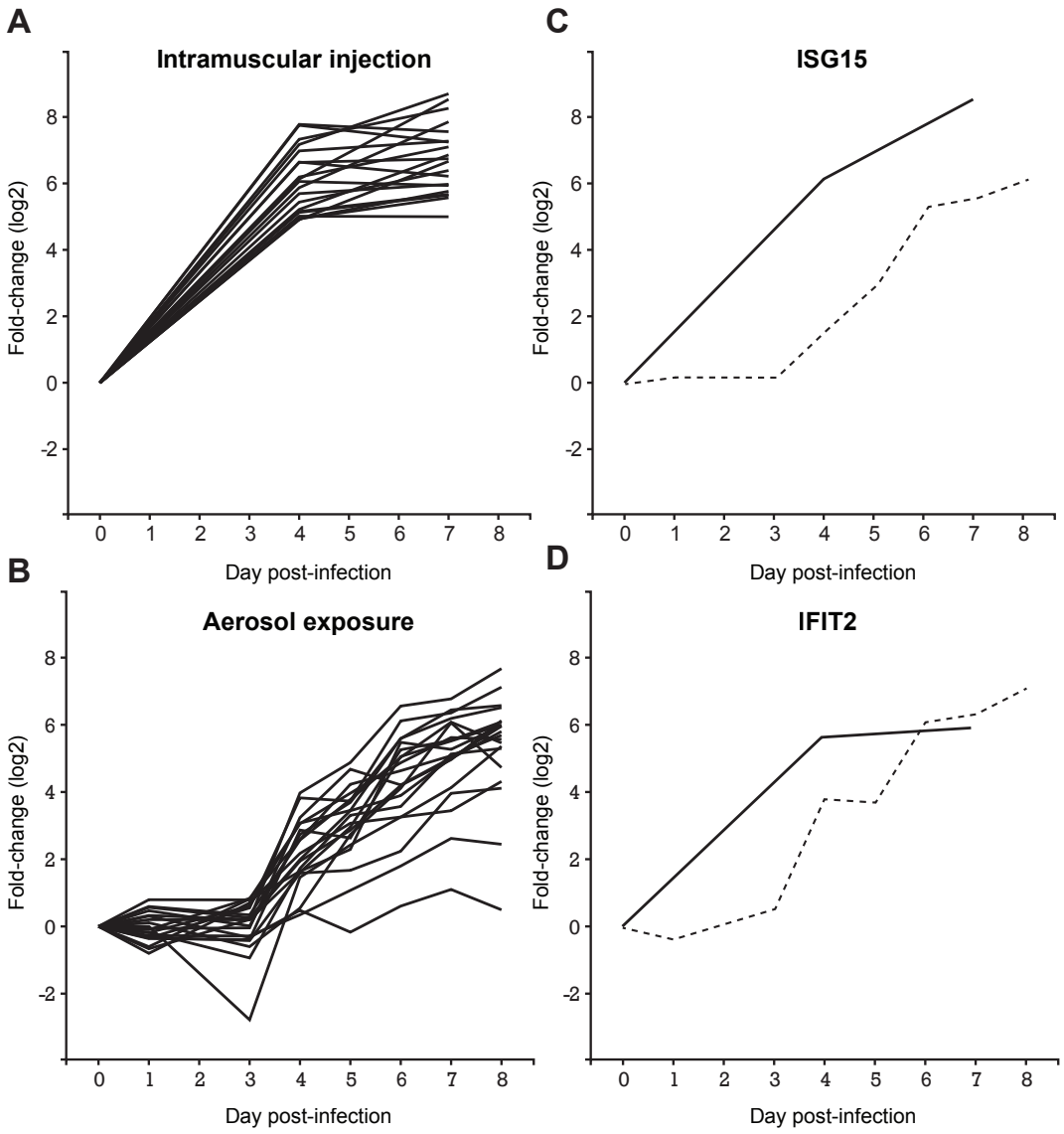


Figure 3.6: Different routes of infection lead to similar transcriptional responses. Each line represents the average fold-change in expression (log2, y-axis) of each gene at different times (x-axis) after Ebola infection. Panel A) corresponds to intramuscular injection (quantified using RNA sequencing), and Panel B) to aerosol exposure (quantified using microarrays). Panels C) and D) illustrate the expression observed in two individual genes (ISG15 and IFIT2) in the aerosol (dashed) and intramuscular groups (solid).

3.4.6 The innate immune response takes place across most tissues

We next compared the expression patterns of the 109 genes that showed the strongest upregulation in PBMCs to their expression patterns in multiple tissues in the macaque model of aerosol exposure to Ebola virus. We found that the early immune response is present across most tissues at varying magnitudes and expression rates. Figure 3.7 shows the average expression rate across each tissue type. PBMCs become activated at 6 days post-infection (Figure 3.4), and show the strongest activation of all tissues at 7 days post-infection. Liver starts becoming activated 3 days post-infection and the average expression level continues to increase by 5 and 8 days post–infection. The gene expression response in the spleen begins to increase at day 3 and also increases to similar levels as in the liver. In the adrenal gland and the pancreas, the expression level increases at 4 days post-infection and remains activated until the end of the infection. The axillary lymph node is the last to become activated, starting around day 6. The brain shows the lowest levels of activation, with a slight increase starting on days 4-6.

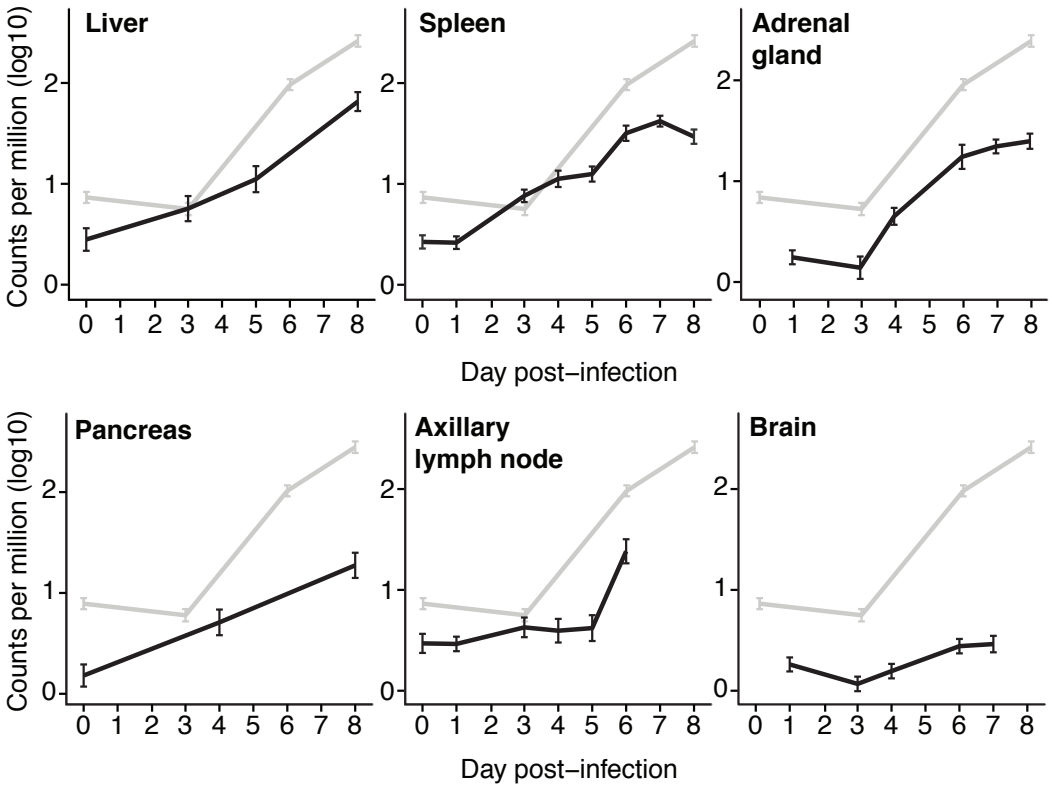


Figure 3.7: Innate transcriptional response in tissues in an aerosol model of infection.
Grey lines represent the average expression values of the 111 genes most highly expressed in PBMCs (grey line). Black lines represent the average expression across all tissues for these same genes. The y-axis represents log10 read counts per million, and the x-axis different times after infection.

3.5 Discussion

One of the most significant results from our analysis of the PBMC response to Ebola infection was the strong activation of innate immune response genes at early times post-infection, most of which are classified as interferon-stimulated genes (ISGs). This finding is consistent with other analyses of host responses to different viral infections

(Caballero et al., 2014; Fink et al., 2007; A. K. Zaas et al., 2009), but in the case of Ebola infection, it is an unexpected result. Previous studies have shown that Ebola virus infection *in vitro* blocks the expression of ISGs in liver cells (Hartman, Ling, et al., 2008; Kash et al., 2006). This interferon-antagonism is based on the ability of VP35 to block the activation of IRF3, which inhibits the expression of interferon beta and other ISGs. Mutation of VP35 in a manner that allows IRF3 signaling has been studied in mice (Hartman, Bird, et al., 2008) and guinea pigs (Prins et al., 2010) using a recombinant Ebola strain containing a single-point mutation in VP35 (R322A). This mutation reduced the virus's ability to replicate and prevented it from inhibiting the interferon response *in vitro* and from leading to pathogenesis in these animal models. These studies are consistent with the hypothesis that Ebola virus causes a systemic inhibition of the interferon response.

A caveat of these studies is that they did not measure the innate immune response during *in vivo* wild type Ebola virus infection. Our analysis shows that almost all tissues examined (with the exception of the brain) undergo a strong expression of interferon-stimulated genes. This implies that an innate immune trigger of IRF3 or interferon signaling is expressed and released into the circulation. Our analysis supports the hypothesis that unlike what happens in an *in vitro* infection, Ebola virus infection leads to a robust, early, and global innate response *in vivo*.

An explanation for why we detect a significant increase in the expression of IRF3 or interferon-stimulated genes at the early stages of Ebola infection could be that infected cells stimulate uninfected neighboring cells to produce interferon through a currently

unknown mechanism, and that this stimulates them to express ISGs and to continue spreading the interferon signal (

Figure 3.8A). Since we see viral titers and ISG expression increasing throughout infection, further studies are required to understand if the populations of cells that are expressing ISGs are different from the ones that are undergoing viral replication. Another hypothesis is that, *in vivo*, some cells may retain their ability to translocate IRF3 to the nucleus and to express interferon, perhaps due to a missing VP35 (

Figure 3.8B). This is consistent with earlier reports showing that there are early immune-associated transcriptional responses in primary target cells exposed to either Ebola virions or VLPs (Wahl-Jensen et al., 2011).

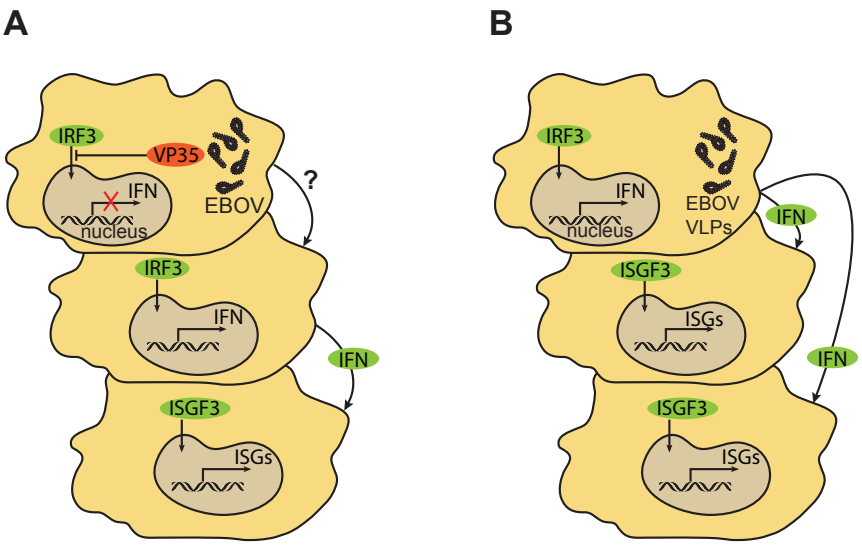


Figure 3.8: Model for the expression of interferon-stimulated genes during Ebola infection.
A) Ebola-infected cell (top) is not able to produce interferon due to the VP35-inhibition of IRF3 translocation. We suggest that the infected cell, through an unknown mechanism, might be able to induce neighboring cells (middle) to translocate IRF3 to the nucleus and start producing interferon. Once interferon is released by neighboring cells, it activates the receptors of additional cells (bottom) and leads to the transcription of ISGs. B) An

alternative model is that some cells (top) can become infected with Ebola VLPs, which are not able to block IRF3 translocation, and therefore they can produce interferon, release it to neighboring cells (middle and bottom) and they in turn start transcribing ISGs.

One of the hallmarks of Ebola infection is dysregulated levels of circulating proinflammatory cytokines. We observed the levels of expression of previously reported cytokines such as IL6, and CXCL8 (IL8) increase between 4 and 7 days post-infection. We also found several cytokines, such as CCL7 and CXCL11 (I-TAC), that showed significant changes in expression but whose protein levels have not been previously measured during Ebola pathogenesis. CXCL10 (IP-10) and CXCL11 were two of the cytokines that showed a significant increase in expression in vaccinated macaques infected with Ebola virus. This suggests that they are part of a conserved response in both the pathogenic and non-pathogenic response to Ebola. Both cytokines are induced by interferon alpha (ISGs), share a common receptor (CXCR3), and are thought to be involved in the recruitment of effector T cells and NK cells (Groom & Luster, 2011). They are also known to be induced during the acute phase in other types of viral infections including dengue (Hsieh et al., 2006), influenza (Fadel, Bromley, Medoff, & Luster, 2008), hepatitis B and C (Dufour et al., 2002; Zeremski et al., 2008), Herpes simplex (Nakanishi, Lu, Gerard, & Iwasaki, 2009) and HIV-1 (Brainard et al., 2007). Upregulation of CXCL10 has also been associated with hemorrhagic manifestations in patients infected with Sudan virus (McElroy & Erickson, 2014). Our analysis suggests that these cytokines may be part of an innate immune response that is seen in both vaccinated and Ebola-naïve animals.

By comparing how different routes of infection affect the cellular circulating immune response in Ebola-infected primates, we observed that infection via an aerosol and via intramuscular (IM) injection resulted in similar patterns of gene expression. Since the earliest samples taken in each group differed by one day (4 dpi in the intramuscular group, and 3 dpi in the aerosol group), it is not possible to determine if there was a time delay.

Our results extend earlier observations of a strong innate immune response and suggest the involvement of new cytokines in Ebola infection. Further analysis of the cells responsible for driving this response and for producing the different cytokine signals will be important to understand the ability of the virus to replicate virtually unchecked in many tissues, and to identify what cells are undergoing that uncontrolled ISG response.

CHAPTER 4

VALIDATION OF HOST IMMUNE RESPONSES ASSOCIATED WITH VIRAL HEMORRHAGIC FEVERS

4.1 Introduction

Hemorrhagic fever viruses are major human pathogens that often result in severe disease and high mortality rates. The mechanisms of pathogenesis of these viruses are not well characterized, but they are thought to be caused by the dysregulation of normal host immune responses. One method of determining the nature of the host response to infection is to identify the gene expression changes in the circulating immune system after infection.

In chapters 2 and 3, I showed that the circulating immune cells of macaques infected with Lassa and Ebola virus show strong increases in the expression of interferon-stimulated genes and other infection-specific genes early after infection. Validating these findings requires analyzing independent datasets to ensure that the detected patterns are reproducible. In this chapter, I compare the patterns of expression of two datasets: one from macaques infected with different strains of the arenaviruses Lassa virus and Lujo virus, and another dataset from macaques infected with Ebola virus.

The greatest strength of the arenavirus dataset is that it makes it possible to compare the patterns of expression across three different strains of the Lassa virus, including the Josiah strain, which I previously analyzed using RNA sequencing. Although this dataset does not have any pre-infection samples, the earliest samples were

taken 1 day post-infection. Based on previous analyses, the patterns of expression at this time are still very similar to those of pre-infection samples. The main advantage of the Ebola dataset is that samples were taken every day between pre-infection and 6 days post-infection. This increases the resolution and allows for a more comprehensive view of the patterns of expression throughout disease. Unfortunately, the microarray platform that was used to quantify this expression is missing probes for several thousand genes, and the efficiency of many of the existing probes is likely to not be optimized. Regardless of the limitations of these datasets, they can both provide useful information that can be used to validate the patterns of expression that are associated with viral hemorrhagic fevers.

4.2 Methods

4.2.1 Macaque model of intramuscular exposure to different Lassa virus strains and Lujo virus

As described in (Rasmussen et al., 2014), cynomolgus macaques were infected via intramuscular injection with either Lujo virus or with one of three different Lassa virus strains (Josiah (Wulff & Johnson, 1979), Z132 (P B Jahrling, Frame, Smith, & Monson, 1985; Safronetz et al., 2010), Soromba (P B Jahrling et al., 1985)). Blood samples were taken at 1, 4, 7, and 10 days post-infection (n=3 per timepoint/condition). Additionally, pre-infection samples were also obtained for the Lujo-infected group. Peripheral blood mononuclear cells were isolated from these samples, RNA was extracted and hybridized

to Agilent Rhesus macaque one-color microarrays. The microarray data was processed using limma (Smyth, 2004) and normalized using LOESS (Smyth & Speed, 2003). The raw data is available in GEO (Barrett et al., 2013) under accession GSE49838.

4.2.2 Macaque model of Ebola virus

Fifteen cynomolgus macaques were inoculated via intramuscular injection with Ebola virus (Rubins et al., 2007). Blood samples were taken before infection and every day between days 1 and 6 post-infection. RNA was extracted and hybridized to Agilent Rhesus macaque two-color microarrays with probes for 18,000 genes. The microarray data was processed using limma (Smyth, 2004) and normalized using LOESS (Smyth & Speed, 2003). Expression values were obtained by calculating the log-ratio between the intensities of the red channel (which corresponds to experimental samples) and the green channel (which corresponds to Human Universal Reference RNA (Novoradovskaya et al., 2004)). The raw data is available in GEO (Barrett et al., 2013) under accession GSE8317.

4.3 Results

4.3.1 Infection with different strains of Lassa virus leads to similar host immune transcriptional responses

To determine the effects that different strains of Lassa virus would have in the immune response, I compared the host patterns of expression from two different studies where macaques were infected with different Lassa virus strains. The first study used

RNA sequencing to quantify the host gene expression of macaques infected with the Josiah strain of Lassa virus (Chapter 2, Caballero et al., 2014). The second study used microarrays to quantify the expression of macaques infected with one of three strains of Lassa virus (Josiah, Soromba and Z132) and Lujo virus, a virus from the same viral family as Lassa (Arenaviridae) (Rasmussen et al., 2014). Only 7 genes, out of the 12 that we had previously identified in Chapter 2, have corresponding microarray probes. Figure 4.1 shows the patterns of expression of these genes, along with the housekeeping gene TPT1, at several times after infection with the Josiah strain of Lassa virus. The dark blue lines correspond to the log2 fold-changes in expression for a specific gene in the RNA sequencing dataset, while the light blue lines correspond to those in the microarray dataset. OAS1, SIGLEC1 and TNFSF10 are three of the genes that increase their levels of expression early after Lassa virus infection. IGJ, BIRC3, TNK2 and NR4A2 are four of the genes whose levels of expression remain relatively constant or decrease after infection. SIGLEC1 seems to show an increase in expression 3-4 days post-infection in both platforms, but the associated microarray probes report a faster decrease of expression than that quantified via RNA sequencing. Conversely, IGJ shows a sharp decrease in the microarray platform, while the early fold change reported in the RNA sequencing study does not appear to vary.

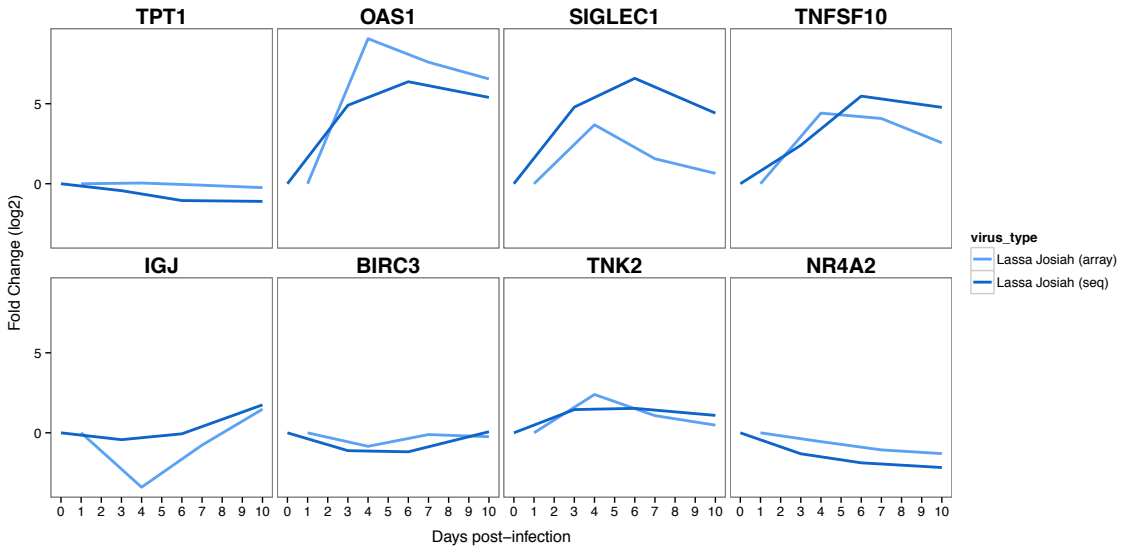


Figure 4.1 Gene expression in an independent dataset of samples infected with Lassa virus (Josiah strain).
Each line represents the average fold-change in expression (log2, y-axis) of a gene at different times (x-axis) after infection with Lassa virus (Josiah). Dark blue lines represent sequencing samples from (Caballero et al., 2014), and light blue lines correspond to microarray samples from (Rasmussen et al., 2014).

Once we identified a high level of agreement between the RNA sequencing and microarray platforms for samples infected with the Josiah strain of Lassa virus, we looked at two additional Lassa virus strains, Soromba and Z132. Figure 4.2 compares the level of expression of these three strains. The dark blue line represents the log2 fold-change of Josiah-infected samples, quantified using RNA sequencing, the light blue and purple lines represent the log2 fold-changes of Soromba-infected and Z132-infected samples, respectively. The upregulations of OAS1, SIGLEC1, and TNFSF10 are also seen in these two additional strains. Both show very similar patterns except for IGJ, BIRC3 and NR4A2, which show slight differences in expression, especially 4 days post-infection.

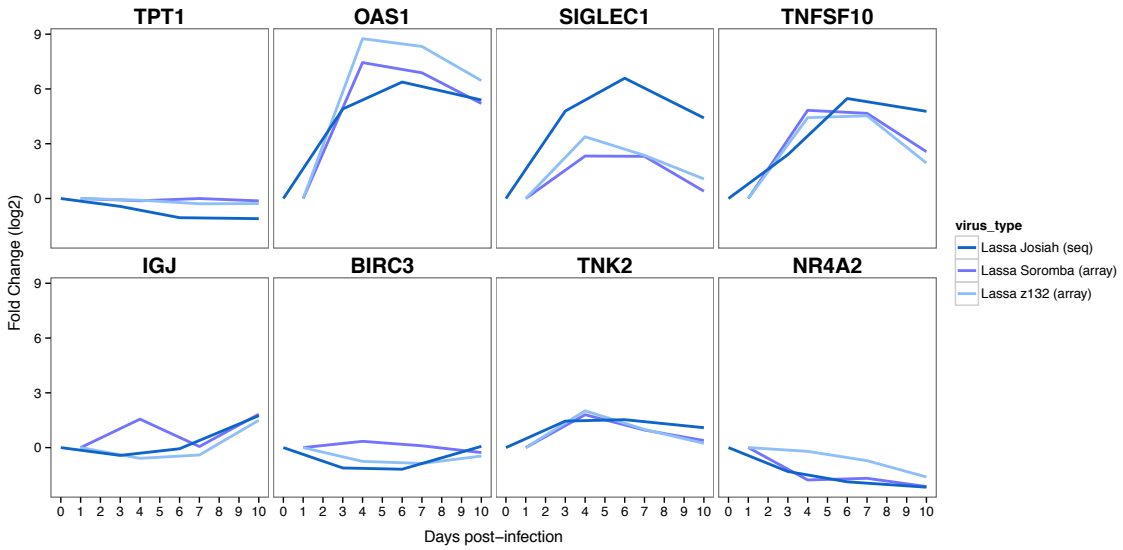


Figure 4.2 Gene expression in an independent dataset of samples infected with Lassa virus (Soromba and Z132 strains).
Each line represents the average fold-change in expression (log2, y-axis) of a gene at different times (x-axis) after infection with Lassa virus. Dark blue lines represent sequencing samples infected with Lassa virus (Josiah) from (Caballero et al., 2014). The other lines correspond to microarray samples from (Rasmussen et al., 2014). Light blue lines correspond to samples infected with Lassa virus (Z132). Purple lines correspond to samples infected with Lassa virus (Soromba).

Finally, we wanted to determine how these similarities that we identified across three different strains of Lassa virus would compare to the patterns of expression of a different arenavirus. Figure 4.3 shows the same genes as in the previous figures, but it compares Josiah-infected samples quantified using RNA sequencing with those infected with Lujo virus and quantified using microarrays. OAS1 shows a similar pattern of upregulation as that seen in the different Lassa strains, but SIGLEC1 and TNFSF10 show lower-magnitude and time-delayed upregulations compared to the Lassa strain values. The low-expression genes show similar patterns of transcription after Lujo virus and Lassa virus infection.

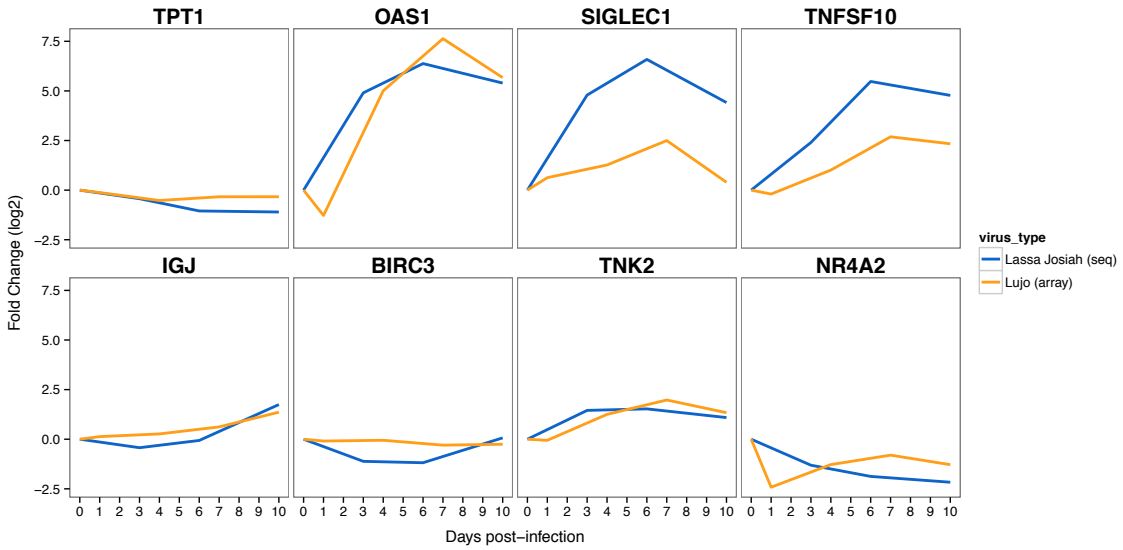


Figure 4.3 Gene expression in an independent dataset of samples infected with Lujo virus.
Each line represents the average fold-change in expression (log2, y-axis) of a gene at different times (x-axis) after infection. Dark blue lines represent sequencing samples infected with Lassa virus (Josiah) from (Caballero et al., 2014). The other lines correspond to microarray samples from (Rasmussen et al., 2014). Orange lines correspond to samples infected with Lujo virus.

The large fold changes in the expression of these genes after infection indicate strong transcriptional differences between uninfected and infected samples. To determine the magnitude of these differences, we applied Primary Component Analysis to the levels of expression of these seven genes and found a clear difference between both groups. Figure 4.4 shows each sample as a point, and the distances between them represent a measure of the level of similarity among the samples. The first two principal components explain 84.46% of the variance of the dataset, and show a clear separation between the uninfected sample cluster (in grey) and the infected samples (in colors), which do not separate into any distinguishable clusters using these markers.

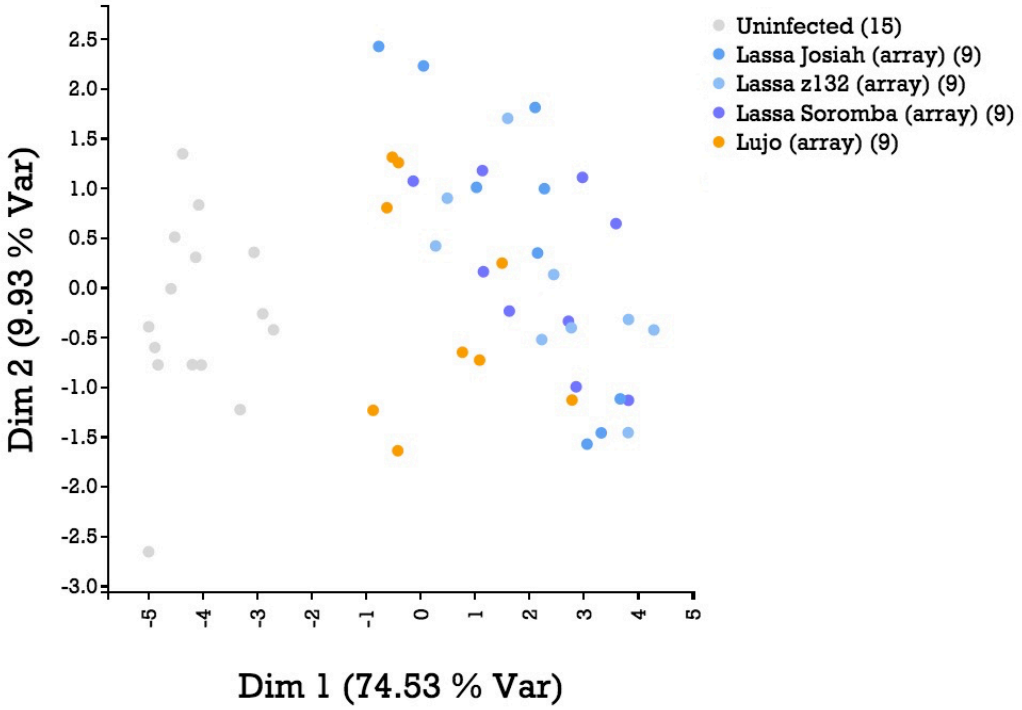


Figure 4.4 Primary Component Analysis of an independent dataset of Arenavirus microarray samples.
Each point corresponds to an RNA sample from (Rasmussen et al., 2014). Lassa-infected samples are shown in shades of blue and purple (Josiah, Z132, and Soromba strains). Lujo-infected samples are shown in orange. Uninfected samples are shown in grey. Distances between points represent the similarity among samples based on the application of Primary Component Analysis (PCA) to the microarray expression values of 7 genes previously identified to be upregulated after Lassa virus infection: OAS1, IGJ, BIRC3, SIGLEC1, TNK2, TNFSF10, NR4A2.

4.3.2 Ebola virus infection leads to the production of robust patterns of host gene expression

To determine if the transcriptional patterns that we identified in Chapter 3 could be reliably detected in an independent study, we analyzed data from a previous study that infected macaques with Ebola virus and used microarrays to measure their gene

expression responses after infection (Rubins et al., 2007). We compared these responses those reported in (Barrenas et al., 2015), which also infected macaques with Ebola virus, but used RNA sequencing to quantify the post-infection patterns of expression. Figure 4.5 shows the log2 fold-changes in expression throughout infection of multiple genes in both datasets. The dark green lines correspond to gene expression changes quantified by RNA sequencing, while the light green lines correspond to expression changes quantified by microarrays. Only genes with corresponding microarray probes were chosen. The Actin B gene (ACTB) is shown first for reference. MX1 and OAS1 are two interferon-stimulated genes that increase quickly and strongly after viral hemorrhagic fever infection. TNFSF10, VCAN, TNK2 and BIRC3 are genes that show an increase in expression around 4 days post-infection in both the sequencing and microarray datasets. Conversely, IGJ shows a decrease in expression in both platforms. The genes HSPA1L, S100A8, HSPA1B and ADAM28 appear to become strongly upregulated in the sequencing dataset, but this upregulation is not present when the data is quantified using microarrays.

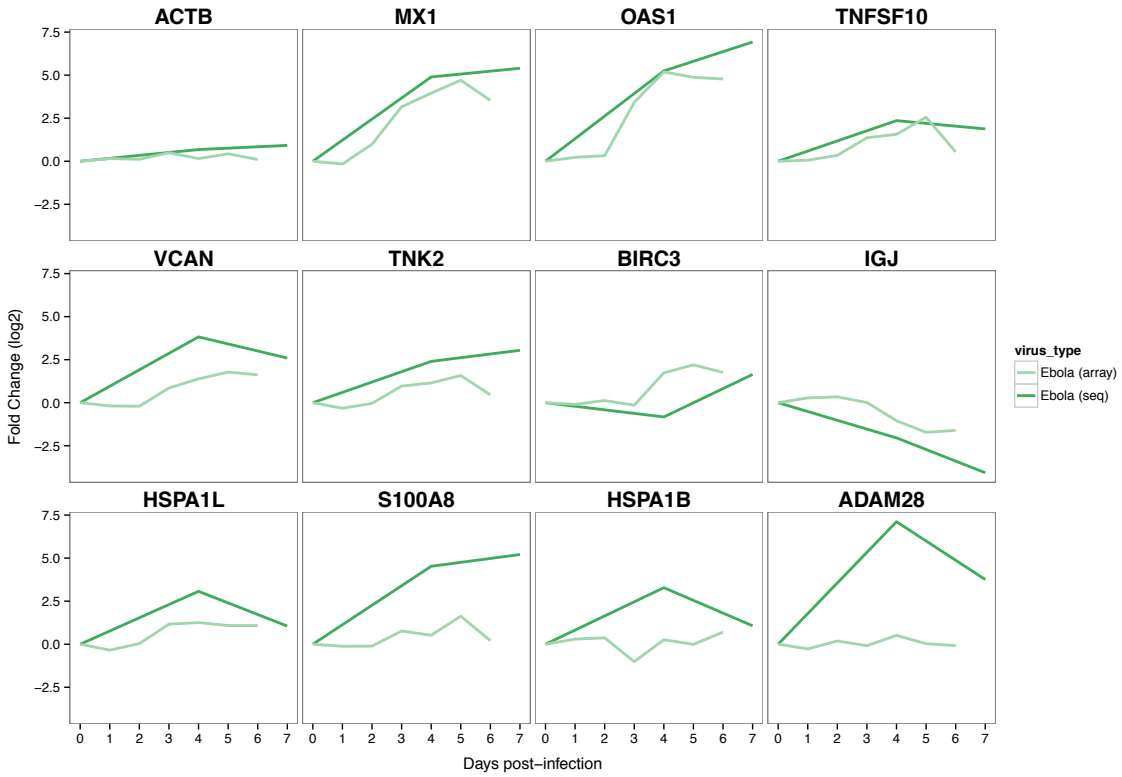


Figure 4.5 Gene expression in an independent dataset of samples infected with Ebola virus.
Each line represents the average fold-change in expression (log2, y-axis) of a gene at different times (x-axis) after infection with Ebola virus. Dark green lines represent sequencing samples from (Barrenas et al., 2015). Light green lines correspond to microarray samples from (Rubins et al., 2007).

Although comparing the sequencing and microarray datasets showed large discrepancies in the expression values of several genes, we wanted to know if the expression values of this set of genes could be used to distinguish between infected and uninfected samples. To verify this, we applied Primary Component Analysis to the levels of expression of these eleven genes and identified a clear difference between both groups. Figure 4.6 shows each sample as a point, and the distances between them represent a measure of the level of similarity among the samples. The first two principal components

explain 79.38% of the variance of the dataset and show two clear clusters. The first cluster is composed of uninfected samples (in grey) and infected samples obtained 1-2 days post-infection (increasingly darker shades of green), before the host transcriptional response gets activated. The second cluster is composed of infected samples obtained 3-6 days post-infection (increasingly darker shades of green), with the 3 dpi samples situated between both clusters along the first principal component (x-axis).

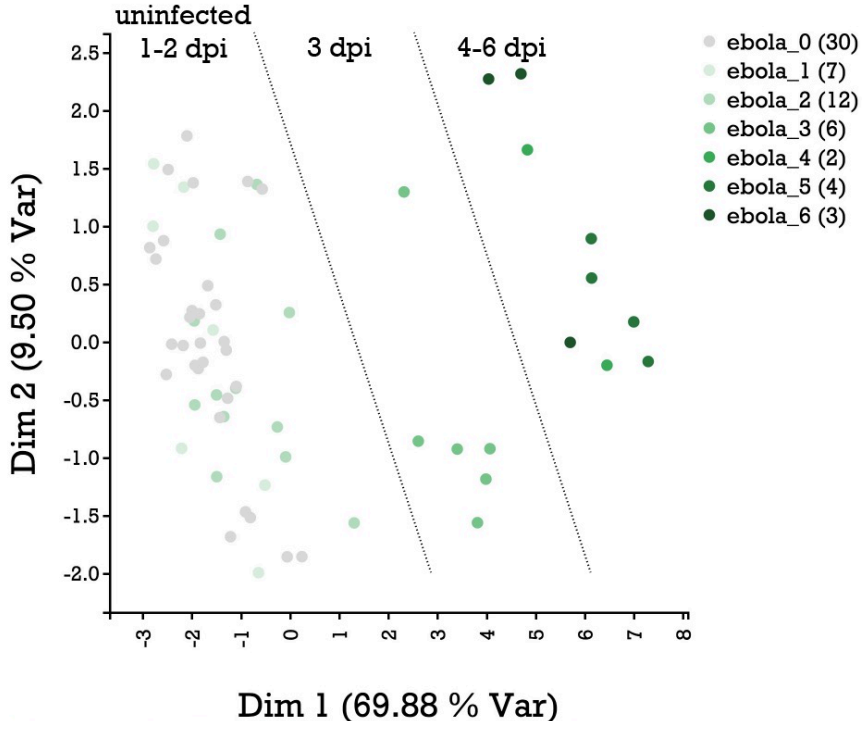


Figure 4.6 Primary Component Analysis of an independent set of Ebola microarray samples.
Each point corresponds to an RNA sample from (Barrenas et al., 2015). Ebola-infected samples are shown in increasingly darker shades of green. Uninfected samples are shown in grey. Distances between points represent the similarity among samples based on the application of Primary Component Analysis (PCA) to the microarray expression values of 11 genes previously identified to be upregulated after Ebola virus infection: MX1, OAS1, HSPA1B, IGJ, HSPA1L, BIRC3, TNK2, TNFSF10, ADAM28, VCAN, and S100A8.

4.4 Discussion

I previously identified a set of genes that undergo unique patterns of expression early after infection with a hemorrhagic fever virus. To ensure that this transcriptional patterns could be reliably detected across a number of experimental conditions, I analyzed two publicly available datasets of cynomolgus macaques infected with different strains of Lassa virus, Lujo virus, or Ebola virus, and compared the patterns of expressing of a subset of genes to those that showed significant changes in expression in the original datasets (Chapters 2 and 3).

Comparing the sequencing and the microarray datasets for the Lassa Josiah strain showed strong agreement in genes that increase 3-4 days post-infection, as well as among genes that remain constant or decrease after infection. A similar result was seen for Lassa Soromba and Z132 strains. The Lujo virus immune response showed similar patterns of upregulation compared to the response to Lassa virus, but with a delay of 72-hours. This is likely a result of the evolutionary distance between both viruses, they are similar enough to be classified in the same viral family, but different enough to be distinct viruses. Several genes that showed strong patterns of infection in the Lassa sequencing dataset could not be compared in the microarray platforms because the corresponding probes for these genes were not present.

In the Ebola dataset, the differences between the gene patterns in the microarray and sequencing platforms were larger than in the Lassa dataset. A number of genes showed strong agreement (MX1, OAS1), while others did not show similar patterns (HSPA1B, ADAM28). Similar to the Lassa dataset, several genes in the Ebola

microarrays lacked corresponding probes, and their expression could not be compared to the sequencing data. This lack of a unified set of genes, along with imperfect probe-gene complementarity, are likely the major factors that explain the lack of agreement between both platforms.

Comparing both platforms has shown that the genes identified in the sequencing datasets can be used to distinguish among infected and uninfected microarray samples (Figure 4.4 and Figure 4.6). This provides additional confidence that these patterns are strong and reliable indicators of an early immune response against these viruses.

CHAPTER 5

THE MASTER REGULATOR OF THE CELLULAR STRESS RESPONSE IS CRITICAL FOR ORTHOPOXVIRUS INFECTION

5.1 Preamble

In this chapter I describe work that I carried out as part of a collaborative effort to define the role of cellular host proteins in poxvirus replication. My contribution to the project was to analyze two different RNA sequencing datasets to identify the expression of host genes after 4-6 hours post infection with Vaccinia virus. I also compiled a list of chaperones and HFS1 targets from previous studies and showed that the majority of upregulated genes in both datasets belonged to these categories. The majority of the results presented in this chapter were previously published in (Filone et al., 2014).

5.2 Introduction

The Poxviridae family is comprised of several human pathogens in the Orthopoxvirus genus, including monkeypox (MPXV) and smallpox (Variola), which was eradicated through vaccination with vaccinia (VACV). With a dramatic increase in human MPXV cases in Africa, the rise of VACV-like orthopoxvirus infection in South America, and concerns about the weaponization of smallpox, it is important to design new strategies for the treatment and prevention of these diseases (Damaso, Esposito, Condit, & Moussatché, 2000; Rimoin et al., 2010). To this end, one valuable method to

understand the mechanism of disease is to determine the virus-host interactions necessary for orthopoxvirus infection.

Orthopoxviruses are large double-stranded DNA viruses with a unique lifecycle in the cytoplasm of the host cell. The viruses enact a cascade of transcriptional responses, with early gene expression occurring from the stages of viral entry to uncoating, intermediate gene expression after DNA replication, followed by late gene expression until the end of the virus lifecycle (Yang et al., 2011; Yang, Bruno, Martens, Porcella, & Moss, 2010). Early in infection, orthopoxviruses express factors that cleave host mRNAs, effectively preventing the expression of most host genes (Guerra et al., 2003; Yang et al., 2010). Poxviruses are also known to use host proteins during their lifecycle. This includes the use of the proteasome to facilitate viral uncoating and DNA replication, the ribosome to translate mRNAs, and specific host factors to help drive late viral transcription events (Becker & Joklik, 1964; Mercer et al., 2012; Satheshkumar, Anton, Sanz, & Moss, 2009; Teale et al., 2009).

Several RNAi screens have been performed in recent years and have expanded our knowledge of the host proteins involved in orthopoxvirus replication. Moser et al. performed a screen of kinase genes in Drosophila cells and found that modulation of the actin cytoskeleton by AMPK is important for VACV entry (Moser, Jones, Thompson, Coyne, & Cherry, 2010). Mercer et al. screened the 7,000 genes comprising the "druggable genome" and revealed the role of the proteasome in viral uncoating and of the Cullin3 ubiquitin ligase in initiating viral DNA replication (Mercer et al., 2012). Finally, Sivan et al. performed two siRNA screens targeting over 18,000 genes to reveal the

importance of nuclear pore genes in viral morphogenesis (Sivan et al., 2013). All of these important new insights were based on an arrayed RNAi screen format. Notably, these screens generated hit lists with some overlap on the protein or functional level, but also significant numbers of unique hits. This is presumably due to substantial false negative rates, false positive rates, and the distinct model systems and readouts used to assess VACV infection, suggesting that more host protein factors remain to be discovered.

Here, we used two complimentary and unbiased assays to identify host proteins necessary for orthopoxvirus infection. First, we developed a pooled-cell lentiviral shRNA screen in human cells based on screening formats previously utilized to determine pathways important in cancer biology (Cheung et al., 2011; Macosko et al., 2015). Strengths of the pooled screen format are the ease of scaling to larger screening sets and the ability to enable multiple screening paradigms. Furthermore, cells are cultured in standard low-throughput format, rather than in multiwell plates, and thus can be easily passaged and otherwise manipulated. As a second assay, we used RNASeq to analyze the host transcriptional responses elicited by poxviral infection. These data identified host mRNAs that were upregulated during infection, suggesting that they may facilitate virus infection.

By comparing the data from these two orthogonal datasets, we identified proteins involved in the heat shock response as critical factors for orthopoxvirus infection. In particular, we found that heat shock factor 1 (HSF1), the ancient master regulator of the cytoprotective heat shock response, is necessary for orthopoxvirus infection. We find that depletion and knockout of HSF1 or its pharmacologic inhibition significantly reduces

VACV infection. Moreover, the principal targets of HSF1 transcription are upregulated during VACV infection, even as global host gene expression is suppressed. Our findings define a set of host factors that are necessary for orthopoxvirus infection and suggest that poxviruses have evolved to utilize host stress responses to their own advantage.

5.3 Materials and Methods

5.3.1 Cell Culture and Viruses

A549 cells (CCL-85), HFF-1 (SCRC-1041) and HeLa (CCL-2) cells were obtained from the ATCC. The VACV used in this study was strain Western Reserve or a derivative thereof (K. Dower et al., 2012; Ken Dower, Rubins, Hensley, & Connor, 2011). MPXV experiments were completed with modified MPXV Zaire 1979 at USAMRIID under appropriate containment conditions (Johnston et al., 2012).

5.3.2 Pooled shRNA Screen

Screen. A549 cells were infected in 4 replicates with the lentiviral 90,000 shRNA library from the Broad Institute (http://www.broadinstitute.org/rnai/public/resources/screening) (Whittaker et al., 2013). Cells were counted and resuspended at 1.44E8 cells/replicate in media with 4 mg/ml polybrene solution. 1 ml of the cell/ polybrene mixture was put into each well of a 12 well plate (12 plates total). 50 ml lentivirus library was added to each well and spun at 2000 rpm (9306g) for 1 hour at room temperature. This concentration gave an infection rate of 28%. Plates were incubated at 37uC overnight. The next day, cells were

trypsinized and each replicate was pooled into T225 flask. Cells were allowed to sit at RT for 30 min, then incubated at 37uC for 1 hour before adding puromycin for selection. Cells were selected for 5 days; on day four pools were counted and split into 3 flasks for 1E8 cells/replicate (12 flasks total). On day 5, cells were infected with VACV-A4L at an MOI 5 for 12 hours. Cells were fixed with 4% formaldehyde, washed and resuspended in FACS buffer (PBS, 1% BSA, 0.05% sodium azide). Fixed cells were sorted on a MoFlo2 (Beckman Coulter) cell sorter; Venus-negative cells were collected with gates set on a control uninfected cell population. Collected cells were processed for sequencing as follows.

Illumina deep sequencing method. The shRNA region from the integrated lentiviral genome was PCR amplified from the purified cellular genomic DNA using the following conditions: 5 uL primary PCR primer mix, 4 mL dNTP mix, 16 Ex Taq buffer, 0.75 mL of Ex TaqDNA polymerase (Takara), and up to 3 mg genomic DNA in a total reaction volume of 50 mL. Up to 5 primary reactions were carried out in parallel for each sample. Thermal cycler PCR conditions consisted of heating samples to 95uC for 5 min; 15 cycles of 94uC for 30 sec, 65uC for 30 sec, and 72uC for 20 sec; and 72uC for 5 min. PCR reactions were then pooled per sample. A secondary PCR step was performed containing 5 mM of common barcoded 39 primer, 8 mL dNTP mix, 16 Ex Taq buffer, 1.5 mL Ex Taq DNA polymerase, and 30 mL of the primary PCR mix for a total volume of 90 mL. 10 mL of independent 59 barcoded primers were then added into each reaction, after which the 100 mL total volume was divided into two 50 mL final reactions. Thermal cycler conditions for secondary PCR were as follows: 95uC for 5 min; 15 cycles of 94uC

for 30 sec, 58uC for 30 sec, and 72uC for 20 sec; and 72uC for 5 min. Individual 50 mL reactions were then re-pooled. Reactions were run on a 2% agarose gel and intensity-normalized. Equal amounts of samples were mixed and gel-purified using a 2% agarose gel. Samples were sequenced using a custom sequencing primer using standard Illumina conditions.

5.3.3 Arrayed shRNA Secondary Screen

A549 cells were seeded in 96-well plates at low density the previous day. The lentivirus vectors were added to each well to achieve an MOI ,1. Infection was allowed to proceed overnight, then puromycin selection was applied for 5 days. The knockdown cells were infected with VACV-LREV at MOI 1 or 0.01. Cells were fixed with 4% formaldehyde 16–19 hpi, then read on a Tecan infinite M1000 for Venus (excitation: 515 nm and emission: 528 nm) and mCherry (excitation: 587 nm and emission: 610 nm). The secondary screen was completed with 3 independent experiments. Each plate was background corrected by subtracting the average of empty wells and normalized to 100% by the total RFU across the plate for early and late. An example plate had an average early Venus signal of 20,000 RLU, with an average background around 700 RLU and a reading of the GFP shRNA of 4,000 RLU. The average late mCherry signal was 3,200 RLU, with an average background signal around 75 RLU and an example positive hit of 1300 RLU. The hits were determined by comparing the normalized data across all three replicates for hairpins that decreased fluorescence more than 50%.

5.3.4 Cell Viability Assay

A549 cells were seeded in 96-well plates the previous day and infected with the arrayed lentiviral vectors at MOI ,1. After 5 days of puromycin selection, the cells were lysed and luciferase read according to manufacturer's instructions using CellTiter-Glo Luminescent Cell Viability Assay system (Promega) on a LUMIstar Omega luminometer (BMG Labtech) for 1 second/well.

5.3.5 Immunoblot

A549 cells were infected at an MOI specified in text. At times indicated, cells were lysed in RIPA buffer with protease and phosphatase inhibitors (1 mM PMSF, 1 mM benzamidine, 100 nM okadaic acid, 100 nM microcysin and 100 nM sodium fluoride). 20 mg of total lysate were separated on a 4–15% SDS-PAGE gel and transferred to PVDF (Bio-Rad 162-0177). Blots were probed with polyclonal antibodies specific to HSP27 (abcam [G3.1] antibody ab2790), Virostat anti-VACV virion (Virostat 8101), VACV I3L (mAb 10D2, generous gift of Dr. David Evans, University of Alberta, Edmonton), HSF1, including anti-HSF1 (phospho S326) (HSF1-PS326; abcam [EP1713Y] ab76076), anti-HSF1-J7F9 (abcam [J7F9] ab115303), anti-HSF1 (phospho S303) (abcam ab47369), and anti-HSF1 #4356 (Cell Signaling).

5.3.6 Vaccinia Infections with Inhibitors

A549 or HeLa cells were seeded in 96-well plates the previous day and infected with modified VACV viruses as specified in the text at an MOI of 1 (fluorescence readout) or 0.1 (viral titer). For drug treatment, compounds were added at specified

concentrations prior to virus addition. Inhibitor compounds: Triptolide, Quercetin (Tocris Bioscience), KRIBB11 (EMD Millipore), KRIBB3, Pifithrin-m, Myricetin (Sigma Aldrich), and Heat Shock Protein Inhibitor I/KNK437 (Santa Cruz Biotechnology, Inc.) For fluorophore assays, cells were fixed at 18 hpi with 4% formaldehyde. Plates were read on a Tecan infinite M1000 for Venus (excitation: 515 nm and emission: 528 nm) and mCherry (excitation: 587 nm and emission: 610 nm). For plaque assays, virus was collected 24 hpi and titered by plaque assay.

5.3.7 Plaque Assay

Virus was collected at specified timepoints post infection. Virus was freeze/thawed and sonicated 36. Viruses were then serially diluted in 10 fold dilutions and added to confluent BSC-40 cells. 24–48 hours post infection, cells were fixed and stained with crystal violet to visualize plaques.

5.3.8 Immunofluorescence

HFF-1 or A549 cells were seeded on coverslips the previous day and infected with VACV at MOI 1, mock infected or heat shocked for 2 hours (HFF-1) or 1 hour (A549) at 42uC. VACV and mock cells were fixed at 5 hpi (HFF-1) or 24 hpi (A549) with 4% formaldehyde. Heat shocked cells were recovered at 37uC for 30 minutes (HFF-1) or 1 hour (A549), then fixed. Cells were stained for HSF1 with HSF1-phospho-S326 or HSF1-phospho-S303 and DAPI to delineate nuclei. Coverslips were mounted using ProLong Gold antifade reagent with DAPI (Invitrogen) and imaged on Axiovert 200M microscope (Zeiss).

5.3.9 Host Transcriptome RNASeq

HeLa cells were grown to 90–100% confluency in 6-well plates, then inoculated with VACV-WR (MOI 10) in DMEM+2% FBS, and incubated at 37uC for 1 hour. After the 1 hour incubation, virus was removed and cells were washed three times with PBS before fresh media (with 2% FBS) was added (0 hpi time point) and returned to 37uC. At each time point (0, 0.5, 2, 6, 18 hpi), cells from two wells were harvested for nucleic acid extraction. RNA isolation: Total RNA was extracted from cells at each time point with Trizol, following the manufacturer's instructions. cDNA Library preparation: We used the Illumina mRNA seq V2 protocol (April 2008) to generate the cDNA library for next-generation sequencing. In short, mRNA was purified from the total RNA samples and fragmented before proceeding with reverse transcription. Adapters were ligated to the resulting cDNA fragments. Templates were size-selected through gel purification, then enriched by PCR, and the resulting libraries were validated on an Agilent Bioanalyzer DNA 1000 chip.

5.3.10 RNASeq Analysis

Sequencing. Sequencing five samples of HeLa cells infected with VACV using the Illumina platform generated an average of 16 million single-end 36-base pair reads. The sequencing data is available at the Sequence Read Archive under accession number SRP026257 (http://trace.ncbi.nlm.nih.gov/Traces/sra/sra. cgi?study = SRP026257) Yang et al had previously sequenced an additional five samples (WTA-A experiment) using the SOLiD platform and generated an average of 41 million single-end 50-base pair reads. In both datasets, the reads were aligned to the human genome (version hg19) using TopHat

(Kim et al., 2013), as specified in https://github.com/nachocab/vaccinia_filone_2013/blob/master/sequencing.sh.

Differential expression. We calculated the total number of reads that aligned to each protein-coding gene specified in GENCODE v14 (Harrow et al., 2012), normalized it using the Trimmed Mean M-values method (Robinson et al., 2010), and considered each count to represent the amount of expression for a specific gene at a given timepoint. We compared the counts at 6 hours post-infection with the pre-infection counts and ranked the genes in two dimensions: by fold change and by average abundance (Figure 5.6). This helped us establish two reasonable cutoffs to determine strongly upregulated genes (fold-change greater than 2, and normalized counts at 6 hpi greater than 20). Additionally, we selected upregulated genes that were also previously identified as HSF1 target genes by (Mendillo et al., 2012).

The code is available at https://github.com/nachocab/vaccinia_filone_2013/.

5.3.11 Monkeypox Infection

A549 cells were seeded in 96 well plates the previous day. Cells were infected with shRNA lentiviruses at MOI ,1 overnight, then selected with puromycin for 3 days. Cells were infected with modified MPXV (Johnston et al., 2012) at MOI 1 for 48 hours. Cells were fixed with 10% neutral buffered formalin, then read on a SpectraMax M5.

5.4 Results

5.4.1 Pooled RNAi Screen to Identify Host Factors Necessary for Orthopoxvirus Infection

To identify host factors necessary for orthopoxvirus infection we completed a whole-genome scale, pooled RNAi screen using lentiviral vectors. This method delivered ,90,000 short hairpin RNAs (shRNAs) with 5 or more independent shRNAs targeting ,17,000 human genes to our target cells (Figure 5.1) (Whittaker et al., 2013). Human A549 cells were transduced in four replicates with the shRNA lentivirus library at a multiplicity of infection (MOI) of #1 to generate cell populations with predominantly no more than one shRNA expressed in each cell. Non-transduced cells were eliminated following selection with puromycin. Each replicate was then infected at an MOI of 5 with a modified vaccinia virus (VACV) that expressed a fusion of the core protein A4L and Venus yellow fluorescent protein (K. Dower et al., 2012). This dose of VACV infected 100% of control cells at 12 hours (data not shown). At 12 hours post infection (hpi), cells were fixed and sorted for Venus-negative cells, with gates set on uninfected cells to collect Venus-negative pools. This population was selected to enrich for cells in which a host protein essential for orthopoxvirus replication, but not essential for host cell survival, had been suppressed.

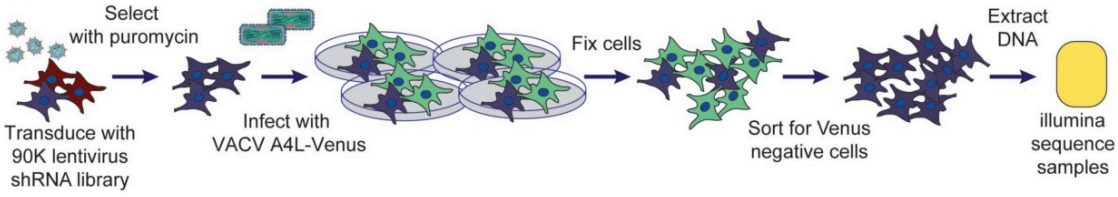


Figure 5.1 Pooled-cell shRNA screen revealed host factors necessary for orthopoxvirus infection.
Schematic of the primary pooled shRNA screen.

To determine which host genes were being suppressed in the Venus-negative cells, the cell population was analyzed to resolve hairpin sequences that were enriched in abundance, in essence using the hairpin sequence as a barcode to indicate shRNA treatment. The abundance of each hairpin was assessed by next generation sequencing (Ashton et al., 2012) using the Illumina GAIIx system. For each replicate, hairpins with fewer than 15 raw reads were not considered. The remaining hairpins were normalized to the total read depth for each individual replicate to eliminate variation in read depth across replicates. The fold-change enrichment of each hairpin within the Venus-negative sorted cells was determined by comparison to the initial abundance of each hairpin observed in the plasmid DNA pool used to generate the pooled lentivirus library. These fold changes were used to rank the enrichment of each hairpin in the Venus-negative cell population in each replicate. Using the RIGER (Luo et al., 2008) algorithm within the Gene-E software tool (http://www.broadinstitute.org/cancer/software/ GENE-E/), the weighted-second-best metric was used to rank the enriched target genes within each replicate. This method uses the pre-calculated ranked hairpin lists for each replicate, and then ranks the candidate genes based on the first and second most enriched hairpin for

each gene in each replicate. Therefore, at least two hairpins against each gene were enriched in the original screen, providing evidence for the specificity of the target gene in VACV infection. The target genes identified in the top 500 genes in each replicate were considered candidate hits (Table S1).

To better understand the candidate hits from our screen, we categorized the cellular pathways represented in our dataset for both functional pathways and biological process gene ontology (GO) terms. First, we used Ingenuity Pathway Analysis (IPA) to analyze the functional and signaling pathways associated with these host genes. We found that genes associated with molecular transport, ion transport and apoptosis were significantly overrepresented. The top categories overrepresented in our dataset correlate well with those identified in other screens for host genes important during vaccinia infection, with cell death and survival ($p = 1.27E-06$) and cell morphology ($p = 3.17E-05$) being significantly overrepresented in both screens analyzed using these parameters (Sivan et al., 2013). These data indicate that the candidate host factors necessary for orthopoxvirus infection are varied, and that several host biological processes act to promote orthopoxvirus infection.

From the initial list of candidate genes identified in our primary screen, a subset of 172 genes was selected for a secondary screen in arrayed format with a different VACV reporter virus system. In the secondary screen, 5–7 distinct shRNAs targeting each gene of interest were used to assess the effect of decreased host protein expression on VACV early and late gene expression. The shRNA lentiviral vectors were arrayed in a 96 well plate format, with a different shRNA in each well. A549 cells were transduced at

an MOI of ,1 with the lentivirus vectors expressing each shRNA, selected with puromycin for lentiviral integration, and then infected with a modified vaccinia virus that expressed soluble Venus under an early promoter and soluble mCherry under a late promoter (VACV-LREV) (K. Dower et al., 2012; Ken Dower et al., 2011). Cells were fixed 20 hours post VACV-LREV infection and fluorescence was measured from each well (Figure 5.2). The secondary screen was carried out with 3 independent biological replicates.

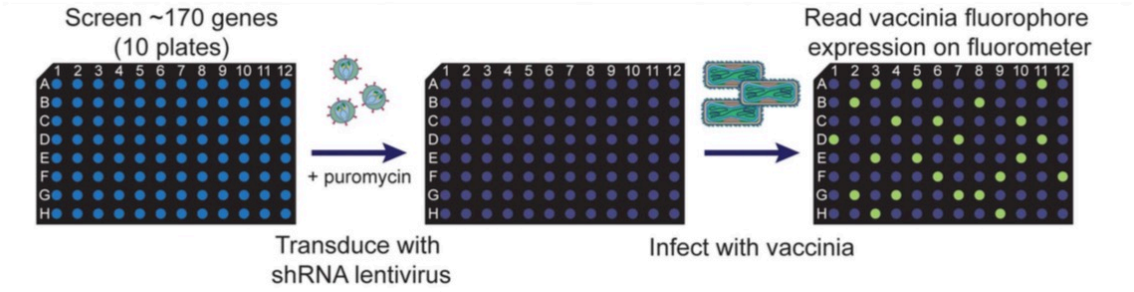


Figure 5.2 High confidence hits identified in secondary screen.
Schematic of arrayed shRNA lentivirus secondary screen.

Genes were considered hits if shRNA expression led to at least a 50% decrease in either early or late virus promoter-dependent fluorescence production (1) with more than one hairpin in a replicate or (2) in at least 2 replicates of the secondary screen without significant toxicity to the cells, as determined independently by cell viability assay (CellTiterGlo; data not shown). In most cases, knocking down host factors with shRNA blocked VACV-LREV late gene expression and not early expression, indicating that the host genes were not necessary for VACV entry and early gene expression. We compiled

a list of 34 genes that validated in the secondary screen (20% of those tested; Table S4). There were 7 genes that were positive hits in all replicates of the secondary screen: transcriptions factors HSF1 and SKI, the integrin binding protein ITGB1BP1, the aminophospholipid transporter ATP8B1, the Notch ligand JAG1, the nuclear transporter TNPO3 and the chemokine receptor CCR9. We considered these seven positives the highest-confidence hits emerging from the initial pooled RNAi screen.

5.4.2 Deep RNA Sequencing to Identify Host Cell Transcriptome during Orthopoxvirus Infection

Among the pooled RNAi screen hits, as well as previously published RNAi screen hits, were a large number of proteins that localize to the nucleus, including transcription factors, suggesting that VACV requires host systems that operate in the nucleus for its own replication (Mercer et al., 2012; Sivan et al., 2013). To investigate the effects of VACV infection on transcriptional responses, we analyzed host mRNA expression 6hours post VACV infection using RNASeq (Figure 5.3A). For each gene, we calculated the difference in normalized read counts (from the Illumina sequencing) between the pre-infection and the 6 hpi samples and compared it with the average number of Illumina read counts across these samples. Consistent with prior reports of a profound suppression of host mRNA following VACV infection, we also saw an overall decrease in the amount of host mRNA (Becker & Joklik, 1964; Guerra et al., 2003; Yang et al., 2010). After the normalization protocol, most genes showed a decrease or no difference in expression (genes in gray; Figure 5.3B). In contrast, 611 host genes were upregulated during VACV infection, as defined by at least a two-fold change in transcript abundance at 6 hours

(genes in black; Figure 5.3B). We consider these genes to be actively expressed during VACV infection to counteract the nonspecific decay of host mRNA during poxvirus infection (Guerra et al., 2003; Parrish & Moss, 2007; Parrish, Resch, & Moss, 2007; Yang et al., 2010).

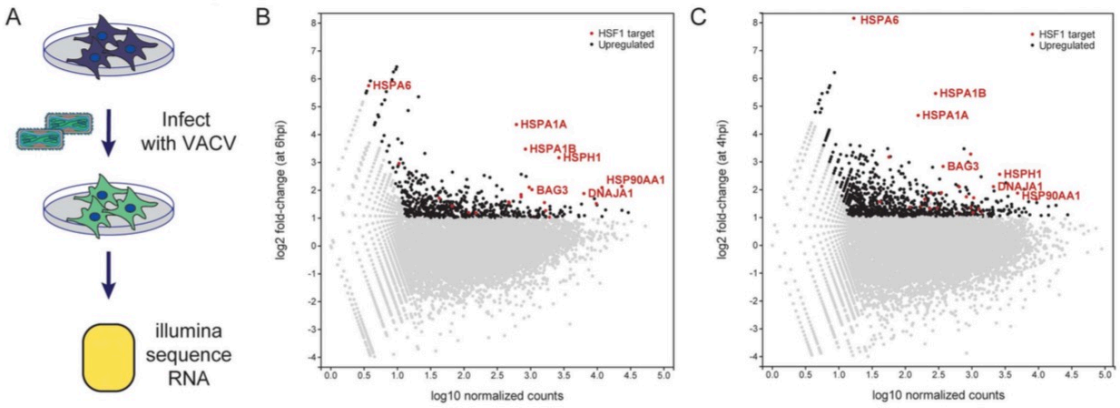


Figure 5.3 Host mRNA transcripts upregulated during VACV infection.
(A) Schematic of RNASeq Experiment. (B) Dots represent the change in expression of genes from 0 to 6 hpi (x-axis) and the average number of sequencing reads that align to each gene in both timepoints (y-axis). Genes with a fold change greater than 2 and more than 20 counts at 6 hpi are considered upregulated (black). HSF1-regulated genes upregulated during VACV infection are labeled in red (a subset is labeled with gene names). (C) Same analysis as in (B) using the WTA-A dataset at 4 hpi from (Yang et al., 2010).

We examined the upregulated genes for transcription factor targets (TFT) and GO biological process terms using the Molecular Signatures Database (MSigDB) (Subramanian, Tamayo, Mootha, Mukherjee, & Ebert, 2005). Strikingly, the set of 611 upregulated genes was very strongly enriched for genes regulated by HSF1 ($p = 3.39E-15$) and the stress response ($p = 1.44E-14$). Because HSF1 was also one of the high-confidence hits from the RNAi screen, we assessed the set of upregulated host genes

during VACV infection for enrichment of HSF1-responsive genes (using a specifically defined set of 61 genes that have at least a two-fold increase in expression and have HSF1 bound to their promoters in multiple cell lines following a 42uC heat shock) (Mendillo et al., 2012). Remarkably, there were 25 HSF1-regulated genes enriched at least two-fold at 6 hpi in our VACV dataset, which encompassed 41% of the HSF1-regulated gene list (genes in red, selected genes labeled; Figure 5.3B). HSF1-regulated genes highly expressed during VACV infection include 83% of the HSPs regulated by HSF1. This included HSPA6, which is not upregulated in cancer cells addicted to HSF1, but is strongly upregulated during a bona fide heat shock response (Mendillo et al., 2012). A number of HSF1-regulated HSP activators and cochaperones (AHSA1, BAG3, CHORDC1, STIP1) were also expressed.

To determine whether the increase in HSF1-regulated gene transcription was observed during VACV infection in other next generation sequencing datasets, we analyzed the data described by (Yang et al., 2010). In that study, HeLa cells were infected with a high MOI of VACV and total polyadenylated RNA was collected at 0 and 4 hpi (termed the Whole Transcriptome Analysis (WTA) dataset A). We analyzed the WTA-A dataset, and found 981 genes upregulated over two-fold at 4 hours post VACV infection (genes in black; Figure 5.3C), with an approximately 30% overlap with the genes upregulated at least two-fold in our dataset. Of the 61 HSF1-responsive genes we previously reported, 28 were upregulated (46%) in the Yang WTA-A dataset (genes in red with selected genes labeled; Figure 5.3C). This correlates well with our data at 6 hpi, with 19 of the 25 HSF1-regulated genes expressed in both datasets (Table S7). The

overlap has good representation of the HSP70/HSP110 superfamily and HSF1-regulated HSP cochaperones and activators. Together, these data indicate that HSF1-transcribed genes are upregulated during VACV infection. Previously published HSP data and a retrospective analysis of microarray experiments tracking host gene expression following poxvirus infection showed an association with the maintenance or upregulation of HSF1-regulated genes (Brum, Lopez, Varela, Baker, & Moyer, 2003; Guerra et al., 2003, 2006; Kowalczyk, Guzik, Slezak, Dziedzic, & Rokita, 2005; Sedger et al., 1996). These findings establish that the HSF1-regulated gene expression program is a dominant host transcriptional event stimulated by VACV infection.

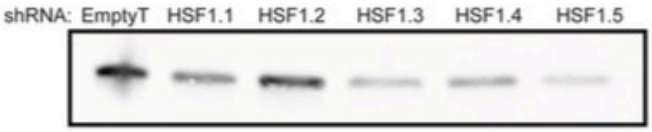
5.4.3 Orthopoxvirus Replication Requires Cell Stress Responses

The integrated analysis of both the pooled RNAi screen and the RNAseq host transcription data indicated that HSF1 was an important host factor. Therefore, we began to investigate the potential role of HSF1 in controlling orthopoxvirus infection. HSF1, the master transcriptional regulator of the heat shock response, controls the expression of most heat shock genes both under basal conditions and following proteotoxic cellular stress (Anckar & Sistonen, 2011; Mendillo et al., 2012; Morimoto, 1998; Page et al., 2006; Pirkkala, Nykänen, & Sistonen, 2001; Santoro, 2000; Trinklein, Murray, Hartman, Botstein, & Myers, 2004). The heat shock protein family is comprised of a large number of heat shock proteins (HSPs) with a broad range of chaperone functions. They are often designated by their molecular weight: HSPB (small HSPs), DNAJ (HSP40), HSPD, HSPA (HSP70), HSPC (HSP90) and HSPH, with most families containing multiple isoforms (Daugaard, Rohde, & Jäättelä, 2007; Easton, Kaneko, & Subjeck, 2000;

Kampinga et al., 2009; Shaner & Morano, 2007; Vos, Hageman, Carra, & Kampinga, 2008). Our RNAseq data supported the hypothesis that these genes were actively transcribed during VACV infection. Together with the RNAi data, our results suggested that HSF1 is critical for orthopoxvirus replication; thus, we investigated the role of HSF1 during orthopoxvirus infection more rigorously.

We used five shRNA lentiviral vectors to create five independent stable cell lines with depleted HSF1. The knockdown efficacy of the shRNAs targeting HSF1 varied, with 23–61% of HSF1 remaining after selection. A representative immunoblot is shown in Figure 5.4A; the percent of HSF1 remaining after shRNA knockdown was quantified using four distinct anti-HSF1 antibodies (Figure 5.4B). The stable knockdown cells were infected with a VACV expressing Venus under an early promoter and TagBFP under a late promoter (Figure 5.4C) (Ken Dower et al., 2011). Hairpins that reduced HSF1 levels inhibited VACV gene expression, significantly decreasing both early and late gene expression when compared to a control hairpin (p,0.002). To validate this finding, cell lines with HSF1 knocked down at least 50% were infected with VACV at MOI 0.01 to measure viral growth in the absence of HSF1. We observed a ,1 log10 decrease in viral titer (90% inhibition of viral growth) at 24 hpi in the knockdown cells compared to control shRNA cells (Figure 5.4D). These data strongly support the HSF1 target specificity of the phenotype, indicating shRNA knockdown of HSF1 is limiting viral gene expression and viral growth.

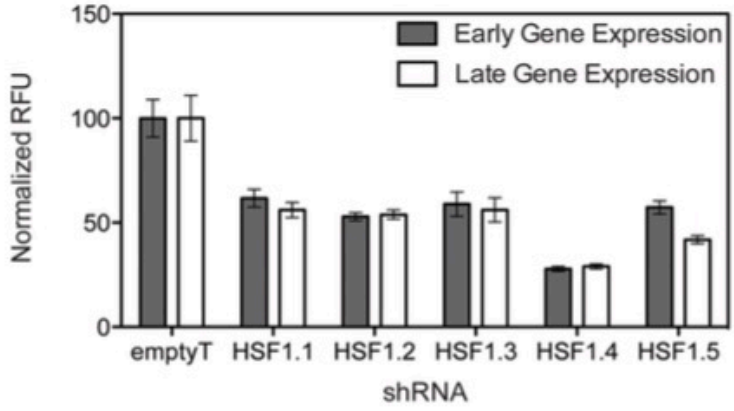
A



B

shRNA	Target	Sequence	% HSF1 Remaining
EmptyT	N/A	N/A	100
HSF1.1	CDS	GCACATTCCATGCCCAAGTAT	48 (4)
HSF1.2	CDS	CCAGCAACAGAAAGTCGTCAA	61 (8)
HSF1.3	CDS	GCCCAAGTACTTCAAGCACAA	24 (7)
HSF1.4	CDS	CAGTGACCACTTGGATGCTAT	41 (6)
HSF1.5	3UTR	GCAGGTTGTTCATAGTCAGAA	23 (6)

C



D

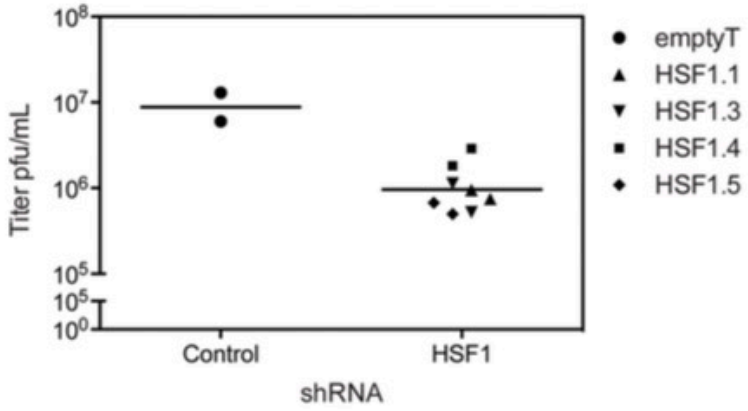


Figure 5.4 HSF1 knockdown inhibits VACV infection.
Five independent shRNAs targeting HSF1 were used to knockdown protein expression in A549 cells. The levels of HSF1 remaining after selection were quantified by immunoblot with several antibodies. (A) Representative immunoblot (anti-HSF1 J7F9). (B) The

average % HSF1 (6 standard error) remaining, quantified by immunoblot, is shown for 4 different anti-HSF1 antibodies (see Materials and Methods). (C) HSF1-knockdown cells infected with VACV-TrpV expressing Venus under an early promoter and TagBFP under a late promoter show a significant decrease in VACV infection, as measured by early and late gene expression. Three independent experiments were completed in triplicate; this is a representative plot showing relative fluorescent units (RFU) with standard error. (D) Plaque assay showing VACV titer reduced by ,1 log when HSF1 protein level is reduced over 50%.

We further confirmed the importance of HSF1 for VACV replication by analyzing virus infection in knockout mouse embryonic fibroblasts (MEFs) lacking HSF1 (Dai, Whitesell, Rogers, & Lindquist, 2007; McMillan, Xiao, Shao, Graves, & Benjamin, 1998). Infecting at MOI 0.01 with VACV-TrpV expressing Venus under an early promoter, mCherry under an intermediate promoter, and TagBFP under a late promoter, we observed that early, intermediate and late viral gene expression was inhibited in HSF1 null MEFs (Figure 5.5A). Images in Figure 5.5B show the expected cytopathic effects (CPE) induced by vaccinia virus infection in wild type Hsf1+/+ MEFs, but no CPE in the absence of HSF1 (Hsf12/2 MEFs). This effect was specific to HSF1, as knockout of HSF2 had equivalent levels of infection and corresponding CPE to WT counterparts (data not shown).

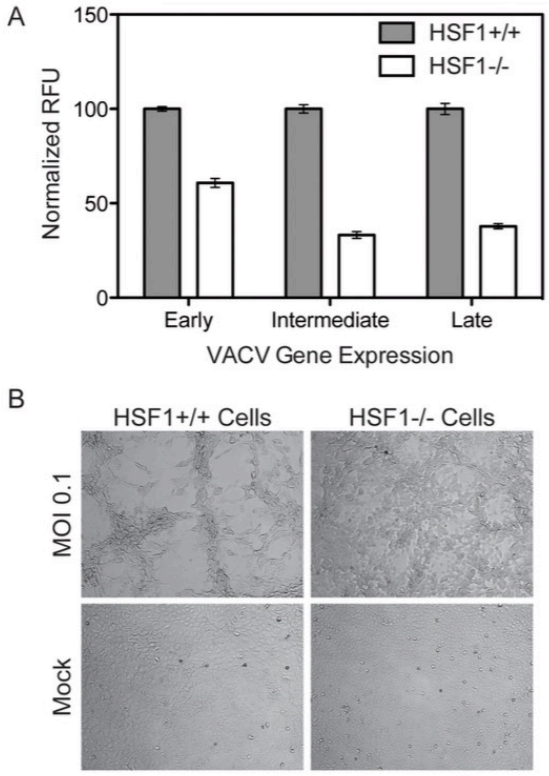


Figure 5.5 HSF1 null MEF cells support significantly less VACV infection. (A) HSF1 null MEF cells infected with VACV-TrpV show significantly less early (Venus), intermediate (mCherry) and late (TagBFP) gene expression compared to wild type MEFs. Three independent experiments were completed in triplicate; this is a representative plot showing normalized relative fluorescent units (RFU) with standard error. (B) Brightfield images show HSF1 null MEFs exhibit less cytopathic effects than wild type MEFs when infected with VACV at an MOI of 0.1 at 18 hpi. Mock infected HSF1 null and wild type MEFs are included for comparison.

In both the Hsf12/2 MEFs and the shRNA-knockdown cells, the depletion of HSF1 reduces VACV early, intermediate and late gene expression. This indicates that HSF1 is necessary for the entire VACV lifecycle, which is unexpected since orthopoxviruses package most of the viral factors necessary for early gene expression within the virion. Orthopoxviruses may need HSF1 directly or may activate its

transcriptional activity to enhance production of an HSF1-regulated target that is necessary for infection.

5.4.4 Orthopoxvirus Infection Activates HSF1

Our results demonstrating a role for HSF1 in vaccinia replication suggested that HSF1 was being activated following infection. In unstressed cells, HSF1 has been shown to exist as an inactive monomer in the cytoplasm, often in complexes with chaperone proteins. HSF1 undergoes an extensive set of posttranslational modifications, including phosphorylation, acetylation and sumoylation (Anckar & Sistonen, 2011). Upon activation, HSF1 is hyperphosphorylated and translocates to the nucleus to promote transcription of target genes (Guettouche, Boellmann, Lane, & Voellmy, 2005; Holmberg, Tran, Eriksson, & Sistonen, 2002; Jolly, Usson, & Morimoto, 1999; Santoro, 2000). We investigated whether HSF1 was activated during VACV infection in a manner similar to its activation by heat shock. When cells were exposed to elevated temperatures (42uC), an increase in the phosphorylated form of HSF1 is observed (Figure 5.6A, lane 1), when compared to the basal level of phosphorylated HSF1 in cells grown at 37uC (lane 2). Basal levels of HSF1 phosphorylation are seen in VACV-infected cells at 30 minutes post infection (lane 3) suggesting that there is no immediate change in HSF1 activation during virus entry. However, at later times in infection, levels of phosphorylated HSF1 strongly increased, similar to that seen following heat-shock (Figure 5.6A and Figure 5.6B). This demonstrated that VACV infection results in HSF1 phosphorylation, an established marker of HSF1 activation (Holmberg et al., 2002).

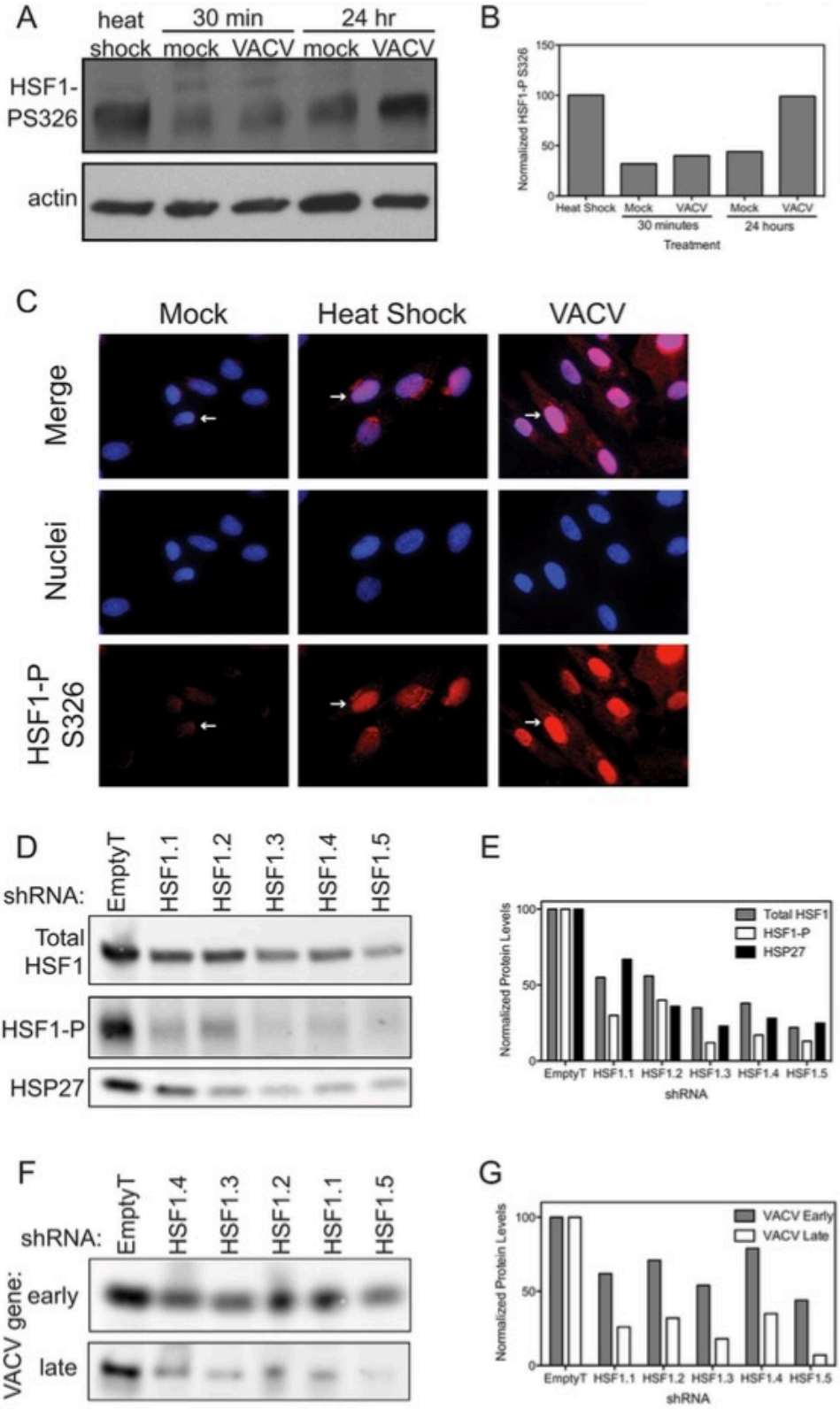

A
heat shock
30 min
24 hr
mock
VACV
mock
VACV
HSF1-PS326
actin
B
Normalized HSF1-P S326
150
100
50
0
Heat Shock
Mock
VACV
Mock
VACV
30 minutes
24 hours
Treatment
C
Mock
Heat Shock
VACV
Merge
Nuclei
HSF1-P S326
D
shRNA:
EmptyT
HSF1.1
HSF1.2
HSF1.3
HSF1.4
HSF1.5
Total HSF1
HSF1-P
HSP27
E
Normalized Protein Levels
100
50
0
Total HSF1
HSF1-P
HSP27
EmptyT
HSF1.1
HSF1.2
HSF1.3
HSF1.4
HSF1.5
shRNA
F
shRNA:
EmptyT
HSF1.4
HSF1.3
HSF1.2
HSF1.1
HSF1.5
VACV gene:
early
late
G
Normalized Protein Levels
100
50
0
VACV Early
VACV Late
EmptyT
HSF1.1
HSF1.2
HSF1.3
HSF1.4
HSF1.5
shRNA

Figure 5.6 HSF1 is activated during VACV infection.
(A) Immunoblot showing HSF1 is phosphorylated on S326 following heat shock (42uC; lane 1). Basal levels of HSF1 phosphorylation are seen in mock and VACV infected cells 30 minutes post infection (lanes 2 and 3). Basal levels of phosphorylation are seen in mock-infected cells at 24 hpi (lane 4), while HSF1 phosphorylation increases during VACV infection by 24 hpi (lane 5). The actin loading control is shown for all samples. (B) VACV induces HSF1 phosphorylation to similar levels as heat shock. Graph shows densitometry quantification of the phosphorylated HSF1 present in each sample normalized to the actin loading control with heat shock phosphorylation levels set at 100%. (C) Immunofluorescence images in HFF-1 cells of HSF1 phosphorylation on S326 (red, white arrows) show protein localization in mock cells (low levels, cytoplasm), heat shocked cells (increased phosphorylation, nucleus) or VACV-infected cells 5 hours post infection with MOI 1 (increased phosphorylation, nucleus). DAPI staining (blue) identifies nuclei. (D–G) A549 cells with HSF1 knocked down by five independent shRNA lentiviral vectors or a control vector were infected with VACV at MOI 0.1 for 18 hours. (D) Cell lysates were immunoblotted for host proteins. Total HSF1, phosphorylated on S326 (activated) HSF1 and HSP27 are shown. The cells transduced with HSF1 shRNA express lower levels of HSF1 and HSP27 and show reduced phosphorylation of HSF1 upon VACV infection. (E) Graph shows densitometry quantification of the bands in (D). (F) Cell lysates were immunoblotted for VACV-expressed proteins. Shown are I3L, an early protein, and a late protein recognized by a polyclonal antibody that recognizes late viral proteins. When HSF1 levels are reduced, there are lower levels of VACV proteins expressed. (G) Graph shows densitometry quantification of the bands in (F).

To determine if phosphorylated HSF1 is relocating to the nucleus upon infection, we undertook immunofluorescence analysis of HSF1. In cells grown at 37uC, the HSF1 antibody recognizing phosphorylation at S326 showed HSF1 located in the cytoplasm of primary human foreskin fibroblast (HFF-1) cells and A549 cells (white arrows, Figure 5.6C). During heat shock, phosphorylated HSF1 signal increases as HSF1 localizes to the nucleus (Figure 5.6C). Similarly, during VACV infection, phosphorylated HSF1 translocated to the nucleus, indicating that VACV is activating HSF1 in a manner similar to heat shock. A different HSF1 antibody, recognizing pS303, shows a distinct staining pattern, with HSF1 in the nucleus in cells grown at 37uC and the development of nuclear

stress granules as evidenced by bright foci in the nucleus, upon heat shock or VACV infection. These data demonstrate that phosphorylated HSF1 is in the nucleus during VACV infection, with staining patterns similar to heat shock, consistent with activation of this transcription factor.

In A549 cells with HSF1 depleted by shRNA knockdown, lower levels of total HSF1 correspond to a decrease in phosphorylated HSF1 during VACV infection (Figure 5.6D, quantitated in Figure 5.6E). The decrease in HSF1 activation following VACV infection corresponded with a decrease in expression of HSF1-transcribed genes, including HSP27 (Figure 5.6D and Figure 5.6E). The lack of HSF1 activation and downstream effectors led to a decrease in VACV gene expression as measured by fluorophores expressed from VACV promoters (Figure 5.4C), as well as expression of native VACV proteins measured by immunoblot. Here, the early viral protein I3 and a late protein recognized by a polyclonal antibody raised to virions, which are composed of predominantly late viral proteins, both show a decrease in protein levels when HSF1 is knocked down (Figure 5.6F, quantitated in Figure 5.6G). We see an inhibition of both early and late gene expression, with more inhibition of late gene expression than early gene expression. These data indicate that HSF1 activation, and perhaps transcription of downstream HSPs, is necessary for viral protein expression during VACV infection.

5.4.5 HSF1 Inhibitors Block Orthopoxvirus Infection

When HSF1 is depleted from the cell, the cellular milieu may be altered such that it is non-permissive for orthopoxvirus infection, or alternatively the virus may directly require active HSF1 transcription during infection. To differentiate between these

options, we pharmacologically inhibited HSF1 activity coincident with virus infection, for acute inhibition of HSF1 activity. We treated cells with several reported HSF1 inhibitors, including triptolide (Westerheide, Kawahara, Orton, & Morimoto, 2006), KNK437 (Ohnishi, Takahashi, Yokota, & Ohnishi, 2004; Voyer & Heikkila, 2008; Yokota, Kitahara, & Nagata, 2000), quercetin (Elia, Amici, Rossi, & Santoro, 1996; Hosokawa et al., 1992; Nagai, Nakai, & Nagata, 1995), and KRIBB11 (Yoon et al., 2011). The first three compounds do not bind HSF1 directly and likely influence HSF1 activity indirectly, along with the activity of other cellular systems (Santagata et al., 2013), while KRIBB11 has been reported to bind HSF1 directly (Yoon et al., 2011). One hour after drug treatment, the cells were infected with VACV expressing Venus from an early promoter and mCherry from a late promoter (VACV-LREV). All four drugs reduced viral gene expression from both early and late promoters in A549 cells (Figure 5.7A). More inhibition of late gene expression was observed compared to early gene expression; this may be due to the cascade transcription mechanism employed by poxviruses or HSF1 may be more important for late stages of infection than early.

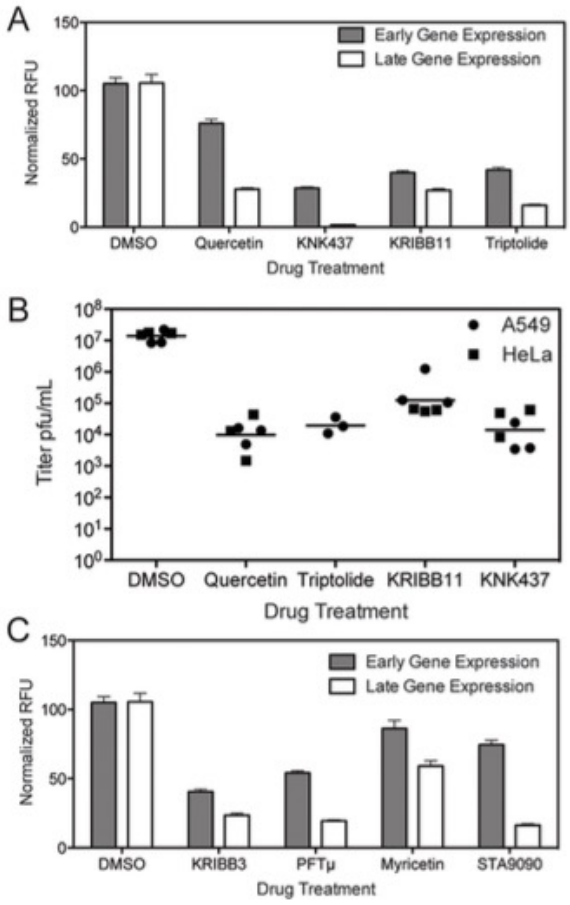


Figure 5.7 Inhibitors that target HSF1 or HSF1-transcribed genes block VACV infection.

(A) Inhibitors targeting HSF1 activity block VACV gene expression in A549 cells. Cells treated with 1 mM Triptolide, 50 mM Quercetin, 50 mM KNK437 or 50 mM KRIBB11 were infected with VACV-LREV at MOI 1 for 16 hours. Cells were fixed and fluorescence read to measure early and late gene expression. Three independent experiments were completed in triplicate; this is a representative plot showing RFU normalized to DMSO control at 100% with standard error. (B) Inhibitors targeting HSF1 activity block VACV infection in A549 and HeLa cells. Cells were treated with compounds at same concentrations as (A), and infected with VACV for 24 hours. Virus was collected and titered by plaque assay. HSF1 inhibitors block viral titers by 2–3 logs. (C) Inhibitors targeting HSF1-transcribed HSPs block VACV gene expression in A549 cells. Cells treated with 1 mM STA9090, 50 mM KRIBB3, 50 mM PFTm and 100 mM Myricetin were infected with VACV-LREV at MOI 1 for 16 hours. Cells were fixed and fluorescence read to measure early and late gene expression. Three independent experiments were completed in triplicate; this is a representative plot showing RFU normalized to DMSO control at 100% with standard error.

All four HSF1-inhibitory drugs also blocked virus replication as measured by viral titer. The compounds inhibited viral growth by 2 to 3 log10 in both A549 and HeLa cells (Figure 5.7B). Although the inhibitors each have off target effects, the drugs all function to block HSF1 with different mechanisms, strengthening the conclusion that active HSF1 transcription is necessary for orthopoxvirus replication, and that inhibition of HSF1 has antiviral effects.

We also tested pharmacologic inhibitors of some of the heat shock proteins transcriptionally controlled by HSF1 and expressed during VACV infection, including HSP90, HSP70 and HSP27. PFTm interacts with HSP70 and prevents its activity (Leu, Pimkina, Frank, Murphy, & George, 2009), KRIBB3 prevents the phosphorylation of HSP27 (Shin et al., 2005, 2008), Ganetespib (STA-9090) binds to the ATP-binding domain in the N-terminus of HSP90 (Lin et al., 2008; Y. Wang, Trepel, Neckers, & Giaccone, 2010), while myricetin may block the interaction between members of the HSP40 and HSP70 families (L. Chang et al., 2011). Acute inhibition of heat shock protein activity significantly decreased VACV-LREV infection as determined by fluorophore expression from early and late gene promoters (Figure 5.7C). Similar to the HSF1-inhibitory drugs, late gene expression was more inhibited than early gene expression. Together, these data suggest that not only is HSF1 important for VACV infection, but that several major HSF1-regulated targets are important as well.

5.4.6 Monkeypox Requires HSF1 for Infection

We next evaluated whether HSF1 could be a potential therapeutic target for other orthopoxviruses, in particular monkeypox, which currently leads to outbreaks in the human population (Rimoin et al., 2010). Knocking down HSF1 protein levels over 50% in A549 cells using four different shRNA sequences (Figure 5.4B) significantly inhibited MPXV early and late gene expression at 48 hpi. The expression of both eGFP, driven by an early MPXV promoter, and dsTomato Red, driven by a late MPXV promoter (Johnston et al., 2012), were significantly decreased when HSF1 levels were decreased (p,0.01; Figure 5.8A).

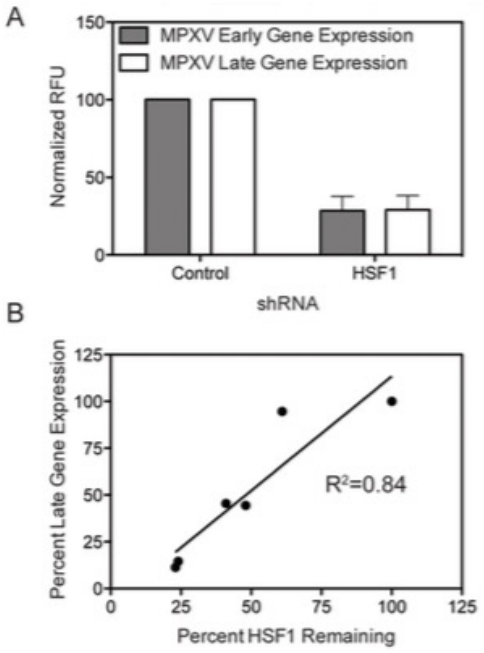


Figure 5.8 HSF1 is also a critical host factor for monkeypox infection.
(A) HSF1-knockdown cells infected with a modified MPXV expressing GFP under an early promoter and dsTomato Red under a late promoter show a decrease in MPXV infection, as measured by early and late gene expression normalized to MPXV infection in cells with no protein depletion. (B) The level of HSF1 protein remaining in the cell

following knockdown (x-axis) correlates with the inhibition of MPXV late gene expression during infection (y-axis).

These data strongly correlate with the efficacy of HSF1 knockdown for each shRNA, showing a clear relationship between the level of HSF1 present in the cell (Figure 5.4A and Figure 5.4B) and the ability of MPXV to express dsTomato Red from a late gene promoter (Figure 5.8B). For example, when HSF1 is knocked down with 48% protein remaining, MPXV late gene expression is 44.4%, while HSF1 knockdown with only 24% remaining results in 14.5% MPXV late gene expression. These data position HSF1 as a conserved host requirement of orthopoxvirus replication and as a potential pan-orthopoxvirus target for future therapeutic development.

5.5 Discussion

Here we identify HSF1, the master regulator of the host transcriptional response to proteotoxic cellular stress, as a critical host factor for orthopoxvirus infection. Identifying HSF1 as important resulted from the combined use of two unbiased experimental approaches: RNASeq and pooled shRNA screening. Pooled shRNA libraries have not yet been widely used to identify host factors important for virus replication, but may be a useful tool for probing the virus-host interaction on a genomic scale. While pooled screening is subject to the same false hit rates and cell toxicity issues as comparable arrayed format screens, our results show that both screening approaches identify components of similar cellular pathways.

Of our high-confidence hits, half (HSF1, JAG1, TNPO3, SKI) have a nuclear function or signal to transcription factors, which is interesting for a cytoplasmic virus. We also have a strong correlation with recently published host factors necessary for orthopoxvirus infection. Sivan et al recently described the importance of nuclear pore proteins in viral morphogenesis; we identified the nuclear import protein TNPO3 in our screen. JAG1, a Notch signaling molecule, is regulated by the Wnt pathway. The Wnt pathway was recently published to be important for Myxoma leporipoxvirus infection (Teferi, Dodd, Maranchuk, Favis, & Evans, 2013). We also identified ITGB1BP1, or ICAP1, which specifically binds to the cytoplasmic domain of beta1 integrin. Beta1 integrin was recently shown to interact with VACV on the cell surface and signal through PI3K/Akt to facilitate VACV entry (Izmailyan et al., 2012).

Previously published RNAi screens for necessary host factors during poxvirus infection of mammalian cells all also identified members of the heat shock response pathway, however the candidate genes were not validated or further developed in those publications (Mercer et al., 2012; Sivan et al., 2013; Teferi et al., 2013). Our screen was the only one to identify the master regulator of the pathway, HSF1, as important. The reason that our screen identified HSF1 and other screens did not is not immediately apparent, but it is notable that all of the screens enforce the idea that HSPs are important for viral replication.

The activation of HSF1 helps unify a mechanism for how poxviruses control the expression of many host proteins that they utilize. Earlier reports established that several heat shock proteins associate with VACV proteins during infection. HSP90 interacts with

VACV core protein 4a (A10L), and colocalizes with the viral factory during specific stages of the virus lifecycle (Hung, Chung, & Chang, 2002). Evidence suggests several viral proteins are bound by HSP70/72 (Jindal & Young, 1992). HSP27 (HSPB1) binds to three VACV proteins in protein-interaction studies: a truncated TNF-a receptor-like protein (VACV-WR002), C2L kelch-like protein (VACV-WR026) and I4L ribonucleoside-diphosphate reductase large subunit (VACV-WR073) (Van Vliet et al., 2009). Upregulation of HSF1 transcription provides a mechanism for how VACV and other poxviruses ensure sufficient levels of multiple chaperones through the activation of a single host protein.

While other studies have shown that poxviruses can activate host transcription processes (Mazzon et al., 2013), our study suggests that the activation of HSF1 aids orthopoxvirus replication on multiple levels. In addition to the involvement of HSPs as mentioned in the previous paragraph, HSC70, HSP72 and HSP90 have been shown to be packaged within virions (Chung et al., 2006; Manes et al., 2008; Resch, Hixson, Moore, Lipton, & Moss, 2007), suggesting an importance for chaperones in early stages of virus infection. The continuous usage of host cell chaperones at multiple (if not all) stages of the virus lifecycle shows that poxviruses use the heat shock response to extend their genome, activating HSF1 to transcribe essential factors that are encoded by the host.

A prediction of this genome extension hypothesis would be that some members of the family Poxviridae will have evolved to include one or more HSF1-stimulated genes in their own genome, reducing their dependence on host-production of HSPs. Consistent with this, genus Mulluscipoxvirus (Molluscum contagiosum; accession number

AAC55141) and genus Crocodylipoxvirus (Nile crocodilepox virus; accession number YP_784220) encode proteins with homology to the DNAJ/HSP40 chaperone family. Interestingly, other large, cytoplasmically replicating DNA viruses encode HSPs in their genome. This is most striking in the case of mimiviruses, which express HSP70 (MIMI_L254), HSP40/DNAJ (MIMI_R269, MIMI_gp0838) and DNAK (MIMI_L393) homologs, suggesting the requirement for large amounts of these HSPs during viral replication is consistent across cytoplasmic large DNA viruses (Legendre, Santini, Rico, Abergel, & Claverie, 2011; Raoult et al., 2004).

How orthopoxviruses activate HSF1 is an interesting question for future study. In unstressed cells, HSF1 is in an inactive form in the cytoplasm, bound to several proteins including HSP90 and HSP70. There are several proposed mechanisms of activation of HSF1, including the idea that recruitment of HSP90 and/or HSP70 away from HSF1 in the cytoplasm allows the free HSF1 to become post-translationally modified and translocate to the nucleus to begin transcribing genes (Guettouche et al., 2005; Holmberg et al., 2002; Jolly et al., 1999; Santoro, 2000). HSF1 may be activated during VACV infection when cytosolic HSP90 is recruited to the viral factory, as has been previously shown (Hung et al., 2002). Poxviruses also encode kinases, which may act to directly phosphorylate and activate HSF1 during VACV infection. Finally, orthopoxviruses may indirectly activate HSF1 by stimulating the MAPK signaling pathway (Andrade et al., 2004), which in turn strongly drives HSF1 activity (Dai et al., 2012). Understanding these mechanisms may provide insight into how an invading virion can manipulate the levels

of a selective set of host proteins while also deploying proteins that reduce the general level of host mRNAs.

Together, these data unify previously disparate observations regarding individually identified heat shock proteins and poxvirus infection. Earlier studies had illustrated the importance of individual members of the heat-shock response, but had not established whether poxviruses co-opted existing proteins or whether a heat-shock response was activated. Both our shRNA screening data and transcriptomic analysis implicate HSF1 activation as an important aspect of the orthopoxvirus lifecycle. This has implications not only for understanding viral evolution but also offers potential antiviral targets. Furthermore, our studies underscore that the HSF1 pathway is a viable target for broad-spectrum antiviral development, as it is a core cellular process used by multiple viruses, including HIV and EBV (Rawat & Mitra, 2011; F. W. Wang et al., 2011). A more complete understanding of how perturbing cellular homeostasis benefits viral replication will be important for illuminating the biology of virus-host interactions and for recognizing new therapeutic possibilities.

.

CHAPTER 6

VIRAL BIODIVERSITY OF RODENTS IN FRENCH GUIANA

6.1 Preamble

In this chapter, I describe the viral metagenomics work that I carried out in the Pasteur Institute in French Guiana, as part of a four-month internship funded by the International Division of the Pasteur Institute. My project was a pilot study to characterize the viral diversity of rodents living in three different regions of French Guiana. I was in charge of analyzing the sequencing data, comparing the performance of different alignment methodologies, and collaborating with experimental biologists to define best practice guidelines for future studies.

6.2 Introduction

Recent technological advances in the field of metagenomics have made it possible to collect a biological sample from a specific environment and characterize its viral composition. Unlike traditional metagenomic approaches—where ribosomal DNA is used to monitor bacterial diversity—, there is no universally conserved gene present across all viral families. Traditionally, viruses have been detected by looking for cytopathic effects on infected cell cultures, or by using neutralization tests (Niklasson, Jahrling, & Peters, 1984; Russell & Nisalak, 1967). Viral pathogens that are not culturable under these conditions, or for which no quality antiserum exists, have eluded identification efforts. Molecular methods like PCR have been used over the past few decades to overcome some of these hurdles, but these approaches require knowing the sequences of the target viruses before they can be identified. This limitation is starting to

disappear with the development of shotgun metagenomic sequencing (Bibby, 2013), which uses high-throughput sequencing to blindly characterize the viral diversity in an environmental sample.

The process starts by isolating viral particles, extracting genetic material from them, reverse transcribing RNA into cDNA, and fragmenting these nucleic acids until they become small enough to be sequenced. Sequencing generates millions of corresponding reads, which represent the genetic sequences of different organisms present in the sample: the host and all the bacteria, viruses, and eukaryotic parasites that inhabit it. The reads can be optionally assembled into longer contiguous sequences known as contigs, which are then compared with the genomes of previously sequenced species using a combination of local and global alignment. A given viral species is assumed to be present in a sample if enough associated reads or assembled contigs align with high similarity to its reference genome.

Here, I describe two approaches, raw read-based and contig-based, to characterize the viral diversity of serum and spleen samples from mice (Mus musculus) captured in three different regions of French Guiana, corresponding to habitats with different levels of human perturbation: urban, semiurban and unperturbed. The main findings are that host contamination is a critical issue for assessing viral diversity, and that both raw read-based and contig-based approaches are useful methods and provide complementary information.

6.3 Methods

6.3.1 Sample preparation

A serum and a spleen sample were taken from three different mice, one for each geographic location: Cayenne (urban), Macouria (semiurban), and Cavaliers Trois Palétuviers (unperturbed). ApoH-coated magnetic beads (ApoH Technologies, Villeneuve Saint Georges, France) were used to isolate intact viral particles. Beads were equilibrated in PBS and 10 ul of a 50% bead mixture in PBS was used for each 1.5 ml of serum or spleen homogenate. After a two-hour incubation period at 37ºC, beads were recovered using MagneSphere Technology Magnetic Separation Stands (Promega). After being washed three times with PBS, the beads were resuspended in RNA extraction buffer (Sigma-Aldrich) and RNA was extracted according to the manufacturer's protocol. The RNA was reverse-transcribed with the Transplex whole transcriptome amplification (WTA) kit (Sigma-Aldrich), using a quasi-random primer, according to the manufacurer's protocol. Sequencing was carried out on an Illumina Miseq using 250 bp paired-end reads.

6.3.2 Metagenomic Analysis

Trimmomatic (Bolger et al., 2014) was used to discard low-quality reads, trim low-quality bases and remove adapter sequences. The reference alignment and the *de novo* assembly methods detailed below used this subset of processed reads.

<u>Reference alignment using sequencing reads</u>

The blastn (Camacho et al., 2009) tool was used to align processed sequencing reads to the mouse genome. Reads that had alignments with an E-value < 10^{-3} were considered to originate from the host and were discarded from further analysis. The resulting set of reads was aligned using blastn, blastx and usearch (Edgar, 2010) against the viral sequences contained in the nucleotide and amino acid non-redundant databases from NCBI. Each high similarity match (E-value < 10^{-9}) was assigned a corresponding species, kingdom and taxonomic family by combining multiple files from the NCBI taxonomy database.

De novo assembly with and without digital normalization

Processed reads were assembled into contigs using SPAdes (Bankevich et al., 2012) with k-mer size k=41. This assembler performs a previous error-correction step using BayesHammer (Nikolenko, Korobeynikov, & Alekseyev, 2013). Additionally, processed reads were digitally normalized using khmer (Brown, Howe, Zhang, Pyrkosz, & Brom, 2012), a single-pass algorithm that normalizes coverage in shotgun sequencing datasets, which decreases sampling variation, and removes redundant data and the majority of sequencing errors.

6.4 Results

6.4.1 The majority of metagenomic sequences originate from the host

The six sequencing samples that were analyzed contained an average of 3.5 million paired reads. After using blastn to align the reads against the mouse genome,

between 28% to 68% of the reads showed high-score alignments with mouse genes, which likely indicates that these reads were the result of sequencing host nucleic acids, and provides evidence that the viral purification process was not virus-specific. Figure 6.1 shows the proportion of reads based on their host or non-host origin for each of the serum and blood samples in each type of environment.

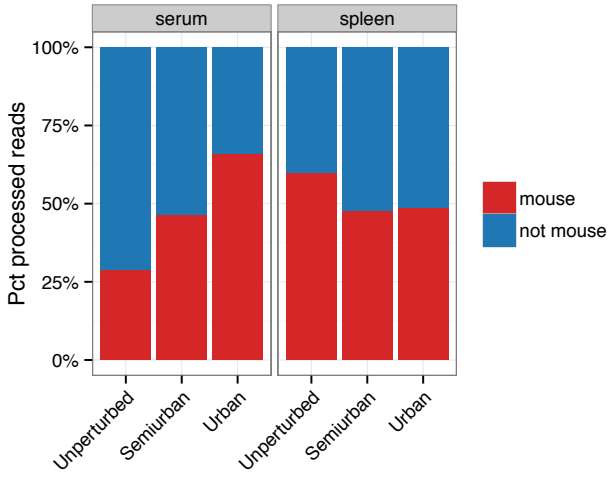


Figure 6.1: Sequencing reads of mouse origin for serum and spleen samples. Red bars indicate the percentage of processed sequencing reads aligning to the *Mus musculus* genome with a high score. Blue bars indicate the percentage of reads that don't align to the mouse genome or that do so with a low score.

6.4.2 Detecting viral sequences by directly aligning reads to a reference genome

After discarding reads that likely originated from the host, I assessed the diversity of viruses in the samples by directly aligning the remaining reads against a database of viral sequences using the usearch aligner. This resulted in multiple high-score alignments, but only five viral genomes accumulated more than 1000 aligned reads in at

least one of the samples (see Table 6.1): Torque teno felis virus (TTV), Lymphocytic choriomeningitis mammarenavirus (LCMV), Human Herpesvirus 6A (HHV-6A), Glypta fumiferanae ichnovirus (GfIV) and Choristoneura occidentalis granulovirus (ChocGV).

Virus	Semiurban Serum	Semiurban Spleen	Urban Serum	Urban Spleen	Unperturbed Serum	Unperturbed Spleen
TTV	0	0	6157	174	0	0
LCMV	0	0	1020	26	0	3
HHV-6A	59	180	71	1374	477	98
GfIV	142	533	186	3534	1262	254
ChocGV	4326	6225	124	432	1750	13342

Table 6.1 Number of reads aligning to viruses in each sample.

TTV is a single-stranded DNA virus from the family Anelloviridae. LCMV is a single-stranded RNA virus from the Arenaviridae family. Both viruses seem to be abundant only in mice captured in the urban environment, especially in the serum, and not in either of the other two environments. Figure 6.2 shows the number of reads that align to the genomes of these two viruses in each environment and tissue.

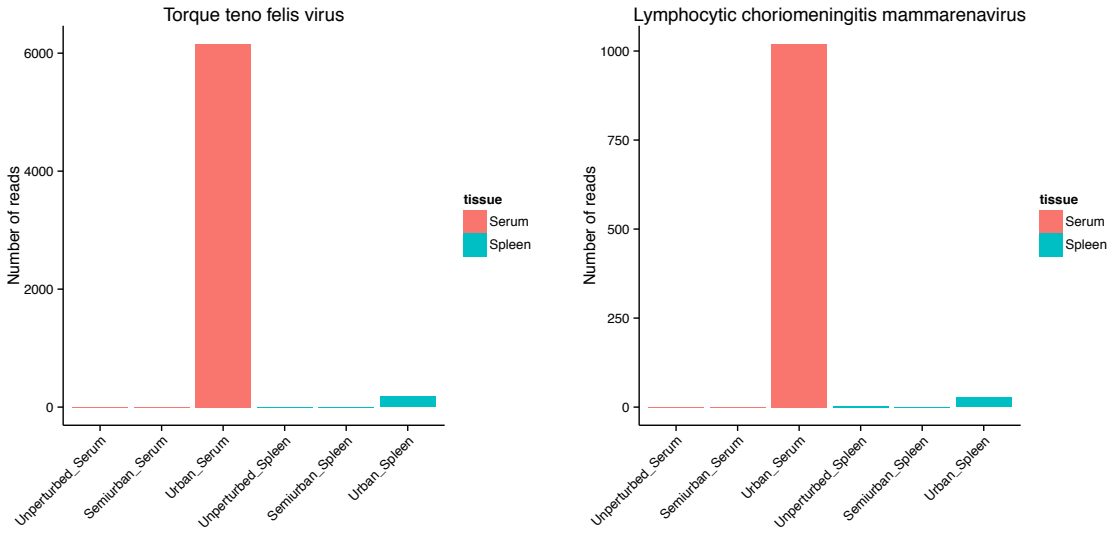


Figure 6.2: Read distribution of TTV and LCMV.
The y-axis shows the number of reads that resulted in a high-score alignment against the Torque teno felis virus genome (left) or the Lymphocytic choriomeningitis mammarenavirus (right). The x-axis indicates the environment and the tissue associated with each sample.

HHV-6A is a double-stranded DNA virus from the Herpesviridae family. GfIV is also a double-stranded DNA virus, but from the Polydnaviridae family. They both appear to be at least marginally present across most samples and environments. For serum, these viruses seem to be more abundant in samples from the unperturbed environment; while for spleen, they are more abundant in samples from the urban environment. Figure 6.3 shows the number of reads that align to the genomes of these two viruses in each environment and tissue.

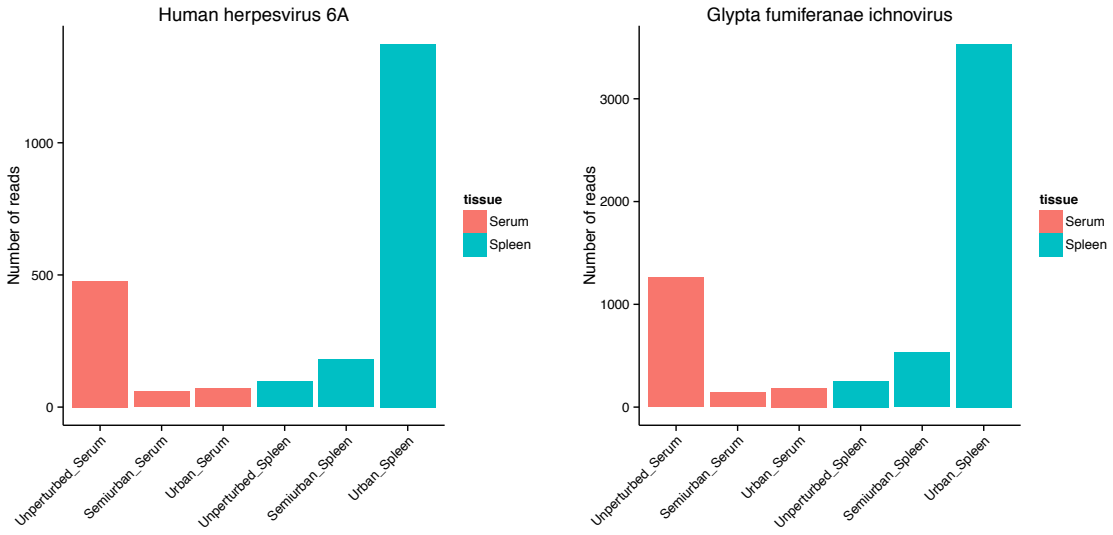


Figure 6.3: Read distribution of HHV-6A and GfIV.
The y-axis shows the number of reads that resulted in a high-score alignment against the Human herpesvirus 6A genome (left) or the Glypta fumiferanae ichnovirus (right). The x-axis indicates the environment and the tissue associated with each sample.

ChocGV is a double-stranded DNA virus from the Baculoviridae family. It appears to be present in unperturbed and semiurban areas, and seems to replicate more easily in the spleen than in serum. Figure 6.4 shows the number of reads that align to this genome in each environment and tissue.

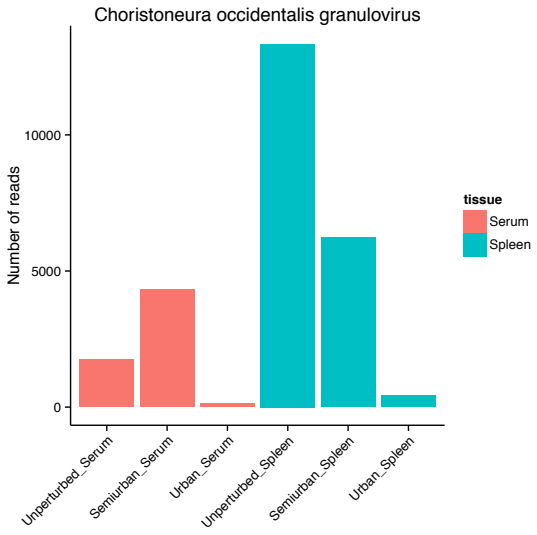


Figure 6.4: Read distribution of ChocGV.
The y-axis shows the number of reads that resulted in a high-score alignment against the Choristoneura occidentalis granulovirus genome. The x-axis indicates the environment and the tissue associated with each sample.

6.4.3 De novo assembly and digital normalization

Previously, I aligned individual reads against a reference genome to determine the viral diversity of a sample. An alternative method is *de novo* assembly, which combines reads into longer contiguous sequences (contigs) and aligns them against the viral genome database. Some of the advantages of this approach include a reduction in the number of sequences to be compared (lowering computing time), more reliable and less ambiguous alignments. Many assemblers are currently being used to process metagenomic sequences (Nagarajan & Pop, 2013). In this chapter I have chosen to compare Velvet (Zerbino & Birney, 2008), one of the most widely adopted *de novo* assemblers, with SPAdes, a relative newcomer in the metagenomic field. In addition to

the previously described raw dataset, I generated a digitally normalized read subset using khmer. As can be seen in Figure 6.5A, the digitally normalized dataset contains 70-77% fewer sequencing reads than the raw dataset. This provides evidence that digital normalization is able to provide optimal performance with a drastically reduced dataset at lower computational cost.

Using Velvet and SPAdes to generate assemblies using the raw and the digitally normalized datasets resulted in strikingly different results. Figure 6.5B shows several metrics that evaluate the performance of each assembler in each dataset. The first observation is that assemblies produced with SPAdes have a larger number of contigs, and of greater average and maximum size than those produced using Velvet. The second observation is that even though there are 70% fewer reads in the digitally normalized dataset, the resulting assemblies have almost identical metrics to those generated using the raw dataset.

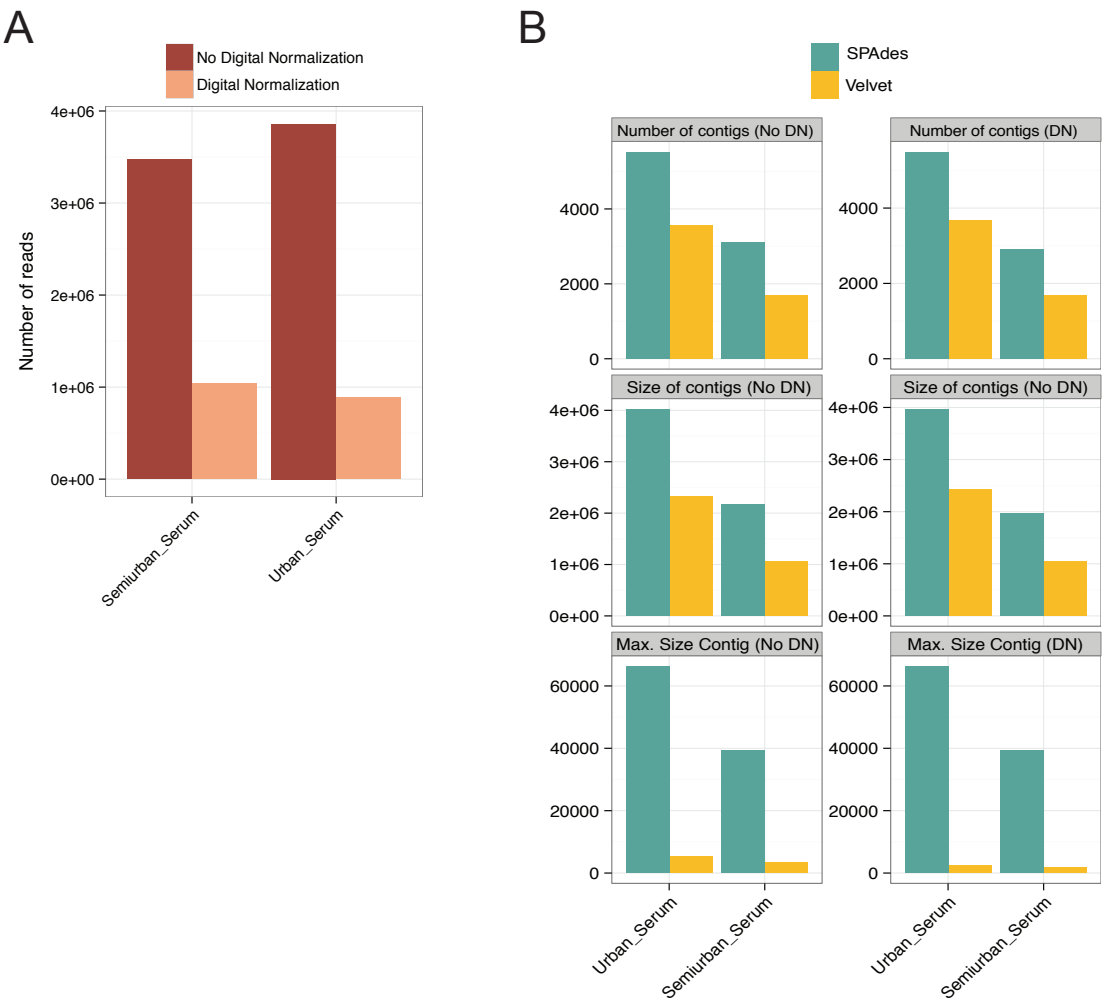


Figure 6.5: De novo assembly metrics.
(A) Number of reads generated from two different serum samples (semiurban and urban). The dark red bars represent the number of reads generated by sequencing (the raw dataset). The pink bars represent the number of reads that were considered non-redundant (the digitally normalized dataset). (B) Assembly metrics for the serum samples. The green bars correspond to the assemblies generated using SPAdes and the orange bars to those generated using Velvet. Boxes in the left correspond to the raw dataset, while those in the right correspond to the digitally normalized dataset.

6.4.4 Digital normalization can lead to more efficient assemblies

After seeing that the contigs generated by SPAdes were more numerous and longer on average than those generated by Velvet, I wanted to confirm the performance for an individual genome. Using BLASTX to align each of the four assemblies (SPAdes-DN, Spades-NoDN, Velvet-DN and Velvet-NoDN) from each sample against the NCBI viral database resulted in several high-score alignments, the longest of which came from the Urban-Serum sample, and matched the Lymphocytic choriomeningitis mammarenavirus. Figure 6.6 shows the size of the reference genome of the virus, along with the corresponding sizes of the longest contigs that aligned each segment in the four assemblies. The SPAdes-digitally normalized assembly contained the largest contigs (L segment: 6302 out of 6680 bp, S segment: 3324 out of 3376 bp), the contigs in the non-digitally normalized were 6% smaller, while the ones generated using Velvet were 96-97% smaller.

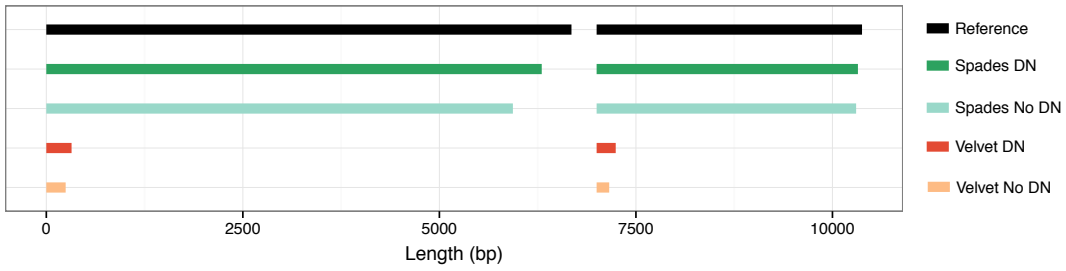


Figure 6.6: Contigs that aligned against the Lymphocytic choriomeningitis mammarenavirus genome in the Urban-Serum sample.
The x-axis shows the lengths (in base pairs) of the longest contig that aligned to the reference LCMV genome (L segment in the left, S segment in the right) in each of the four assemblies (Spades/Velvet and Normalized/Not normalized).

6.5 Discussion

The initial goal of the study was to identify pathogens present in mice that were living in regions of French Guiana that have experienced varying amounts of human perturbation (urban, semiurban and unperturbed). Unfortunately, multiple factors severely limited this goal. The number of samples and replicates from each region was not high enough to make any statistically robust predictions about region-specific viral biodiversity. Additionally, the experimental protocol that was used to prepare the samples for sequencing did not isolate viral particles with high efficiency, which resulted in the overwhelming majority of sequencing reads being generated from host RNA transcripts, with only a small percentage of sequences having a viral origin.

With the limited amount of data available, I was able to apply two different alignment strategies to obtain additional information about the most abundant viral sequences in the samples. Using reference-guided alignment tools, I identified the abundance patterns of five different viruses that had at least 1000 aligned reads in one of the samples. TTV and LCMV seemed to be most abundant in urban settings, while ChocGV was present in semiurban and unperturbed environments. The fact that multiple reads aligned to these genomes doesn't mean that their corresponding viruses are in fact present in the host, but rather that one of the viruses present in the sample shares its genomic sequence with a known, previously-sequenced virus. The actual host species-specific virus might not be identifiable until enough viral reads have accumulated to assemble the entire genome at high coverage.

Additionally, I applied *de novo* assembly methods to assemble reads into contiguous sequences, and assessed the effects of applying digital normalization to significantly reduce the number of reads needed to generate a high-quality assembly. Thanks to this technique, a reduction of approximately 70% in the number of reads used in the assembly generation step led to similar quality metrics than using all of the available reads. The advantage of using this approach is that it speeds up computation, lowers costs, and in some cases, increases performance. This is a direct consequence of the read-evaluation process of digital normalization, which discards reads that likely to contain low-frequency sequencing errors and results in higher-quality assemblies. In fact, the reason why the SPAdes assembler outperforms Velvet is likely to be the additional Bayesian error-correcting step that SPAdes incorporates in its pipeline (BayesHammer). This highlights the dependence between genome assembly and a low-error sequencing platform.

Viral metagenomics is a promising new field that can be a powerful technique to answer important scientific questions related to biodiversity, epidemics, and species discovery. Future studies would benefit from ensuring that host-contamination is minimized and that the experimental design is appropriate for the scientific goals of the project.

CHAPTER 7 BIBLIOGRAPHY

Alves, D. a, Glynn, a R., Steele, K. E., Lackemeyer, M. G., Garza, N. L., Buck, J. G., … Reed, D. S. (2010). Aerosol exposure to the angola strain of marburg virus causes lethal viral hemorrhagic Fever in cynomolgus macaques. *Veterinary Pathology*, *47*(5), 831–51. doi:10.1177/0300985810378597

Anckar, J., & Sistonen, L. (2011). Regulation of HSF1 function in the heat stress response: implications in aging and disease. *Annual Review of Biochemistry*, *80*, 1089–1115. doi:10.1146/annurev-biochem-060809-095203

Anders, S. (2010). HTSeq: Analysing high-throughput sequencing data with Python. *URL Http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html*.

Andrade, A. A., Silva, P. N. G., Pereira, A. C. T. C., De Sousa, L. P., Ferreira, P. C. P., Gazzinelli, R. T., … Bonjardim, C. A. (2004). The vaccinia virus-stimulated mitogen-activated protein kinase (MAPK) pathway is required for virus multiplication. *The Biochemical Journal*, *381*(Pt 2), 437–446. doi:10.1042/BJ20031375

Ashton, J. M., Balys, M., Neering, S. J., Hassane, D. C., Cowley, G., Root, D. E., … Jordan, C. T. (2012). Gene sets identified with oncogene cooperativity analysis regulate in vivo growth and survival of leukemia stem cells. *Cell Stem Cell*, *11*(3), 359–372. doi:10.1016/j.stem.2012.05.024

Baas, T., Baskin, C. R., Diamond, D. L., García-Sastre, a, Bielefeldt-Ohmann, H., Tumpey, T. M., … Katze, M. G. (2006). Integrated molecular signature of disease: analysis of influenza virus-infected macaques through functional genomics and proteomics. *Journal of Virology*, *80*(21), 10813–28. doi:10.1128/JVI.00851-06

Bahassi, E. M., Conn, C. W., Myer, D. L., Hennigan, R. F., McGowan, C. H., Sanchez, Y., & Stambrook, P. J. (2002). Mammalian Polo-like kinase 3 (Plk3) is a multifunctional protein involved in stress response pathways. *Oncogene*, *21*(43), 6633–40. doi:10.1038/sj.onc.1205850

Baize, S., Leroy, E. M., Georges, a J., Georges-Courbot, M.-C., Capron, M., Bedjabaga, I., … Mavoungou, E. (2002). Inflammatory responses in Ebola virus-infected patients. *Clinical and Experimental Immunology*, *128*(1), 163–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1906357&tool=pmcentrez&rendertype=abstract

Baize, S., Leroy, E. M., Georges-Courbot, M. C., Capron, M., Lansoud-Soukate, J., Debré, P., … Georges, a J. (1999). Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nature Medicine*, *5*(4), 423–6. doi:10.1038/7422

Baize, S., Leroy, E. M., Mavoungou, E., & Fisher-Hoch, S. P. (2000). Apoptosis in fatal Ebola infection. Does the virus toll the bell for immune system? *Apoptosis*, *5*(1), 5–7. doi:10.1023/A:1009657006550

Baize, S., Marianneau, P., Loth, P., Reynard, S., Journeaux, A., Chevallier, M., … Contamin, H. (2009). Early and strong immune responses are associated with control of viral replication and recovery in lassa virus-infected cynomolgus monkeys. *Journal of Virology*, *83*(11), 5890–5903. doi:10.1128/JVI.01948-08

Baize, S., Pannetier, D., Oestereich, L., Rieger, T., Koivogui, L., Magassouba, N., … Günther, S. (2014). Emergence of Zaire Ebola Virus Disease in Guinea - Preliminary Report. *The New England Journal of Medicine*, *371*, 418–1425. doi:10.1056/NEJMoa1404505

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. a, Dvorkin, M., Kulikov, A. S., … Pevzner, P. a. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology*, *19*(5), 455–77. doi:10.1089/cmb.2012.0021

Barrenas, F., Green, R. R., Thomas, M. J., Law, G. L., Proll, S. C., Engelmann, F., … Katze, M. G. (2015). Next generation sequencing reveals a controlled immune response to Zaire Ebola virus challenge in cynomolgus macaques immunized with VSVΔG/EBOVgp. *Clinical and Vaccine Immunology : CVI*, (January). doi:10.1128/CVI.00733-14

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., … Soboleva, A. (2013). NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Research*, *41*(D1), 991–995. doi:10.1093/nar/gks1193

Basler, C. F. (2015). Innate immune evasion by filoviruses. *Virology*, *480*, 122–130. doi:10.1016/j.virol.2015.03.030

Basler, C. F., & Amarasinghe, G. K. (2009). Evasion of interferon responses by Ebola and Marburg viruses. *Journal of Interferon & Cytokine Research : The Official Journal of the International Society for Interferon and Cytokine Research*, *29*(9), 511–20. doi:10.1089/jir.2009.0076

Basler, C. F., Wang, X., Mühlberger, E., Volchkov, V., Paragas, J., Klenk, H. D., … Palese, P. (2000). The Ebola virus VP35 protein functions as a type I IFN antagonist. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 12289–12294. doi:10.1073/pnas.220398297

Basler, C., & Mikulasova, A. (2003). The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. *Journal of …*, *77*(14), 7945–7956. doi:10.1128/JVI.77.14.7945

Bausch, D. G., Nichol, S. T., Muyembe-Tamfum, J. J., Borchert, M., Rollin, P. E., Sleurs, H., … Swanepoel, R. (2006). Marburg hemorrhagic fever associated with multiple genetic lineages of virus. *The New England Journal of Medicine*, *355*(9), 909–19. doi:10.1056/NEJMoa051465

Bausch, D. G., Rollin, P. E., Demby, A. H., Coulibaly, M., Kanu, J., Conteh, A. S., … Ksiazek, T. G. (2000). Diagnosis and Clinical Virology of Lassa Fever as Evaluated

by Enzyme-Linked Immunosorbent Assay, Indirect Fluorescent-Antibody Test, and Virus Isolation. *Journal of Clinical Microbiology*, *38*(7), 2670–2677. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=86994&tool=pmcentrez&rendertype=abstract

Becker, Y., & Joklik, W. K. (1964). Messenger Rna in Cells Infected With Vaccinia Virus. *Proceedings of the National Academy of Sciences of the United States of America*, *51*(1961), 577–585.

Bibby, K. (2013). Metagenomic identification of viral pathogens. *Trends in Biotechnology*, *31*(5), 275–279. doi:10.1016/j.tibtech.2013.01.016

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–20. doi:10.1093/bioinformatics/btu170

Bosio, C. M., Aman, M. J., Grogan, C., Hogan, R., Ruthel, G., Negley, D., … Schmaljohn, A. (2003). Ebola and Marburg viruses replicate in monocyte-derived dendritic cells without inducing the production of cytokines and full maturation. *The Journal of Infectious Diseases*, *188*(11), 1630–1638. doi:10.1086/379199

Bradfute, S. B., & Bavari, S. (2011). Correlates of immunity to filovirus infection. *Viruses*, *3*(7), 982–1000. doi:10.3390/v3070982

Bradfute, S. B., Swanson, P. E., Smith, M. a, Watanabe, E., McDunn, J. E., Hotchkiss, R. S., & Bavari, S. (2010). Mechanisms and consequences of ebolavirus-induced lymphocyte apoptosis. *Journal of Immunology (Baltimore, Md. : 1950)*, *184*(1), 327–35. doi:10.4049/jimmunol.0901231

Brainard, D. M., Tager, A. M., Misdraji, J., Frahm, N., Lichterfeld, M., Draenert, R., … Luster, A. D. (2007). Decreased CXCR3+ CD8 T cells in advanced human immunodeficiency virus infection suggest that a homing defect contributes to cytotoxic T-lymphocyte dysfunction. *Journal of Virology*, *81*(16), 8439–8450. doi:10.1128/JVI.00199-07

Brauburger, K., Hume, A. J., Mühlberger, E., & Olejnik, J. (2012). Forty-five years of marburg virus research. *Viruses*, *4*(10), 1878–1927. doi:10.3390/v4101878

Bray, M. (2001). The role of the Type I interferon response in the resistance of mice to filovirus infection. *The Journal of General Virology*, *82*(Pt 6), 1365–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11369881

Brown, C. T., Howe, A., Zhang, Q., Pyrkosz, A. B., & Brom, T. H. (2012). A Reference-Free Algorithm for Computational Normalization of Shotgun Sequencing Data. *arXiv*, *1203.4802*(v2), 1–18. doi:10.1128/genomeA.00802-14.Copyright

Brum, L. M., Lopez, M. C., Varela, J. C., Baker, H. V., & Moyer, R. W. (2003). Microarray analysis of A549 cells infected with rabbitpox virus (RPV): A

comparison of wild-type RPV and RPV deleted for the host range gene, SPI-1. *Virology*, *315*(2), 322–334. doi:10.1016/S0042-6822(03)00532-4

Buckley, S. M., & Casals, J. (1970). Lassa fever, a new virus disease of man from West Africa. 3. Isolation and characterization of the virus. *The American Journal of Tropical Medicine and Hygiene*, *19*(4), 680–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4987547

Bwaka, M. a, Bonnet, M. J., Calain, P., Colebunders, R., De Roo, A., Guimard, Y., … Van den Enden, E. (1999). Ebola hemorrhagic fever in Kikwit, Democratic Republic of the Congo: clinical observations in 103 patients. *The Journal of Infectious Diseases*, *179 Suppl*(Suppl 1), S1–S7. doi:10.1086/514308

Caballero, I. S., Bonilla, G., Yen, J. Y., & Connor, J. H. (2012). Diagnosing Lassa virus infection by tracking the antiviral response. *BMC Bioinformatics*, *13*(Suppl 18), A13.

Caballero, I. S., Yen, J. Y., Hensley, L. E., Honko, A. N., Goff, A. J., & Connor, J. H. (2014). Lassa and Marburg viruses elicit distinct host transcriptional responses early after infection. *BMC Genomics*, *15*(1), 960. doi:10.1186/1471-2164-15-960

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421. doi:10.1186/1471-2105-10-421

Cao, W., Henry, M. D., Borrow, P., Yamada, H., Elder, J. H., Ravkov, E. V, … Oldstone, M. B. a. (1998). Identification of a-Dystroglycan as a receptor for lymphocytic choriomeningitis virus and lassa fever virus. *Science*, *282*(5396), 2079–2081. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9851928

Cárdenas, W. B., Loo, Y.-M., Gale, M., Hartman, A. L., Kimberlin, C. R., Martínez-Sobrido, L., … Basler, C. F. (2006). Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. *Journal of Virology*, *80*(11), 5168–5178. doi:10.1128/JVI.02199-05

Centers for Disease Control and Prevention. (2015a). 2014 Ebola Outbreak in West Africa. Retrieved July 15, 2015, from http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/case-counts.html

Centers for Disease Control and Prevention. (2015b). Chronology of Marburg Hemorrhagic Fever Outbreaks. Retrieved from http://www.cdc.gov/vhf/marburg/resources/outbreak-table.html

Chang, L., Miyata, Y., Ung, P. M. U., Bertelsen, E. B., McQuade, T. J., Carlson, H. A., … Gestwicki, J. E. (2011). Chemical screens against a reconstituted multiprotein complex: Myricetin blocks DnaJ regulation of DnaK through an allosteric mechanism. *Chemistry and Biology*, *18*(2), 210–221. doi:10.1016/j.chembiol.2010.12.010

Chang, T. H., Kubota, T., Matsuoka, M., Jones, S., Bradfute, S. B., Bray, M., & Ozato, K. (2009). Ebola Zaire virus blocks type I interferon production by exploiting the host SUMO modification machinery. *PLoS Pathogens*, *5*(6). doi:10.1371/journal.ppat.1000493

Chen, M., Carlson, D., Zaas, A., Woods, C. W., Ginsburg, G. S., Hero, A., … Carin, L. (2011). Detection of viruses via statistical gene expression analysis. *IEEE Transactions on Bio-Medical Engineering*, *58*(3), 468–79. doi:10.1109/TBME.2010.2059702

Cheung, H. W., Cowley, G. S., Weir, B. A., Boehm, J. S., Rusin, S., Scott, J. A., … Hahn, W. C. (2011). Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(30), 12372–12377. doi:10.1073/pnas.1109363108

Chung, C.-S., Chen, C.-H., Ho, M.-Y., Huang, C.-Y., Liao, C.-L., & Chang, W. (2006). Vaccinia virus proteome: identification of proteins in vaccinia virus intracellular mature virion particles. *Journal of Virology*, *80*(5), 2127–2140. doi:10.1128/JVI.80.5.2127-2140.2006

Cilloniz, C., Ebihara, H., Ni, C., Neumann, G., Korth, M. J., Kelly, S. M., … Katze, M. G. (2011). Functional genomics reveals the induction of inflammatory response and metalloproteinase gene expression during lethal Ebola virus infection. *Journal of Virology*, *85*(17), 9060–8. doi:10.1128/JVI.00659-11

Crocker, P. R., Paulson, J. C., & Varki, A. (2007). Siglecs and their roles in the immune system. *Nature Reviews. Immunology*, *7*(4), 255–66. doi:10.1038/nri2056

Cummins, D. (1990). Acute Sensorineural Deafness in Lassa Fever. *JAMA: The Journal of the American Medical Association*, *264*(16), 2093. doi:10.1001/jama.1990.03450160063030

Cummins, D., Fisher-Hoch, S. P., Walshe, K. J., Mackie, I. J., McCormick, J. B., Bennett, D., … Machin, S. J. (1989). A plasma inhibitor of platelet aggregation in patients with Lassa fever. *British Journal of Haematology*, *72*(4), 543–548. doi:10.1111/j.1365-2141.1989.tb04321.x

Cunningham, F., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., … Flicek, P. (2014). Ensembl 2015. *Nucleic Acids Research*, *43*(D1), D662–D669. doi:10.1093/nar/gku1010

Dai, C., Santagata, S., Tang, Z., Shi, J., Cao, J., Kwon, H., … Lindquist, S. (2012). Loss of tumor suppressor NF1 activates HSF1 to promote carcinogenesis. *Journal of Clinical Investigation*, *122*(10), 3742–3754. doi:10.1172/JCI62727

Dai, C., Whitesell, L., Rogers, A. B., & Lindquist, S. (2007). Heat Shock Factor 1 Is a Powerful Multifaceted Modifier of Carcinogenesis. *Cell*, *130*(6), 1005–1018. doi:10.1016/j.cell.2007.07.020

Damaso, C. R., Esposito, J. J., Condit, R. C., & Moussatché, N. (2000). An emergent poxvirus from humans and cattle in Rio de Janeiro State: Cantagalo virus may derive from Brazilian smallpox vaccine. *Virology*, *277*(2), 439–449. doi:10.1006/viro.2000.0603

Daugaard, M., Rohde, M., & Jäättelä, M. (2007). The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. *FEBS Letters*. doi:10.1016/j.febslet.2007.05.039

Djavani, M. M., Crasta, O. R., Zapata, J. C., Fei, Z., Folkerts, O., Sobral, B., … Salvato, M. S. (2007). Early blood profiles of virus infection in a monkey model for Lassa fever. *Journal of Virology*, *81*(15), 7960–73. doi:10.1128/JVI.00536-07

Dower, K., Filone, C. M., Hodges, E. N., Bjornson, Z. B., Rubins, K. H., Brown, L. E., … Connor, J. H. (2012). Identification of a Pyridopyrimidinone Inhibitor of Orthopoxviruses from a Diversity-Oriented Synthesis Library. *Journal of Virology*. doi:10.1128/JVI.05416-11

Dower, K., Rubins, K. H., Hensley, L. E., & Connor, J. H. (2011). Development of Vaccinia Reporter Viruses for Rapid, High Content Analysis of Viral Function at All Stages of Gene Expression. *Antiviral Research*, *91*(1), 72–80. doi:10.1016/j.biotechadv.2011.08.021.Secreted

Drosten, C., Kümmerer, B. M., Schmitz, H., & Günther, S. (2003). Molecular diagnostics of viral hemorrhagic fevers. *Antiviral Research*, *57*(1-2), 61–87. doi:10.1016/S0166-3542(02)00201-2

Dufour, J. H., Dziejman, M., Liu, M. T., Leung, J. H., Lane, T. E., & Luster, A. D. (2002). IFN-gamma-inducible protein 10 (IP-10; CXCL10)-deficient mice reveal a role for IP-10 in effector T cell generation and trafficking. *Journal of Immunology (Baltimore, Md. : 1950)*, *168*(7), 3195–204. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11907072

Easton, D. P., Kaneko, Y., & Subjeck, J. R. (2000). The hsp110 and Grp1 70 stress proteins: newly recognized relatives of the Hsp70s. *Cell Stress & Chaperones*, *5*(4), 276–290. doi:10.1379/1466-1268(2000)005<0276:THAGSP>2.0.CO;2

Ebihara, H., Rockx, B., Marzi, A., Feldmann, F., Haddock, E., Brining, D., … Feldmann, H. (2011). Host response dynamics following lethal infection of rhesus macaques with Zaire ebolavirus. *The Journal of Infectious Diseases*, *204 Suppl*(Suppl 3), S991–9. doi:10.1093/infdis/jir336

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. doi:10.1093/bioinformatics/btq461

Elia, G., Amici, C., Rossi, A., & Santoro, M. G. (1996). Modulation of prostaglandin A1-induced thermotolerance by quercetin in human leukemic cells: role of heat shock protein 70. *Cancer Research*, *56*(1), 210–217.

Fadel, S. a., Bromley, S. K., Medoff, B. D., & Luster, A. D. (2008). CXCR3-deficiency protects influenza-infected CCR5-deficient mice from mortality. *European Journal of Immunology*, *38*(12), 3376–3387. doi:10.1002/eji.200838628

Feldmann, H., & Geisbert, T. W. (2011). Ebola haemorrhagic fever. *Lancet*, *377*(9768), 849–62. doi:10.1016/S0140-6736(10)60667-8

Feldmann, H., Jones, S. M., Daddario-DiCaprio, K. M., Geisbert, J. B., Ströher, U., Grolla, A., … Geisbert, T. W. (2007). Effective post-exposure treatment of ebola infection. *PLoS Pathogens*, *3*(1), 0054–0061. doi:10.1371/journal.ppat.0030002

Filone, C. M., Caballero, I. S., Dower, K., Mendillo, M. L., Cowley, G. S., Santagata, S., … Connor, J. (2014). The Master Regulator of the Cellular Stress Response (HSF1) Is Critical for Orthopoxvirus Infection. *PLoS Pathogens*, *10*(2), e1003904. doi:10.1371/journal.ppat.1003904

Fink, J., Gu, F., Ling, L., Tolfvenstam, T., Olfat, F., Chin, K. C., … Hibberd, M. L. (2007). Host gene expression profiling of dengue virus infection in cell lines and patients. *PLoS Neglected Tropical Diseases*, *1*(2). doi:10.1371/journal.pntd.0000086

Fisher-Hoch, S. P., McCormick, J. B., Sasso, D., & Craven, R. B. (1988). Hematologic dysfunction in Lassa fever. *Journal of Medical Virology*, *26*(2), 127–135. doi:10.1002/jmv.1890260204

Fisher-Hoch, S. P., Platt, G. S., Neild, G. H., Southee, T., Baskerville, a, Raymond, R. T., … Simpson, D. I. (1985). Pathophysiology of shock and hemorrhage in a fulminating viral infection (Ebola). *The Journal of Infectious Diseases*, *152*(5), 887–894.

Fisher-Hoch, S. P., Tomori, O., Nasidi, a, Perez-Oronoz, G. I., Fakile, Y., Hutwagner, L., & McCormick, J. B. (1995). Review of cases of nosocomial Lassa fever in Nigeria: the high price of poor medical practice. *BMJ (Clinical Research Ed.)*, *311*(7009), 857–859. doi:10.1136/bmj.311.7009.857

Fritz, E. a, Geisbert, J. B., Geisbert, T. W., Hensley, L. E., & Reed, D. S. (2008). Cellular immune response to Marburg virus infection in cynomolgus macaques. *Viral Immunology*, *21*(3), 355–63. doi:10.1089/vim.2008.0023

Geisbert, T. W., & Feldmann, H. (2011). Recombinant vesicular stomatitis virus-based vaccines against Ebola and Marburg virus infections. *The Journal of Infectious Diseases*, *204 Suppl* (Suppl 3), S1075–81. doi:10.1093/infdis/jir349

Geisbert, T. W., Hensley, L. E., Gibb, T. R., Steele, K. E., Jaax, N. K., & Jahrling, P. B. (2000). Apoptosis induced in vitro and in vivo during infection by Ebola and Marburg viruses. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *80*(2), 171–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10701687

Geisbert, T. W., Hensley, L. E., Larsen, T., Young, H. a, Reed, D. S., Geisbert, J. B., … Davis, K. J. (2003). Pathogenesis of Ebola hemorrhagic fever in cynomolgus

macaques: evidence that dendritic cells are early and sustained targets of infection. *The American Journal of Pathology*, *163*(6), 2347–70. doi:10.1016/S0002-9440(10)63591-2

Geisbert, T. W., Young, H. a, Jahrling, P. B., Davis, K. J., Kagan, E., & Hensley, L. E. (2003). Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. *The Journal of Infectious Diseases*, *188*(11), 1618–29. doi:10.1086/379724

Geisbert, T. W., Young, H. a, Jahrling, P. B., Davis, K. J., Larsen, T., Kagan, E., & Hensley, L. E. (2003). Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. *The American Journal of Pathology*, *163*(6), 2371–82. doi:10.1016/S0002-9440(10)63592-4

Gerszten, R. E., Garcia-Zepeda, E. a, Lim, Y. C., Yoshida, M., Ding, H. a, Gimbrone, M. a, … Rosenzweig, A. (1999). MCP-1 and IL-8 trigger firm adhesion of monocytes to vascular endothelium under flow conditions. *Nature*, *398*(April), 718–723. doi:10.1038/19546

Green, A. (2014). West Africa struggles to contain Ebola outbreak. *The Lancet*, *383*(9924), 1196. doi:10.1016/S0140-6736(14)60579-1

Grolla, a, Lucht, a, Dick, D., Strong, J. E., & Feldmann, H. (2005). Laboratory diagnosis of Ebola and Marburg hemorrhagic fever. *Bulletin de La Société de Pathologie Exotique (1990)*, *98*(3), 205–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16267962

Groom, J. R., & Luster, A. D. (2011). CXCR3 ligands: redundant, collaborative and antagonistic functions. *Immunology and Cell Biology*, *89*(2), 207–215. doi:10.1038/icb.2010.158

Guerra, S., López-Fernández, L. A., Pascual-Montano, A., Muñoz, M., Harshman, K., & Esteban, M. (2003). Cellular gene expression survey of vaccinia virus infection of human HeLa cells. *Journal of Virology*, *77*(11), 6493–6506. doi:10.1128/JVI.77.11.6493-6506.2003

Guerra, S., López-Fernández, L. A., Pascual-Montano, A., Nájera, J. L., Zaballos, A., & Esteban, M. (2006). Host response to the attenuated poxvirus vector NYVAC: upregulation of apoptotic genes and NF-kappaB-responsive genes in infected HeLa cells. *Journal of Virology*, *80*(2), 985–998. doi:10.1128/JVI.80.2.985-998.2006

Guettouche, T., Boellmann, F., Lane, W. S., & Voellmy, R. (2005). Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. *BMC Biochem*, *6*, 4. doi:10.1186/1471-2091-6-4

Günther, S., & Lenz, O. (2004). *Lassa virus. Critical reviews in clinical laboratory sciences* (Vol. 41). doi:10.1080/10408360490497456

Gupta, M., Mahanty, S., Ahmed, R., & Rollin, P. E. (2001). Monocyte-derived human macrophages and peripheral blood mononuclear cells infected with ebola virus secrete MIP-1alpha and TNF-alpha and inhibit poly-IC-induced IFN-alpha in vitro. *Virology*, *284*(1), 20–5. doi:10.1006/viro.2001.0836

Harcourt, B. H., Sanchez, A., & Margaret, K. (1999). Ebola Virus Selectively Inhibits Responses to Interferons , but Not to Interleukin-1 β , in Endothelial Cells Ebola Virus Selectively Inhibits Responses to Interferons , but Not to Interleukin-1 NL , in Endothelial Cells.

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., … Hubbard, T. J. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Research*, *22*(9), 1760–74. doi:10.1101/gr.135350.111

Hartman, A. L., Bird, B. H., Towner, J. S., Antoniadou, Z.-A., Zaki, S. R., & Nichol, S. T. (2008). Inhibition of IRF-3 activation by VP35 is critical for the high level of virulence of ebola virus. *Journal of Virology*, *82*(6), 2699–2704. doi:10.1128/JVI.02344-07

Hartman, A. L., Ling, L., Nichol, S. T., & Hibberd, M. L. (2008). Whole-genome expression profiling reveals that inhibition of host innate immune response pathways by Ebola virus can be reversed by a single amino acid change in the VP35 protein. *Journal of Virology*, *82*(11), 5348–58. doi:10.1128/JVI.00215-08

Hartman, A. L., Towner, J. S., & Nichol, S. T. (2004). A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza a virus. *Virology*, *328*(2), 177–184. doi:10.1016/j.virol.2004.07.006

Hayes, M., & Salvato, M. (2012). Arenavirus evasion of host anti-viral responses. *Viruses*, *4*(10), 2182–96. doi:10.3390/v4102182

Hensley, L. E., Alves, D. a, Geisbert, J. B., Fritz, E. a, Reed, C., Larsen, T., & Geisbert, T. W. (2011). Pathogenesis of Marburg hemorrhagic fever in cynomolgus macaques. *The Journal of Infectious Diseases*, *204 Suppl* (Suppl 3), S1021–31. doi:10.1093/infdis/jir339

Hensley, L. E., Smith, M. a, Geisbert, J. B., Fritz, E. a, Daddario-DiCaprio, K. M., Larsen, T., & Geisbert, T. W. (2011). Pathogenesis of Lassa fever in cynomolgus macaques. *Virology Journal*, *8*(1), 205. doi:10.1186/1743-422X-8-205

Hensley, L. E., Stevens, E. L., Yan, S. B., Geisbert, J. B., Macias, W. L., Larsen, T., … Geisbert, T. W. (2007). Recombinant human activated protein C for the postexposure treatment of Ebola hemorrhagic fever. *The Journal of Infectious Diseases*, *196 Suppl*(Suppl 2), S390–S399. doi:10.1086/520598

Hensley, L. E., Young, H. a, Jahrling, P. B., & Geisbert, T. W. (2002). Proinflammatory response during Ebola virus infection of primate models: possible involvement of

the tumor necrosis factor receptor superfamily. *Immunology Letters*, *80*(3), 169–79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11803049

Hink-Schauer, C., Estébanez-Perpiñá, E., Kurschus, F. C., Bode, W., & Jenne, D. E. (2003). Crystal structure of the apoptosis-inducing human granzyme A dimer. *Nature Structural Biology*, *10*(7), 535–40. doi:10.1038/nsb945

Holmberg, C. I., Tran, S. E. F., Eriksson, J. E., & Sistonen, L. (2002). Multisite phosphorylation provides sophisticated regulation of transcription factors. *Trends in Biochemical Sciences*. doi:10.1016/S0968-0004(02)02207-7

Hosokawa, N., Hirayoshi, K., Kudo, H., Takechi, H., Aoike, A., Kawai, K., & Nagata, K. (1992). Inhibition of the activation of heat shock factor in vivo and in vitro by flavonoids. *Molecular and Cellular Biology*, *12*(8), 3490–3498.

Howlin, J., Rosenkvist, J., & Andersson, T. (2008). TNK2 preserves epidermal growth factor receptor expression on the cell surface and enhances migration and invasion of human breast cancer cells. *Breast Cancer Research : BCR*, *10*(2), R36. doi:10.1186/bcr2087

Hsieh, M.-F., Lai, S.-L., Chen, J.-P., Sung, J.-M., Lin, Y.-L., Wu-Hsieh, B. a, … Liao, F. (2006). Both CXCR3 and CXCL10/IFN-inducible protein 10 are required for resistance to primary infection by dengue virus. *Journal of Immunology (Baltimore, Md. : 1950)*, *177*(3), 1855–1863. doi:177/3/1855 [pii]

Hu, X., Yu, J., Crosby, S. D., & Storch, G. a. (2013). Gene expression profiles in febrile children with defined viral and bacterial infection. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(31), 12792–7. doi:10.1073/pnas.1302968110

Huang, Y., Zaas, A. K., Rao, A., Dobigeon, N., Woolf, P. J., Veldman, T., … Hero, A. O. (2011). Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. *PLoS Genetics*, *7*(8), e1002234. doi:10.1371/journal.pgen.1002234

Hung, J.-J., Chung, C.-S., & Chang, W. (2002). Molecular chaperone Hsp90 is important for vaccinia virus growth in cells. *Journal of Virology*, *76*(3), 1379–1390. doi:10.1128/JVI.76.3.1379-1390.2002

Izmailyan, R., Hsao, J.-C., Chung, C.-S., Chen, C.-H., Hsu, P. W.-C., Liao, C.-L., & Chang, W. (2012). Integrin 1 Mediates Vaccinia Virus Entry through Activation of PI3K/Akt Signaling. *Journal of Virology*. doi:10.1128/JVI.06860-11

Izquierdo-Useros, N., Lorizate, M., Puertas, M. C., Rodriguez-Plata, M. T., Zangger, N., Erikson, E., … Martinez-Picado, J. (2012). Siglec-1 is a novel dendritic cell receptor that mediates HIV-1 trans-infection through recognition of viral membrane gangliosides. *PLoS Biology*, *10*(12), e1001448. doi:10.1371/journal.pbio.1001448

Jahrling, P. B., Frame, J. D., Rhoderick, J. B., & Monson, M. H. (1985). Endemic lassa fever in liberia. IV. Selection of optimally effective plasma for treatment by passive

immunization. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *79*(3), 380–384. doi:10.1016/0035-9203(85)90388-8

Jahrling, P. B., Frame, J. D., Smith, S. B., & Monson, M. H. (1985). Endemic Lassa fever in Liberia. III. Characterization of Lassa virus isolates. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *79*(3), 374–379. doi:10.1016/0035-9203(85)90386-4

Jahrling, P. B., Peters, C. J., & Stephen, E. L. (1984). Enhanced treatment of Lassa fever by immune plasma combined with ribavirin in cynomolgus monkeys. *The Journal of Infectious Diseases*, *149*(3), 420–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6715898

Jindal, S., & Young, R. A. (1992). Vaccinia virus infection induces a stress response that leads to association of Hsp70 with viral proteins. *Journal of Virology*, *66*(9), 5357–5362.

Johnson, K. M., McCormick, J. B., Webb, P. a, Smith, E. S., Elliott, L. H., & King, I. J. (1987). Clinical virology of Lassa fever in hospitalized patients. *The Journal of Infectious Diseases*, *155*(3), 456–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3805773

Johnston, S. C., Lin, K. L., Connor, J. H., Ruthel, G., Goff, A., & Hensley, L. E. (2012). In vitro inhibition of monkeypox virus production and spread by Interferon-β. *Virology Journal*, *9*(1), 5. doi:10.1186/1743-422X-9-5

Johnstone, R. W., Frew, A. J., & Smyth, M. J. (2008). The TRAIL apoptotic pathway in cancer onset, progression and therapy. *Nature Reviews. Cancer*, *8*(10), 782–798. doi:10.1038/nrc2465

Jolly, C., Usson, Y., & Morimoto, R. I. (1999). Rapid and reversible relocalization of heat shock factor 1 within seconds to nuclear stress granules. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(12), 6769–6774. doi:10.1073/pnas.96.12.6769

Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., … Hightower, L. E. (2009). Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress & Chaperones*, *14*(1), 105–11. doi:10.1007/s12192-008-0068-7

Kash, J., Mühlberger, E., & Carter, V. (2006). Global suppression of the host antiviral response by Ebola-and Marburgviruses: increased antagonism of the type I interferon response is associated with enhanced virulence. *Journal of Virology*. doi:10.1128/JVI.80.6.3009

Kawai, T., & Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunology*, *11*(5), 373–84. doi:10.1038/ni.1863

Ke, N., Claassen, G., Yu, D.-H., Albers, A., Fan, W., Tan, P., … Li, Q.-X. (2004). Nuclear hormone receptor NR4A2 is involved in cell transformation and apoptosis. *Cancer Research*, *64*(22), 8208–12. doi:10.1158/0008-5472.CAN-04-2134

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36. doi:10.1186/gb-2013-14-4-r36

Kodama, Y., Shumway, M., & Leinonen, R. (2012). The sequence read archive: Explosive growth of sequencing data. *Nucleic Acids Research*, *40*(D1), 2011–2013. doi:10.1093/nar/gkr854

Kortepeter, M. G., Bausch, D. G., & Bray, M. (2011). Basic clinical and laboratory features of filoviral hemorrhagic fever. *The Journal of Infectious Diseases*, *204 Suppl*(Suppl 3), S810–6. doi:10.1093/infdis/jir299

Kowalczyk, A., Guzik, K., Slezak, K., Dziedzic, J., & Rokita, H. (2005). Heat shock protein and heat shock factor 1 expression and localization in vaccinia virus infected human monocyte derived macrophages. *Journal of Inflammation (London, England)*, *2*, 12. doi:10.1186/1476-9255-2-12

Ksiazek, T. G., Rollin, P. E., Williams, a J., Bressler, D. S., Martin, M. L., Swanepoel, R., … Peters, C. J. (1999). Clinical virology of Ebola hemorrhagic fever (EHF): virus, virus antigen, and IgG and IgM antibody findings among EHF patients in Kikwit, Democratic Republic of the Congo, 1995. *The Journal of Infectious Diseases*, *179 Suppl* , S177–87. doi:10.1086/514321

Kuhn, J. H., Bao, Y., Bavari, S., Becker, S., Bradfute, S., Brister, J. R., … Nichol, S. T. (2012). Virus nomenclature below the species level: a standardized nomenclature for natural variants of viruses assigned to the family Filoviridae. *Archives of Virology*. doi:10.1007/s00705-012-1454-0

Kunz, S., Rojek, J. M., Kanagawa, M., Spiropoulou, C. F., Barresi, R., Campbell, K. P., & Oldstone, M. B. a. (2005). Posttranslational modification of alpha-dystroglycan, the cellular receptor for arenaviruses, by the glycosyltransferase LARGE is critical for virus binding. *Journal of Virology*, *79*(22), 14282–14296. doi:10.1128/JVI.79.22.14282-14296.2005

Lecompte, E., Fichet-Calvet, E., Daffis, S., Koulémou, K., Sylla, O., Kourouma, F., … Ter Meulen, J. (2006). Mastomys natalensis and Lassa fever, West Africa. *Emerging Infectious Diseases*, *12*(12), 1971–1974. doi:10.3201/eid1212.060812

Legendre, M., Santini, S., Rico, A., Abergel, C., & Claverie, J.-M. (2011). Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. *Virology Journal*, *8*, 99. doi:10.1186/1743-422X-8-99

Leroy, E., Baize, S., & Volchkov, V. (2000). Human asymptomatic Ebola infection and strong inflammatory response. *The Lancet*, *355*, 2210–2215. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673600024053

Leroy, E. M., Baize, S., Debre, P., Lansoud-Soukate, J., & Mavoungou, E. (2001). Early immune responses accompanying human asymptomatic Ebola infections. *Clinical and Experimental Immunology*, *124*(3), 453–60. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1906073&tool=pmcentrez&rendertype=abstract

Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., … Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, *438*(7068), 575–576. doi:10.1038/438575a

Leu, J. I.-J., Pimkina, J., Frank, A., Murphy, M. E., & George, D. L. (2009). A small molecule inhibitor of inducible heat shock protein 70. *Molecular Cell*, *36*(1), 15–27. doi:10.1016/j.molcel.2009.09.023

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–60. doi:10.1093/bioinformatics/btp324

Lin, T. Y., Bear, M., Du, Z., Foley, K. P., Ying, W., Barsoum, J., & London, C. (2008). The novel HSP90 inhibitor STA-9090 exhibits activity against Kit-dependent and -independent malignant mast cell tumors. *Experimental Hematology*, *36*(10), 1266–1277. doi:10.1016/j.exphem.2008.05.001

Lubaki, N. M., Ilinykh, P., Pietzsch, C., Tigabu, B., Freiberg, A. N., Koup, R. a, & Bukreyev, A. (2013). The lack of maturation of Ebola virus-infected dendritic cells results from the cooperative effect of at least two viral domains. *Journal of Virology*, *87*(13), 7471–85. doi:10.1128/JVI.03316-12

Luo, B., Cheung, H. W., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., … Root, D. E. (2008). Highly parallel identification of essential genes in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20380–20385. doi:10.1073/pnas.0810485105

Macher, A., & Wolfe, M. (2006). Historical Lassa Fever Reports and 30-year Clinical Update. *Emerging Infectious Diseases*, *12*(5), 835–837. Retrieved from http://europepmc.org/articles/PMC3374442

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., … McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, *161*(5), 1202–1214. doi:10.1016/j.cell.2015.05.002

Mahanty, S., Bausch, D. G., Thomas, R. L., Goba, a, Bah, a, Peters, C. J., & Rollin, P. E. (2001). Low levels of interleukin-8 and interferon-inducible protein-10 in serum are associated with fatal infections in acute Lassa fever. *The Journal of Infectious Diseases*, *183*(12), 1713–21. doi:10.1086/320722

Mahanty, S., & Bray, M. (2004). Reviews Pathogenesis of filoviral haemorrhagic fevers, *4*(August), 487–498.

Mahanty, S., Gupta, M., Paragas, J., Bray, M., Ahmed, R., & Rollin, P. E. (2003). Protection from lethal infection is determined by innate immune responses in a mouse model of Ebola virus infection. *Virology*, *312*(2), 415–424. doi:10.1016/S0042-6822(03)00233-2

Mahanty, S., Hutchinson, K., Agarwal, S., McRae, M., Rollin, P. E., & Pulendran, B. (2003). Cutting edge: impairment of dendritic cells and adaptive immunity by Ebola and Lassa viruses. *Journal of Immunology (Baltimore, Md. : 1950)*, *170*(6), 2797–801. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12626527

Malhotra, S., Yen, J. Y., Honko, A. N., Garamszegi, S., Caballero, I. S., Johnson, J. C., … Connor, J. H. (2013). Transcriptional Profiling of the Circulating Immune Response to Lassa Virus in an Aerosol Model of Exposure. *PLoS Neglected Tropical Diseases*, *7*(4), e2171. doi:10.1371/journal.pntd.0002171

Manes, N. P., Estep, R. D., Mottaz, H. M., Moore, R. J., Clauss, T. R. W., Monroe, M. E., … Smith, R. D. (2008). Comparative proteomics of human monkeypox and vaccinia intracellular mature and extracellular enveloped virions. *Journal of Proteome Research*, *7*(3), 960–968. doi:10.1021/pr070432+

Marzi, A., Engelmann, F., Feldmann, F., Haberthur, K., Shupert, W. L., Brining, D., … Messaoudi, I. (2013). Antibodies are necessary for rVSV/ZEBOV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 1893–8. doi:10.1073/pnas.1209591110

Mazzon, M., Peters, N. E., Loenarz, C., Krysztofinska, E. M., Ember, S. W. J., Ferguson, B. J., & Smith, G. L. (2013). A mechanism for induction of a hypoxic response by vaccinia virus. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(30), 12444–9. doi:10.1073/pnas.1302140110

McCormick, J. B., & Fisher-Hoch, S. P. (2002). Lassa fever. *Current Topics in Microbiology and Immunology*, *262*, 75–109.

McCormick, J. B., King, I. J., Webb, P. a, Johnson, K. M., O'Sullivan, R., Smith, E. S., … Tong, T. C. (1987). A case-control study of the clinical diagnosis and course of Lassa fever. *The Journal of Infectious Diseases*, *155*(3), 445–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3805772

McCormick, J., & Webb, P. (1987). A Prospective Study of the Epidemiology and Ecology of Lassa Fever. *Journal of Infectious …*, *155*(3). Retrieved from http://jid.oxfordjournals.org/content/155/3/437.short

McElroy, A., & Erickson, B. (2014). Ebola hemorrhagic fever: novel biomarker correlates of clinical outcome. *Journal of Infectious …*. Retrieved from http://jid.oxfordjournals.org/content/early/2014/02/12/infdis.jiu088.short

McMillan, D. R., Xiao, X., Shao, L., Graves, K., & Benjamin, I. J. (1998). Targeted disruption of heat shock transcription factor 1 abolishes thermotolerance and

protection against heat-inducible apoptosis. *Journal of Biological Chemistry*, *273*(13), 7523–7528. doi:10.1074/jbc.273.13.7523

Mejias, A., Dimo, B., Suarez, N. M., Garcia, C., Suarez-Arrabal, M. C., Jartti, T., … Ramilo, O. (2013). Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Medicine*, *10*(11), e1001549. doi:10.1371/journal.pmed.1001549

Mendillo, M. L., Santagata, S., Koeva, M., Bell, G. W., Hu, R., Tamimi, R. M., … Lindquist, S. (2012). HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. *Cell*, *150*(3), 549–62. doi:10.1016/j.cell.2012.06.031

Mercer, J., Snijder, B., Sacher, R., Burkard, C., Bleck, C. K. E., Stahlberg, H., … Helenius, A. (2012). RNAi Screening Reveals Proteasome- and Cullin3-Dependent Stages in Vaccinia Virus Infection. *Cell Reports*, *2*(4), 1036–1047. doi:10.1016/j.celrep.2012.09.003

Misasi, J., & Sullivan, N. J. (2014). Camouflage and Misdirection: The Full-On Assault of Ebola Virus Disease. *Cell*, *159*(3), 477–486. doi:10.1016/j.cell.2014.10.006

Mohamadzadeh, M., Chen, L., & Schmaljohn, A. L. (2007). How Ebola and Marburg viruses battle the immune system. *Nature Reviews. Immunology*, *7*(7), 556–67. doi:10.1038/nri2098

Monath, T. P., Newhouse, V. F., Kemp, G. E., Setzer, H. W., & Cacciapuoti, a. (1974). Lassa virus isolation from Mastomys natalensis rodents during an epidemic in Sierra Leone. *Science (New York, N.Y.)*, *185*(147), 263–265. doi:10.1126/science.185.4147.263

Morimoto, R. I. (1998). Regulation of the heat-shock transcriptional response: cross talk between a family of heat-shock factors, molecular chaperones, and negative regulators. *Genes Dev*, *12*, 3788–3796. doi:10.1101/gad.12.24.3788

Moser, T. S., Jones, R. G., Thompson, C. B., Coyne, C. B., & Cherry, S. (2010). A kinome RNAi screen identified AMPK as promoting poxvirus entry through the control of actin dynamics. *PLoS Pathogens*, *6*(6). doi:10.1371/journal.ppat.1000954

Nagai, N., Nakai, A., & Nagata, K. (1995). Quercetin suppresses heat shock response by down regulation of HSF1. *Biochemical and Biophysical Research Communications*, *208*(3), 1099–1105. doi:10.1006/bbrc.1995.1447

Nagarajan, N., & Pop, M. (2013). Sequence assembly demystified. *Nature Reviews. Genetics*, *14*(3), 157–67. doi:10.1038/nrg3367

Nakanishi, Y., Lu, B., Gerard, C., & Iwasaki, A. (2009). CD8(+) T lymphocyte mobilization to virus-infected tissue requires CD4(+) T-cell help. *Nature*, *462*(7272), 510–513. doi:10.1038/nature08511

Niklasson, B. S., Jahrling, P. B., & Peters, C. J. (1984). Detection of Lassa virus antigens and Lassa virus-specific immunoglobulins G and M by enzyme-linked

immunosorbent assay. *Journal of Clinical Microbiology*, *20*(2), 239–44. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=271295&tool=pmcentrez&rendertype=abstract

Nikolenko, S. I., Korobeynikov, A. I., & Alekseyev, M. a. (2013). BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics*, *14 Suppl 1*(Suppl 1), S7. doi:10.1186/1471-2164-14-S1-S7

Novoradovskaya, N., Whitfield, M. L., Basehore, L. S., Novoradovsky, A., Pesich, R., Usary, J., … Braman, J. (2004). Universal Reference RNA as a standard for microarray experiments. *BMC Genomics*, *5*(1), 20. doi:10.1186/1471-2164-5-20

Ohnishi, K., Takahashi, A., Yokota, S., & Ohnishi, T. (2004). Effects of a heat shock protein inhibitor KNK437 on heat sensitivity and heat tolerance in human squamous cell carcinoma cell lines differing in p53 status. *Int.J.Radiat.Biol.*, *80*(0955-3002 LA - eng PT - Journal Article RN - 0 (Benzhydryl Compounds) RN - 0 (DNA-Binding Proteins) RN - 0 (Heat-Shock Proteins) RN - 0 (KNK 437) RN - 0 (Protein p53) RN - 0 (Pyrrolidinones) RN - 136111-36-9 (heat shock factor, human) RN - 9007-49-2 (DNA)), 607–614.

Paessler, S., & Walker, D. H. (2013). Pathogenesis of the viral hemorrhagic fevers. *Annual Review of Pathology*, *8*, 411–40. doi:10.1146/annurev-pathol-020712-164041

Page, T. J., Sikder, D., Yang, L., Pluta, L., Wolfinger, R. D., Kodadek, T., & Thomas, R. S. (2006). Genome-wide analysis of human HSF1 signaling reveals a transcriptional program linked to cellular adaptation and survival. *Molecular bioSystems*, *2*(12), 627–639. doi:10.1039/b606129j

Panning, M., Emmerich, P., Olschläger, S., Bojenko, S., Koivogui, L., Marx, A., … Drosten, C. (2010). Laboratory diagnosis of Lassa fever, liberia. *Emerging Infectious Diseases*, *16*(6), 1041–3. doi:10.3201/eid1606.100040

Parrish, S., & Moss, B. (2007). Characterization of a second vaccinia virus mRNA-decapping enzyme conserved in poxviruses. *Journal of Virology*, *81*(23), 12973–12978. doi:10.1128/JVI.01668-07

Parrish, S., Resch, W., & Moss, B. (2007). Vaccinia virus D10 protein has mRNA decapping activity, providing a mechanism for control of host and viral gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(7), 2139–2144. doi:10.1073/pnas.0611685104

Peters, C. J., & LeDuc, J. W. (1999). An introduction to Ebola: the virus and the disease. *The Journal of Infectious Diseases*, *179 Suppl* , ix–xvi. doi:10.1086/514322

Peters, C. J., & Zaki, S. R. (2002). Role of the endothelium in viral hemorrhagic fevers. *Critical Care Medicine*, *30*(5 Suppl), S268–S273. doi:10.1097/00003246-200205001-00016

Pigott, D. M., Golding, N., Mylne, a., Huang, Z., Weiss, D. J., Brady, O. J., … Hay, S. I. (2015). Mapping the zoonotic niche of Marburg virus disease in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, (March), 366–378. doi:10.1093/trstmh/trv024

Pirkkala, L., Nykänen, P., & Sistonen, L. (2001). Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *The FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *15*(7), 1118–1131. doi:10.1096/fj00-0294rev

Prins, K. C., Cárdenas, W. B., & Basler, C. F. (2009). Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1. *Journal of Virology*, *83*(7), 3069–3077. doi:10.1128/JVI.01875-08

Prins, K. C., Delpeut, S., Leung, D. W., Reynard, O., Volchkova, V. a, Reid, S. P., … Basler, C. F. (2010). Mutations abrogating VP35 interaction with double-stranded RNA render Ebola virus avirulent in guinea pigs. *Journal of Virology*, *84*(6), 3004–3015. doi:10.1128/JVI.02459-09

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., … Claverie, J.-M. (2004). The 1.2-megabase genome sequence of Mimivirus. *Science (New York, N.Y.)*, *306*(5700), 1344–1350. doi:10.1126/science.1101485

Rasmussen, A. L., Tchitchek, N., Safronetz, D., Carter, V. S., Williams, C. M., Haddock, E., … Katze, M. G. (2014). Delayed inflammatory and cell death responses are associated with reduced pathogenicity in Lujo virus-infected cynomolgus macaques. *Journal of Virology*, (December). doi:10.1128/JVI.02246-14

Rawat, P., & Mitra, D. (2011). Cellular heat shock factor 1 positively regulates human immunodeficiency virus-1 gene expression and replication by two distinct pathways. *Nucleic Acids Research*, *39*(14), 5879–5892. doi:10.1093/nar/gkr198

Reed, D. S., Hensley, L. E., Geisbert, J. B., Jahrling, P. B., & Geisbert, T. W. (2004). Depletion of peripheral blood T lymphocytes and NK cells during the course of ebola hemorrhagic Fever in cynomolgus macaques. *Viral Immunology*, *17*(3), 390–400. doi:10.1089/vim.2004.17.390

Reid, S. P., Leung, L. W., Hartman, A. L., Martinez, O., Shaw, M. L., Carbonnelle, C., … Basler, C. F. (2006). Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *Journal of Virology*, *80*(11), 5156–5167. doi:10.1128/JVI.02349-05

Resch, W., Hixson, K. K., Moore, R. J., Lipton, M. S., & Moss, B. (2007). Protein composition of the vaccinia virus mature virion. *Virology*, *358*(1), 233–247. doi:10.1016/j.virol.2006.08.025

Rimoin, A. W., Mulembakani, P. M., Johnston, S. C., Lloyd Smith, J. O., Kisalu, N. K., Kinkela, T. L., … Muyembe, J.-J. (2010). Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic

Republic of Congo. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16262–16267. doi:10.1073/pnas.1005769107

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, *26*(1), 139–40. doi:10.1093/bioinformatics/btp616

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, *11*(3), R25. doi:10.1186/gb-2010-11-3-r25

Rougeron, V., Feldmann, H., Grard, G., Becker, S., & Leroy, E. M. (2015). Ebola and Marburg haemorrhagic fever. *Journal of Clinical Virology*, *64*, 111–119. doi:10.1016/j.jcv.2015.01.014

Rubins, K. H., Hensley, L. E., Wahl-Jensen, V., Daddario DiCaprio, K. M., Young, H. a, Reed, D. S., … Geisbert, T. W. (2007). The temporal program of peripheral blood gene expression in the response of nonhuman primates to Ebola hemorrhagic fever. *Genome Biology*, *8*(8), R174. doi:10.1186/gb-2007-8-8-r174

Russell, P. K., & Nisalak, a. (1967). Dengue virus identification by the plaque reduction neutralization test. *Journal of Immunology (Baltimore, Md. : 1950)*, *99*(2), 291–296.

Russier, M., Reynard, S., Tordo, N., & Baize, S. (2012). NK cells are strongly activated by Lassa and Mopeia virus-infected human macrophages in vitro but do not mediate virus suppression. *European Journal of Immunology*, *42*(7), 1822–1832. doi:10.1002/eji.201142099

Safronetz, D., Lopez, J. E., Sogoba, N., Traore', S. F., Raffel, S. J., Fischer, E. R., … Feldmann, H. (2010). Detection of lassa virus, mali. *Emerging Infectious Diseases*, *16*(7), 1123–1126. doi:10.3201/eid1607.100146

Santagata, S., Mendillo, M. L., Tang, Y., Subramanian, A., Perley, C. C., Roche, S. P., … Lindquist, S. (2013). Tight coordination of protein translation and HSF1 activation supports the anabolic malignant state. *Science (New York, N.Y.)*, *341*(6143), 1238303. doi:10.1126/science.1238303

Santoro, M. G. (2000). Heat shock factors and the control of the stress response. *Biochem Pharmacol*, *59*(1), 55–63. doi:S0006-2952(99)00299-3 [pii]

Satheshkumar, P. S., Anton, L. C., Sanz, P., & Moss, B. (2009). Inhibition of the ubiquitin-proteasome system prevents vaccinia virus DNA replication and expression of intermediate and late genes. *Journal of Virology*, *83*(6), 2469–2479. doi:10.1128/JVI.01986-08

Schnittler, H. J., & Feldmann, H. (1998). Marburg and Ebola hemorrhagic fevers: does the primary course of infection depend on the accessibility of organ-specific macrophages? *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *27*(2), 404–406.

Schoggins, J. W., Wilson, S. J., Panis, M., Murphy, M. Y., Jones, C. T., Bieniasz, P., & Rice, C. M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, *472*(7344), 481–5. doi:10.1038/nature09907

Sedger, L., Ramshaw, I., Condie, A., Medveczky, J., Braithwaite, A., & Ruby, J. (1996). Vaccinia virus replication is independent of cellular HSP72 expression which is induced during virus infection. *Virology*, *225*(2), 423–427. doi:10.1006/viro.1996.0619

Shaffer, J. G., Grant, D. S., Schieffelin, J. S., Boisen, M. L., Goba, A., Hartnett, J. N., … Garry, R. F. (2014). Lassa fever in post-conflict sierra leone. *PLoS Neglected Tropical Diseases*, *8*(3), e2748. doi:10.1371/journal.pntd.0002748

Shaner, L., & Morano, K. A. (2007). All in the family: atypical Hsp70 chaperones are conserved modulators of Hsp70 activity. *Cell Stress & Chaperones*, *12*(1), 1–8. doi:10.1379/CSC-245R.1

Shimojima, M., Stroher, U., Ebihara, H., Feldmann, H., & Kawaoka, Y. (2012). Identification of Cell Surface Molecules Involved in Dystroglycan-Independent Lassa Virus Cell Entry. *Journal of Virology*, *86*(4), 2067–2078. doi:10.1128/JVI.06451-11

Shin, K. D., Lee, M. Y., Shin, D. S., Lee, S., Son, K. H., Koh, S., … Han, D. C. (2005). Blocking tumor cell migration and invasion with biphenyl isoxazole derivative KRIBB3, a synthetic molecule that inhibits Hsp27 phosphorylation. *Journal of Biological Chemistry*, *280*(50), 41439–41448. doi:10.1074/jbc.M507209200

Shin, K. D., Yoon, Y. J., Kang, Y.-R., Son, K.-H., Kim, H. M., Kwon, B.-M., & Han, D. C. (2008). KRIBB3, a novel microtubule inhibitor, induces mitotic arrest and apoptosis in human cancer cells. *Biochemical Pharmacology*, *75*(2), 383–394. doi:10.1016/j.bcp.2007.08.027

Sivan, G., Martin, S. E., Myers, T. G., Buehler, E., Szymczyk, K. H., Ormanoglu, P., & Moss, B. (2013). Human genome-wide RNAi screen reveals a role for nuclear pore proteins in poxvirus morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(9), 3519–24. doi:10.1073/pnas.1300708110

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*(1), Article3. doi:10.2202/1544-6115.1027

Smyth, G. K., & Speed, T. (2003). Normalization of cDNA microarray data. *Methods*, *31*(4), 265–273. doi:10.1016/S1046-2023(03)00155-5

Ströher, U., West, E., Bugany, H., Klenk, H. D., Schnittler, H. J., & Feldmann, H. (2001). Infection and activation of monocytes by Marburg and Ebola viruses. *Journal of Virology*, *75*(22), 11025–11033. doi:10.1128/JVI.75.22.11025-11033.2001

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., & Ebert, B. L. (2005). Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide.

Teale, A., Campbell, S., Van Buuren, N., Magee, W. C., Watmough, K., Couturier, B., … Barry, M. (2009). Orthopoxviruses require a functional ubiquitin-proteasome system for productive replication. *Journal of Virology*, *83*(5), 2099–2108. doi:10.1128/JVI.01753-08

Teferi, W. M., Dodd, K., Maranchuk, R., Favis, N., & Evans, D. H. (2013). A whole-genome RNA interference screen for human cell factors affecting myxoma virus replication. *Journal of Virology*, *87*(8), 4623–41. doi:10.1128/JVI.02617-12

Ter Meulen, J., Badusche, M., Kuhnt, K., Doetze, a, Satoguina, J., Marti, T., … Hoerauf, a. (2000). Characterization of human CD4(+) T-cell clones recognizing conserved and variable epitopes of the Lassa virus nucleoprotein. *Journal of Virology*, *74*(5), 2186–2192. doi:10.1128/JVI.74.5.2186-2192.2000

Ter Meulen, J., Badusche, M., Satoguina, J., Strecker, T., Lenz, O., Loeliger, C., … Hoerauf, A. (2004). Old and New World arenaviruses share a highly conserved epitope in the fusion domain of the glycoprotein 2, which is recognized by Lassa virus-specific human CD4+ T-cell clones. *Virology*, *321*(1), 134–143. doi:10.1016/j.virol.2003.12.013

Timen, A. (2009). Response to Imported Case of Marburg Hemorrhagic Fever, the Netherlands. *Emerging Infectious Diseases*, *15*(8), 1171–1175. doi:10.3201/eid1508.090051

Towner, J. S., Amman, B. R., Sealy, T. K., Reeder Carroll, S. a., Comer, J. a., Kemp, A., … Rollin, P. E. (2009). Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. *PLoS Pathogens*, *5*(7). doi:10.1371/journal.ppat.1000536

Towner, J. S., Khristova, M. L., Sealy, T. K., Vincent, M. J., Erickson, B. R., Bawiec, D. a, … Nichol, S. T. (2006). Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. *Journal of Virology*, *80*(13), 6497–516. doi:10.1128/JVI.00069-06

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)*, *25*(9), 1105–11. doi:10.1093/bioinformatics/btp120

Trinklein, N. D., Murray, J. I., Hartman, S. J., Botstein, D., & Myers, R. M. (2004). The Role of Heat Shock Transcription Factor 1 in the Genome-wide Regulation of the Mammalian Heat Shock Response □. *Molecular Biology of the Cell*, *15*(March), 1254–1261. doi:10.1091/mbc.E03

Twenhafel, N. a, Mattix, M. E., Johnson, J. C., Robinson, C. G., Pratt, W. D., Cashman, K. a, … Honko, a N. (2013). Pathology of experimental aerosol Zaire ebolavirus infection in rhesus macaques. *Veterinary Pathology*, *50*(3), 514–29. doi:10.1177/0300985812469636

Uchiumi, F., Miyazaki, S., & Tanuma, S. (2011). The possible functions of duplicated ets (GGAA) motifs located near transcription start sites of various human genes. *Cellular and Molecular Life Sciences : CMLS*, *68*(12), 2039–51. doi:10.1007/s00018-011-0674-x

Ursic-Bedoya, R., Mire, C. E., Robbins, M., Geisbert, J. B., Judge, A., MacLachlan, I., & Geisbert, T. W. (2014). Protection against lethal Marburg virus infection mediated by lipid encapsulated small interfering RNA. *The Journal of Infectious Diseases*, *209*(4), 562–70. doi:10.1093/infdis/jit465

Valmas, C., Grosch, M. N., Schümann, M., Olejnik, J., Martinez, O., Best, S. M., … Mühlberger, E. (2010). Marburg virus evades interferon responses by a mechanism distinct from Ebola virus. *PLoS Pathogens*, *6*(1). doi:10.1371/journal.ppat.1000721

Van der Kuyl, A. C., van den Burg, R., Zorgdrager, F., Groot, F., Berkhout, B., & Cornelissen, M. (2007). Sialoadhesin (CD169) expression in CD14+ cells is upregulated early after HIV-1 infection and increases during disease progression. *PloS One*, *2*(2), e257. doi:10.1371/journal.pone.0000257

Van Vliet, K., Mohamed, M. R., Zhang, L., Villa, N. Y., Werden, S. J., Liu, J., & McFadden, G. (2009). Poxvirus proteomics and virus-host protein interactions. *Microbiology and Molecular Biology Reviews : MMBR*, *73*(4), 730–749. doi:10.1128/MMBR.00026-09

Villinger, F., Rollin, P. E., Brar, S. S., Chikkala, N. F., Winter, J., Sundstrom, J. B., … Peters, C. J. (1999). Markedly elevated levels of interferon (IFN)-gamma, IFN-alpha, interleukin (IL)-2, IL-10, and tumor necrosis factor-alpha associated with fatal Ebola virus infection. *The Journal of Infectious Diseases*, *179 Suppl* (Il), S188–91. doi:10.1086/514283

Vos, M. J., Hageman, J., Carra, S., & Kampinga, H. H. (2008). Structural and functional diversities between members of the human HSPB, HSPH, HSPA, and DNAJ chaperone families. *Biochemistry*. doi:10.1021/bi800639z

Voyer, J., & Heikkila, J. J. (2008). Comparison of the effect of heat shock factor inhibitor, KNK437, on heat shock- and chemical stress-induced hsp30 gene expression in Xenopus laevis A6 cells. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, *151*(2), 253–261. doi:10.1016/j.cbpa.2008.07.004

Wahl-Jensen, V., Kurz, S., Feldmann, F., Buehler, L. K., Kindrachuk, J., DeFilippis, V., … Feldmann, H. (2011). Ebola virion attachment and entry into human macrophages profoundly effects early cellular gene expression. *PLoS Neglected Tropical Diseases*, *5*(10), e1359. doi:10.1371/journal.pntd.0001359

Walker, D. H., McCormick, J. B., Johnson, K. M., Webb, P. a, Komba-Kono, G., Elliott, L. H., & Gardner, J. J. (1982). Pathologic and virologic study of fatal Lassa fever in man. *The American Journal of Pathology*, *107*(3), 349–356.

Wang, F. W., Wu, X. R., Liu, W. J., Liao, Y. J., Lin, S., Zong, Y. S., … Xie, D. (2011). Heat shock factor 1 upregulates transcription of Epstein-Barr Virus nuclear antigen 1 by binding to a heat shock element within the BamHI-Q promoter. *Virology*, *421*(2), 184–191. doi:10.1016/j.virol.2011.10.001

Wang, Y., Trepel, J. B., Neckers, L. M., & Giaccone, G. (2010). STA-9090, a small-molecule Hsp90 inhibitor for the potential treatment of cancer. *Current Opinion in Investigational Drugs (London, England : 2000)*, *11*(12), 1466–1476.

Warfield, K. L., Bradfute, S. B., Wells, J., Lofts, L., Cooper, M. T., Alves, D. A., … Bavari, S. (2009). Development and characterization of a mouse model for Marburg hemorrhagic fever. *Journal of Virology*, *83*(13), 6404–6415. doi:10.1128/JVI.00126-09

Wauquier, N., Becquart, P., Padilla, C., Baize, S., & Leroy, E. M. (2010). Human fatal zaire ebola virus infection is associated with an aberrant innate immunity and with massive lymphocyte apoptosis. *PLoS Neglected Tropical Diseases*, *4*(10). doi:10.1371/journal.pntd.0000837

Westerheide, S. D., Kawahara, T. L. A., Orton, K., & Morimoto, R. I. (2006). Triptolide, an inhibitor of the human heat shock response that enhances stress-induced cell death. *The Journal of Biological Chemistry*, *281*(14), 9616–9622. doi:10.1074/jbc.M512044200

Whittaker, S. R., Theurillat, J. P., Van Allen, E., Wagle, N., Hsiao, J., Cowley, G. S., … Garraway, L. A. (2013). A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. *Cancer Discovery*, *3*(3), 350–362. doi:10.1158/2159-8290.CD-12-0470

WHO Ebola Response Team. (2014). Ebola Virus Disease in West Africa — The First 9 Months of the Epidemic and Forward Projections. *The England New Journal of Medecine*, 1–15. doi:10.1056/NEJMoa1411100

Wong, G., Kobinger, G. P., & Qiu, X. (2014). Characterization of host immune responses in Ebola virus infections. *Expert Review of Clinical Immunology*, *10*(6), 781–90. doi:10.1586/1744666X.2014.908705

Wulff, H., & Johnson, K. M. (1979). Immunoglobulin M and G responses measured by immunofluorescence in patients with Lassa or Marburg virus infections. *Bulletin of the World Health Organization*, *57*(4), 631–635.

Xie, S., Wu, H., Wang, Q., Cogswell, J. P., Husain, I., Conn, C., … Dai, W. (2001). Plk3 functionally links DNA damage to cell cycle arrest and apoptosis at least in part via the p53 pathway. *The Journal of Biological Chemistry*, *276*(46), 43305–12. doi:10.1074/jbc.M106050200

Yang, Z., Bruno, D. P., Martens, C. A., Porcella, S. F., & Moss, B. (2010). Simultaneous high-resolution analysis of vaccinia virus and host cell transcriptomes by deep RNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11513–8. doi:10.1073/pnas.1006594107

Yang, Z., Reynolds, S. E., Martens, C. A., Bruno, D. P., Porcella, S. F., & Moss, B. (2011). Expression profiling of the intermediate and late stages of poxvirus replication. *Journal of Virology*, *85*(19), 9899–9908. doi:10.1128/JVI.05446-11

Yen, B., Mulder, L. C. F., Martinez, O., & Basler, C. F. (2014). Molecular Basis for Ebolavirus VP35 Suppression of Human Dendritic Cell Maturation. *Journal of Virology*, *88*(21), 12500–12510. doi:10.1128/JVI.02163-14

Yen, J. Y., Garamszegi, S., Geisbert, J. B., Rubins, K. H., Geisbert, T. W., Honko, A., … Hensley, L. E. (2011). Therapeutics of Ebola hemorrhagic fever: whole-genome transcriptional analysis of successful disease mitigation. *The Journal of Infectious Diseases*, *204 Suppl* (Suppl 3), S1043–52. doi:10.1093/infdis/jir345

Yokota, S. I., Kitahara, M., & Nagata, K. (2000). Benzylidene lactam compound, KNK437, a novel inhibitor of acquisition of thermotolerance and heat shock protein induction in human colon carcinoma cells. *Cancer Research*, *60*(11), 2942–2948.

Yoon, Y. J., Kim, J. A., Shin, K. D., Shin, D. S., Han, Y. M., Lee, Y. J., … Han, D. C. (2011). KRIBB11 inhibits HSP70 synthesis through inhibition of heat shock factor 1 function by impairing the recruitment of positive transcription elongation factor b to the hsp70 promoter. *Journal of Biological Chemistry*, *286*(3), 1737–1747. doi:10.1074/jbc.M110.179440

Yun, N. E., & Walker, D. H. (2012). Pathogenesis of lassa fever. *Viruses*, *4*(10), 2031–2048. doi:10.3390/v4102031

Zaas, a. K., Burke, T., Chen, M., McClain, M., Nicholson, B., Veldman, T., … Ginsburg, G. S. (2013). A Host-Based RT-PCR Gene Expression Signature to Identify Acute Respiratory Viral Infection. *Science Translational Medicine*, *5*(203), 203ra126–203ra126. doi:10.1126/scitranslmed.3006280

Zaas, A. K., Chen, M., Varkey, J., Veldman, T., Hero, A. O., Lucas, J., … Ginsburg, G. S. (2009). Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host & Microbe*, *6*(3), 207–17. doi:10.1016/j.chom.2009.07.006

Zampieri, C. A., Sullivan, N. J., & Nabel, G. J. (2007). Immunopathology of highly virulent pathogens: insights from Ebola virus. *Nature Immunology*, *8*, 1159–1164. doi:10.1038/ni1519

Zapata, J. C., Carrion, R., Patterson, J. L., Crasta, O., Zhang, Y., Mani, S., … Salvato, M. S. (2013). Transcriptome Analysis of Human Peripheral Blood Mononuclear Cells Exposed to Lassa Virus and to the Attenuated Mopeia/Lassa Reassortant 29 (ML29), a Vaccine Candidate. *PLoS Neglected Tropical Diseases*, *7*(9), e2406. doi:10.1371/journal.pntd.0002406

Zapata, J. C., & Salvato, M. S. (2015). Genomic profiling of host responses to Lassa virus: therapeutic potential from primate to man. *Future Virology*, *10*(3), 233–256. doi:10.2217/fvl.15.1

Zerbino, D. R., & Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–9. doi:10.1101/gr.074492.107

Zeremski, M., Petrovic, L. M., Chiriboga, L., Brown, Q. B., Yee, H. T., Kinkhabwala, M., … Talal, A. H. (2008). Intrahepatic levels of CXCR3-associated chemokines correlate with liver inflammation and fibrosis in chronic hepatitis C. *Hepatology*, *48*(5), 1440–1450. doi:10.1002/hep.22500

Zhang, S.-J., Liu, C.-J., Shi, M., Kong, L., Chen, J.-Y., Zhou, W.-Z., … Li, C.-Y. (2013). RhesusBase: a knowledgebase for the monkey research community. *Nucleic Acids Research*, *41*(Database issue), D892–905. doi:10.1093/nar/gks835

CHAPTER 8 CURRICULUM VITAE