

UNTARGETED LCMS SERUM METABOLOMICS OF THE SIERRA LEONEAN
LASSA FEVER PATIENT AND METAANALYSIS OF THE VIRION PROTEOME

AN ABSTRACT

SUBMITTED ON THE 8TH OF APRIL, 2016

TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

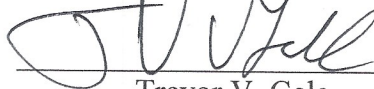
OF THE GRADUATE SCHOOL

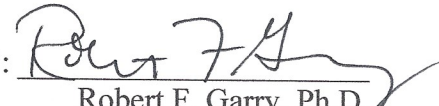
OF TULANE UNIVERISTY

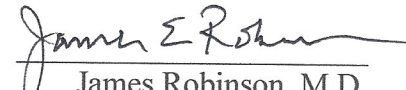
FOR THE DGREE OF DOCTOR

OF

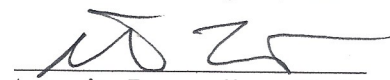
PHILOSOPHY BY


Trevor V. Gale

APPROVED: 
Robert F. Garry, Ph.D.
Director


James Robinson, M.D.


Erik Flemington, Ph.D.


Antonito Panganiban, Ph.D.

ABSTRACT

RNA viruses causing viral hemorrhagic fever (VHF) are responsible for a large global disease incidence where the endemic burden has a disproportionately negative affect on countries with insufficient medical resources. Lassa virus alone causes some 300,000 cases of Lassa fever—a prototypical VHF—annually with fatality exceeding 30% for acute patients. The febrile illness VHF presents with is often indistinguishable from other endemic infectious diseases where the hallmark manifestations of VHF, e.g. profuse hemorrhage, are the exception, not the rule. Tracking the progression of disease poses significant hindrance where no approved prophylactic or pharmaceutical treatment options exist for VHF. Multiple viruses cause VHF, which pose threats through globalization or nefarious release hastening the need for alternative, precise views of the disease progression and proposed treatment options, serving to benefit both the afflicted endemic populations and local preparedness. It is worth interrogating the disease and the viruses that cause VHF to identify biomarkers of disease and targetable treatment opportunities. Systems Biology approaches render information-rich, complex biological matrixes to distinguishable, quantifiable datasets. We employ Liquid Chromatography coupled Mass Spectrometry (LCMS) to characterize the global biomolecular composition from two separate matrixes: metabolite composition of Lassa fever patient sera and protein composition of Lassa virus-like particles. Pairwise comparison of the metabolomes of endemic patient sera will yield differential features that correlate with disease progression while synergizing the protein composition of Lassa virus-like particles with a database of lipid-virion proteins will potentiate targetable host-virus interactions for pharmacological intervention. During the course of Lassa fever global

metabolic flux as a result of acute and terminal disease yields significant disparities between small molecules of multiple chemistries, notably of lipids. Analysis of this data through machine learning identifies biomarkers indicative of disease outcome where a class of pro-inflammatory lipid molecules predicts survival verse terminal disease. Lassa virus-like particles incorporate multiple host proteins observed in diverse enveloped viral particles during biogenesis that are targetable for small molecule inhibitors of viral replication. These results lay the foundational structure for further development of prognostic methods mapping the trajectory of VHF while identifying a priority ‘hit list’ for novel, broad-spectrum antiviral therapies.

UNTARGETED LCMS SERUM METABOLOMICS OF THE SIERRA LEONEAN
LASSA FEVER PATIENT AND METAANALYSIS OF THE VIRION PROTEOME

A DISSERTATION

SUBMITTED ON THE 8TH OF APRIL, 2016

TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

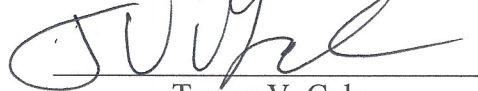
OF THE GRADUATE SCHOOL

OF TULANE UNIVERISTY

FOR THE DEGREE OF DOCTOR

OF

PHILOSOPHY BY



Trevor V. Gale

APPROVED:



Robert F. Garry, Ph.D.

Director



James Robinson, M.D.



Erik Flemington, Ph.D.



Antonito Panganiban, Ph.D.

ACKNOWLEDGEMENTS

There are many people whose support during the process of my professional development has been imperative. Firstly, I owe Dr. Robert Garry a debt of gratitude not only for facilitating the opportunity for me to earn this Ph.D but for his scientific vision and foresight. Secondly, the efforts of Timothy Horton need to be acknowledged here as he could just as easily be listed as an author of this document as it contains work we were partners in developing. I thank the members of my committee consisting of Drs. James Robinson, Erik Flemington, and Antonito Panganiban who were engaged and constructive. Drs. John Schieffelin, Ben Bradley, and Jeff Shaffer and Garry lab member Andrew Hoffmann went out of their way to work with me on shared endeavors and I am grateful. I am especially thankful to Dr. Jessica Hartnett who has never hesitated to offer her efforts on my behalf. Dr. Luis Branco was a particularly committed and unselfish mentor and I owe him many thanks. Finally, without the encouragement and love of my family I may not have ever ventured as far down this path as I've come. My wife Sara Gale is credited with identifying my passion for virology and her encouragement has allowed me to pursue my dreams. My parents Cassie and Kelley Gale only offered nurturing support and I am thankful, as this would have been unachievable without them. I owe my son, Theodore Vernon Gale most of all as he's the driving force behind my efforts. Nothing has given me more happiness and fulfillment than being your father.

TABLE OF CONTENTS

Acknowledgements.....	ii
List of Tables	ix
Introduction.....	1
Preamble	1
Lassa Virus	1
Lassa Fever.....	9
VHF Biomarkers	11
Metabolomics and Viral Disease.....	15
Proteomics and Enveloped Viral Particles	20
Specific Aims	22
Materials and Methods.....	25
Clinical sample derivation.....	25
Preparation of small molecules from patient sera	25
Liquid Chromatography Mass Spectrometry for small molecule detection.....	26
Data analysis.....	27
Mammalian Cell Culture and Antibodies.....	28
Generation of Lassa virus-like particles	29
Purifying Lassa virus-like particles, protein isolation, and LCMS analysis	29
Western blot.....	31

Generation of the Virion Proteome Database.....	32
Taxonomic Tree Generation.....	32
GO Annotation of the Virion Proteome Database.....	33
Compounds.....	33
Screening candidate compounds	34
Toxicity.....	34
Viruses.....	35
Viral titration via plaque assay	35
p24 ELISA	36
Pre-incubation.....	36
Residual/Carryover toxicity assay	37
Plaque inhibition assay	37
Plaque-forming unit reduction assay	38
Results.....	39
Aim 1	39
The non-descript presentation of febrile illness	39
The Sierra Leonean LFW patient serum metabolome is dynamic	46
Putatively identified molecular classes	57
Metabolites with biomarker property	61
Machine learning identifying metabolites with biomarker property	77
Aim 2	84
Host Proteins Exclusively From Extracellular Viral Particles	84
Augmenting the dataset	87
Lassa Virus-like particles and LCMS analysis.....	90
The Virion Proteome Database	95

High(er) throughput screening for novel antiviral activity.....	99
p24 ELISA.....	106
Pre-incubation and carryover toxicity	108
Screening hits in competent viral systems.....	111
Discussion.....	120
Aim 1	120
Aim 2	130
References.....	143

List of Figures

Figure 1	3
Figure 2	5
Figure 3	7
Figure 4	8
Figure 5	14
Figure 6	16
Figure 7	44
Figure 8	45
Figure 9	49
Figure 10	50
Figure 11	53
Figure 12	54
Figure 13	56
Figure 14	59
Figure 15	60
Figure 16	63
Figure 17	64
Figure 18	67

Figure 19	69
Figure 20	71
Figure 21	73
Figure 22	75
Figure 23	76
Figure 24	78
Figure 25	80
Table 6	82
Figure 26	83
Figure 27	86
Figure 28	89
Figure 29	92
Figure 30	94
Figure 31	96
Figure 32	98
Figure 33	101
Figure 34	102
Figure 35	103
Figure 36	104

Figure 37	105
Figure 38	107
Figure 39	110
Figure 40	112
Figure 41	114
Figure 42	116
Figure 43	118

LIST OF TABLES

Table 1	40
Table 2	42
Table 3	48
Table 4	52
Table 5	62
Table 7	91

INTRODUCTION

Preamble

Multiple transmissible agents including multicellular parasites, bacteria, and viruses cause infectious diseases. Despite being some of the simplest replicating biological entities, viruses are responsible for an enormous disease burden in man. From chronic, latent infections to rapid and extraordinarily virulent, viruses are as diverse as the disease they cause. Opposed to infectious disease caused by cellular organisms, viruses pose a substantial problem as they leverage host cellular processes towards to replication of viral progeny, making specific treatments especially challenging. Amongst viruses, Viral Hemorrhagic Fever (VHF) is the most notable disease, characterized by a non-descript febrile illness, i.e. “flu-like” symptoms, where any outward manifestation rarely defines disease. In limited cases, visually detectable symptoms/abnormalities are present such as rash or profuse bleeding from multiple anatomical but almost uniformly at the terminal stages of disease. With four viral families—the *Arenaviridae*, *Bunyaviridae*, *Filoviridae*, and *Flaviviridae*—harboring members causing VHF, there is an impetus for further refinement of the knowledge of VHF pathology. While many of VHF’s are characterized by random, rare, and sporadic outbreaks, Lassa fever—a VHF caused by Lassa Virus—is a poignant exception.

Lassa Virus

The family *Arenaviridae* are enveloped (lipid) viruses endemic to African, American, and European countries. Lassa virus (LV) is an old world Arenavirus

consisting of a bi-segmented negative-sense RNA genome encoding 4 genes yielding 6 proteins[1]. The genome structure and replication strategy of Arenaviruses is illustrated in **Figure 1**. Upon binding to its cellular receptor alpha dystroglycan(α -DG)[2] which mediates the pantropic nature of LV infection given the relative ubiquity of α -DG in extracellular matrix[3] LV enters a permissive cell via endocytosis[4]. Once inside a cell LV replicates in the cytoplasm[1]. Following low-pH dependent fusion[5], the virally encoded RNA dependent RNA polymerase (L) mediates transcription[6].

Figure 1

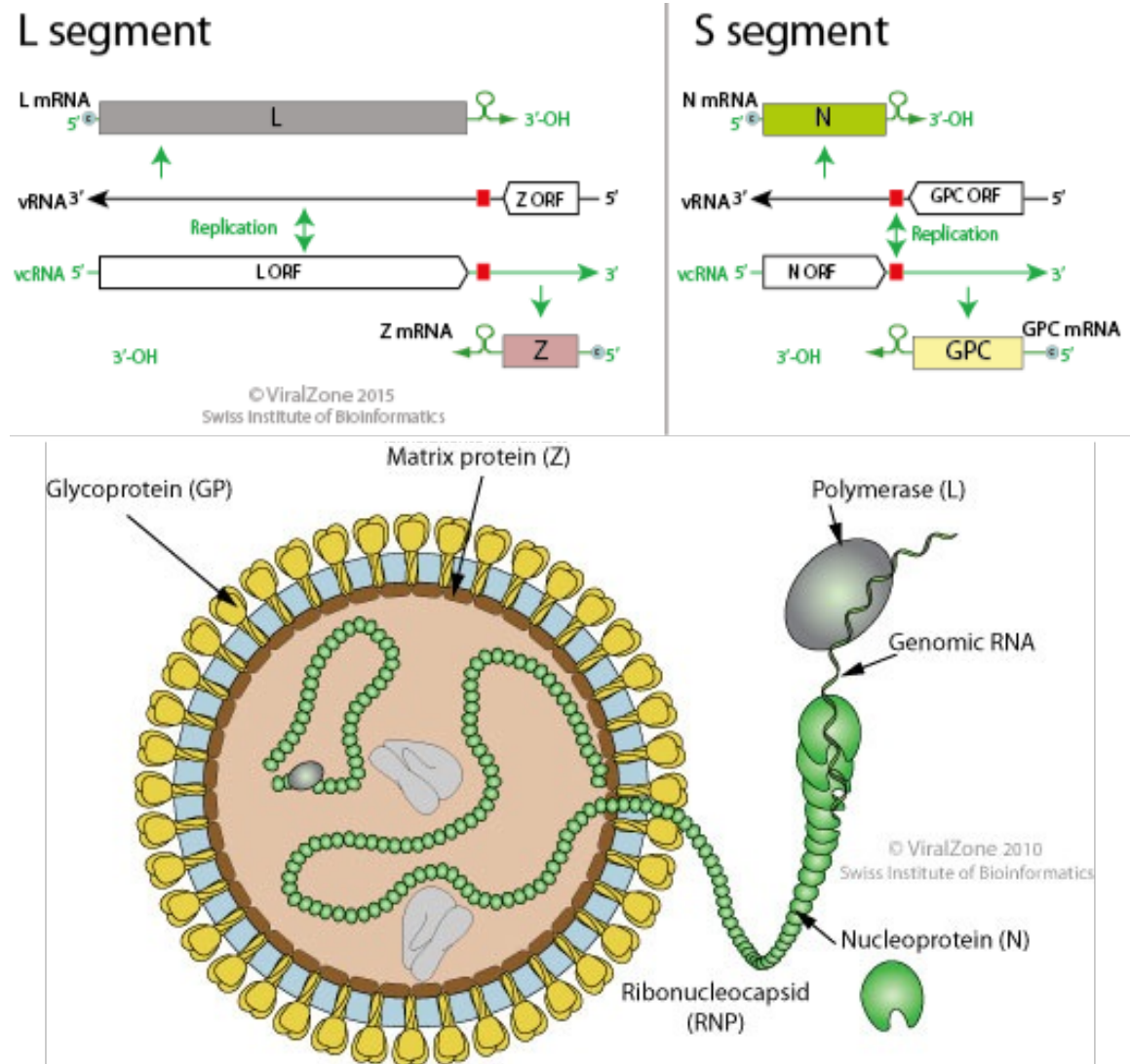


Figure 1 Arenavirus genome structure and replication strategy. Lassa virus encodes 4 genes yielding 6 proteins. Reprinted from [7].

Arenaviral polymerases mediate genomic replication where each of the two genomic segments can yield coding messenger RNA (mRNA) and an L-synthesized antigenomic RNA from which an additional mRNA species are synthesized[1, 8]. The polyprotein precursor of the viral glycoprotein (GP) is trafficked to the lumen of endoplasmic reticulum[9] where it is processed by cellular proteases into the competent GP spike protein which is subsequently acquired through egress at the cellular plasma membrane[10]. Glycoprotein synthesis is illustrated in **Figure 2**. Important features of GP synthesis include the multiple interactions of arenaviral GP with host cell processes[11]. Mature GP spike proteins translocate to areas of the cellular plasma membrane deriving a specific cholesterol content of the budding virion which has a direct effect on resultant infectivity[12].

Figure 2

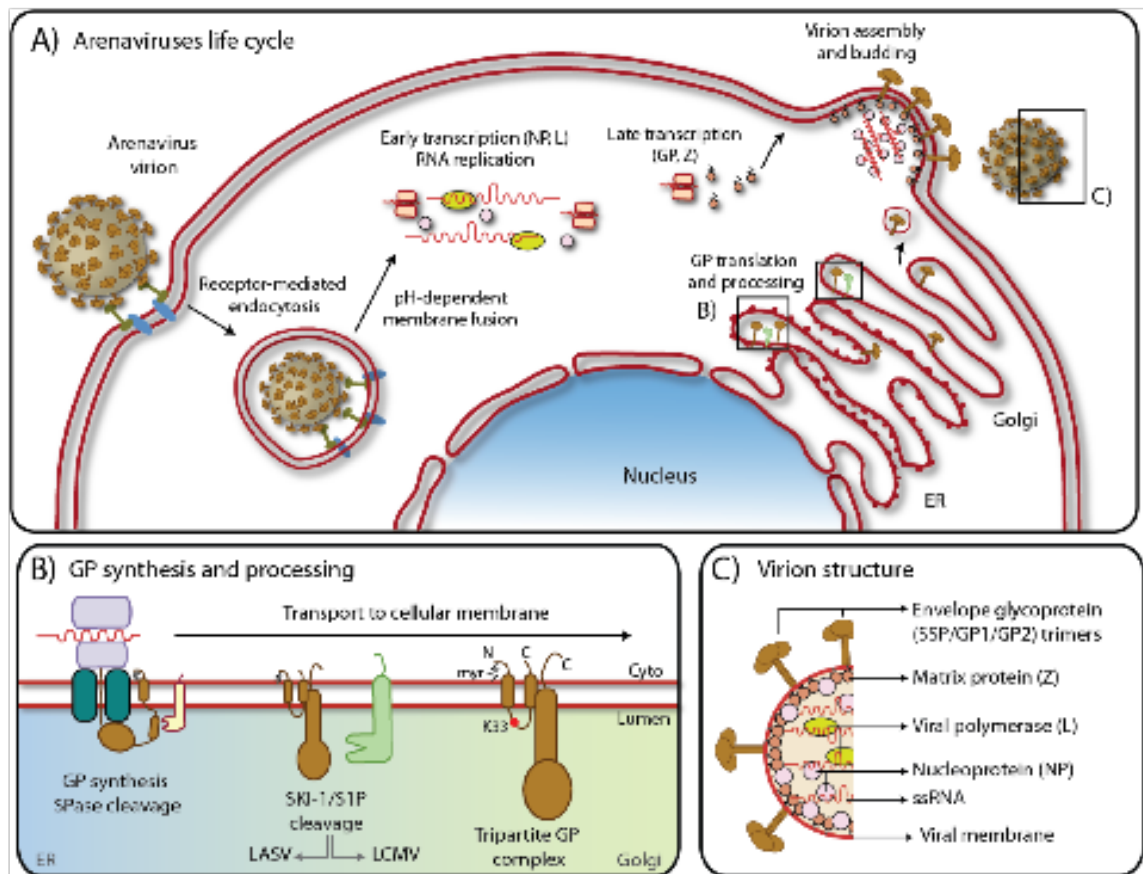


Figure 2 The cellular proteins with roles in glycoprotein synthesis are shown with particular emphasis paid to endoplasmic reticulum trafficking. Reprinted from [11].

The nucleoprotein (NP) is the most abundant viral protein whose intrinsic double-stranded RNA exonuclease activity is implicated in the inhibition of host innate immune response[13] including inhibition of innate signaling response[14] and natural killer cell activity[15]. The zinc finger matrix protein (Z) is responsible for viral egress[16] where cellular expression of Z alone results in the production of virus-like particles[16-18]. As shown in **Figure 3** and **Figure 4**, Z coordinates viral assembly and egress. The multiple functions of Z include the simultaneous and coordinated interaction of viral proteins and with cellular proteins pertinent to viral replication[19]. Virion assembly is presumed coordinated through the activity of Z interacting with NP, encapsulating L associated genome and trafficked to the inner leaflet of the plasma membrane and acquiring mature GP through Z-mediated egress[6, 12, 17]. Mature virions bud from infected cells and are competent for subsequent rounds of infection and replication in a variety of mammalian tissues.

Figure 3

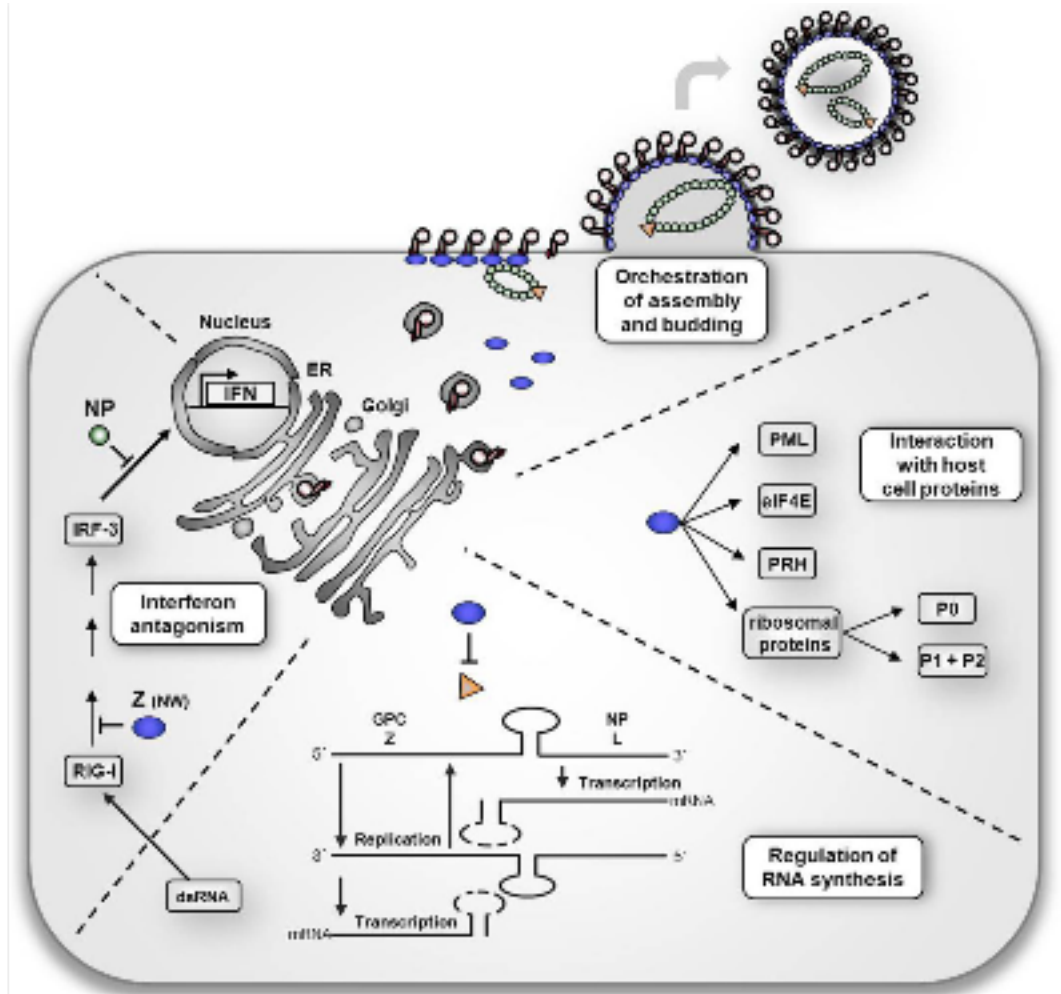


Figure 3 The matrix protein Z of arenaviruses. The matrix protein of arenaviruses drives multiple synthetic processes. Reprinted from [19].

Figure 4

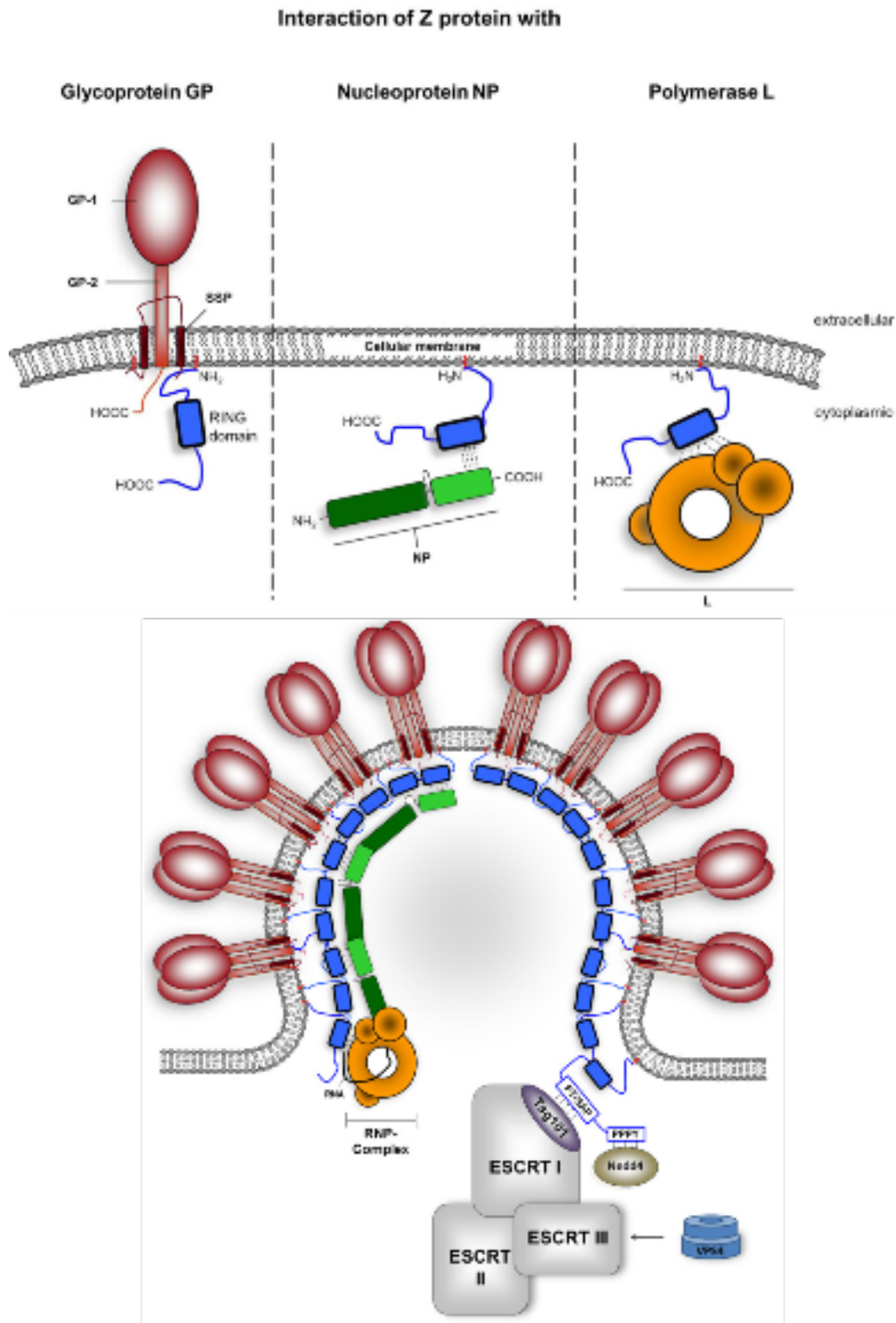


Figure 4 Cellular interacting partners the arenaviral Z protein. The matrix protein coordinates with the 3 other viral proteins (top) and multiple cellular proteins necessary for assembly and egress. Reprinted from [19].

The 40-200 nanometer (nM) Arenaviruses[17] appear grainy or sandy upon visualization with electron microscopy due to the presence of dark staining, cellular organelles—presumed to be ribosomes—decorating the exterior of the virion[18, 20]. Lassa virus is maintained in the natural environment by rodents of the genus *Mastomys*[21] which are presumed chronically infected. These animals account for the zoonotic classification of LF acting as vectors of LV by shedding infectious virus through excreta[22]. Therefore, ecological patterns of *Mastomys* closely correlate with epidemiological patterns of Lassa fever, notably that disease incidences increase in the dry season when resource-seeking behavior sends rodents into households, increasing exposure of humans and rodent excreta[23-25]. The molecular characterization of arenaviruses notwithstanding and in spite of the well-characterized host-vector relationship, LV is still a major causative agent of VHF, a disease whose acuity is striking in afflicted patients where there is no approved, specific prophylactic or therapeutic treatment option.

Lassa Fever

First discovered in 1969, Lassa fever (LF) was immediately recognized as a lethal disease[26]. Falling within the family of maladies known as viral hemorrhagic fever, LF is characterized primarily as a non-descript febrile illness with flu-like symptoms[25, 27, 28] where some patients deteriorate to a fulminant, terminal disease[29]. Lassa fever is predicted to cause hundreds of thousands of cases of disease annually in west and central African nations[30]. Historical data suggest a case fatality rate (CSF) of 14 to 20 percent[28, 30, 31]. For certain patient populations, including the young and pregnant women CFRs are appreciably higher[25, 30]. With the development and implementation

sensitive, modern molecular diagnostics a CFR of 60% is observed for those patients whom have LV antigenemia[25].

Lassa virus is transmitted through direct contact with virus on surfaces[22], contagious patients and fomites facilitating nosocomial infection[32] or via aerosols[33]. Regardless of route of infection, LV shows early replicative events in antigen presenting cells of the innate immune system, particularly Dendritic Cells (DC)[34]. In addition Macrophages ($M\phi$) are permissive to infection and as in DCs, infection by LV leaves the cell unable to produce inflammatory mediators typical for the amplification of innate and specific immune responses[34, 35]. Circulatory monocytes, $M\phi$, and DC are likely to contribute to systemic distribution of LV within the host[36] where the virus is observed pantropic following viremia, found in most tissues except for those of the nervous system[37, 38]. The confluence of LV tissue saturation compounded with dysregulation of the immune system and coagulopathy are hallmarks of LF and more generally, VHF.

Infection and resultant inhibition of inflammatory mediators, particularly of front-line defenses[34, 35] are accompanied by vascular leakage/permeability[39]. Though implicated, disseminated intravascular coagulation (DIC) does not appear to be a primary pathology of LF[39, 40]. Notably, platelet counts are found in the normal range compared to healthy controls and patients not suffering LF but platelet function is found depressed in LF patients[29, 39, 41]. The depression of platelet activity is progressively more pronounced through the course of LF and is mediated by a serum soluble factor. A similar observation is made in the serum of patients suffering Argentine hemorrhagic fever (AHF), caused by Junin virus (JV) a New World Arenavirus, and mechanistically shown to be virus-independent[42]. Hematological abnormalities in LF patients such as

platelet dysregulation or immunomodulatory molecules are candidates for monitoring disease progression when discernable amounts are observed differentially as the disease progresses.

VHF Biomarkers

A biomarker is defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”[43]. Multiple biological phenomena can therefore qualify as biomarkers but those representing *phenotype* are arguably the most desirable. Multiple studies have been implemented to measure the presence and dynamics of biomolecules in VHF, providing insights into the fundamentals of disease as well as examples of biomarkers.

The earliest biomarkers analyzed for correlation with LF outcome are serum infectious viral units, enzyme(s), and anti-Lassa immunoglobulin responses. Likely result of liver damage—a frequent pathology of LF—serum aspartate aminotransferase (AST) is elevated in LF patients and correlates with poor outcome[29, 44-46]. Elevated AST was also recently shown in a newly identified VHF, Lujo hemorrhagic fever (LHF) caused by Lujo virus where poor prognosis correlated with elevated serum AST level[47]. The optimal *diagnostic* biomarker is the presence of viral protein. Modern LV and Ebola diagnostics allow for the detection of viral antigen and have limited dynamic range for serial measure to monitor viremia[25, 46, 48]. Unsurprisingly, there is an inverse relationship between greater serum viral load and poor outcome, also observed in Ebola hemorrhagic fever and LF (EHV)[29, 44-46, 49]. Antibody responses follow the expected time-course for arenaviruses[50] where patients registering early IgM and viral

antigen experience slightly less mortality but in general antibody responses are poor indicators of LF outcome and at least partially confounded by immunoglobulin persistence[44, 46].

Metabolic and inflammatory markers have also been examined as indicators of the disease status in VHF. Multiple serum indicators have been shown to correlate with disease outcome with similarities in both LF and EHF. The cytokines Interleukin 6, 8, and 10 (IL-6, IL-8, IL-10), macrophage inflammatory protein 1 alpha & beta (MIP-1 α/β), and Interferons-alpha/gamma (INF- α/γ) are measured with increased concentration and correlate with poor outcome[44, 46, 49, 51-53].

Coagulopathies are also present in VHF and correlate with disease outcome. The coagulation cascade features multiple targetable proteins and metabolic substrates suitable for monitoring dysregulation of the coagulation cascade in a clinical scenario as highlighted in **Figure 5**. The aforementioned depression of platelet activity correlates with LF outcome where patients with lower platelet activity have poorer outcome[41]. In EHF multiple substrates and mediators of the coagulation cascade have been examined where higher mortality was associated with elevated thrombomodulin and ferritin[49] though there was no statistical difference of thrombomodulin between pediatric and adult Ebola patients[54]. Elevated D-dimer, a product of activated coagulation phenomena[55] has been correlated with increased mortality in EHV and LHF[47, 49]. In Dengue fever (DF) which presents with hemorrhagic manifestations in ~ 5% of cases, there is a presumed correlation between clinical outcome and coagulopathy though it is historically difficult to conclude given the lack of uniformity in available clinical data[56]. Taken together, multiple candidate biomarkers correlate with the severity of disease status in

VHF. These markers are well-characterized proteins and/or metabolites and have been detected with commercially available reagents; therefore these markers could be employed in a clinical setting for utility as prognostic indicators.

Figure 5

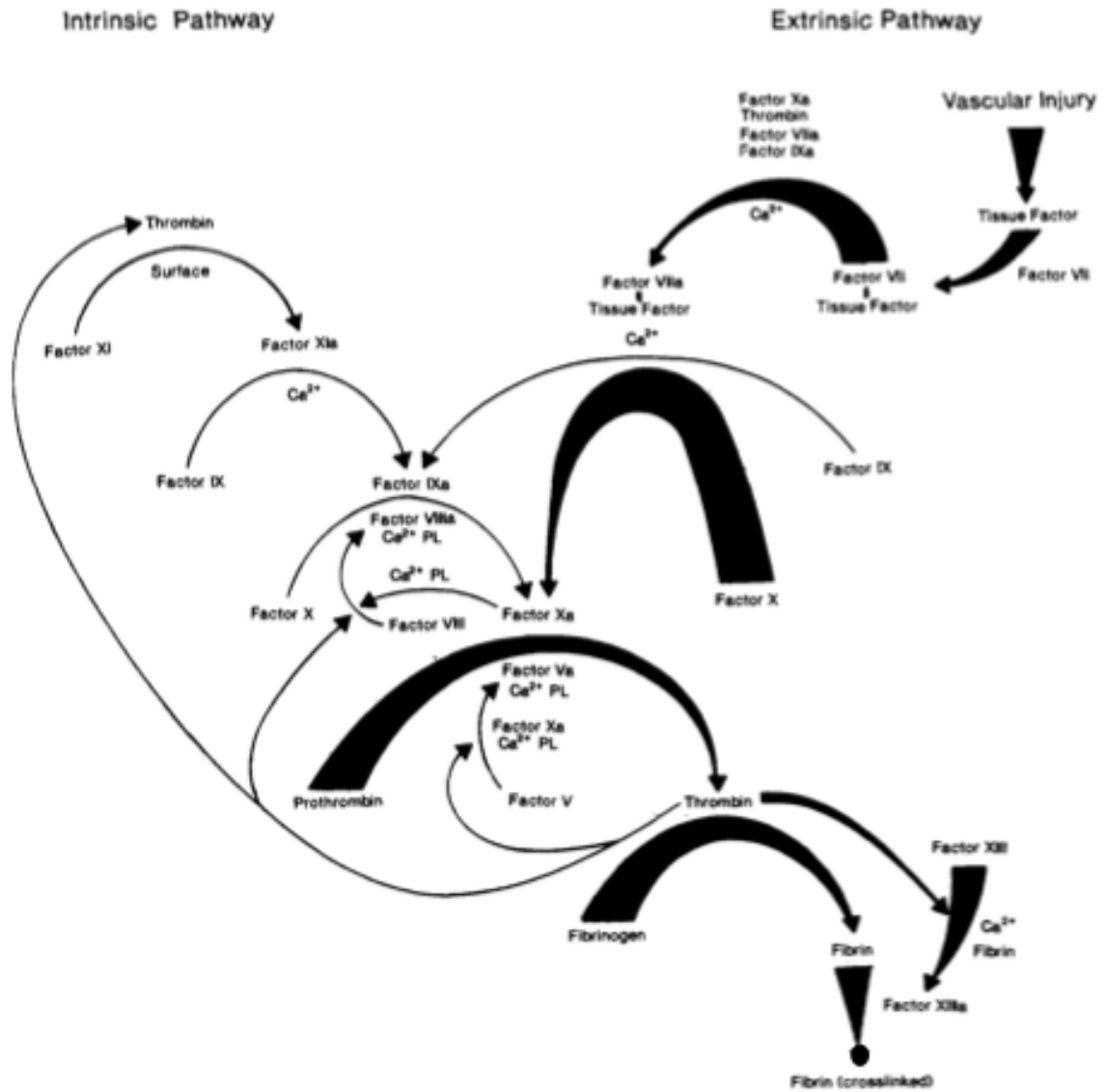


Figure 5 Coagulation cascade. Detailed proteins, substrates, products, and structural features of the coagulation cascade are detailed. Fibrin is the terminal product in this cascade and implicated in viral hemorrhagic fever. Reprinted from [57].

Metabolomics and Viral Disease

Metabolomics is the systems biology discipline seeking to isolate, detect, characterize, and identify all small molecules from a given sample. There are two common rationales in metabolomics research: *targeted* where a select predetermined group of molecules are examined and *untargeted* where all small molecules are examined (**Figure 6**)[58]. When a metabolic pathway of interest is already known a targeted investigation holds more direct hypothesis-testing utility. Consistent with many systems biology research endeavors, untargeted metabolomics is “hypothesis generating” where observations in group-to-group metabolite disparity potentiate targets for further analyses. Untargeted metabolomics investigations have been performed on multiple biological matrixes of clinical origin with serum/plasma providing information-rich, experimentally reproducible[59], and routinely accessed specimen in the case of VFH[25, 48].

Figure 6

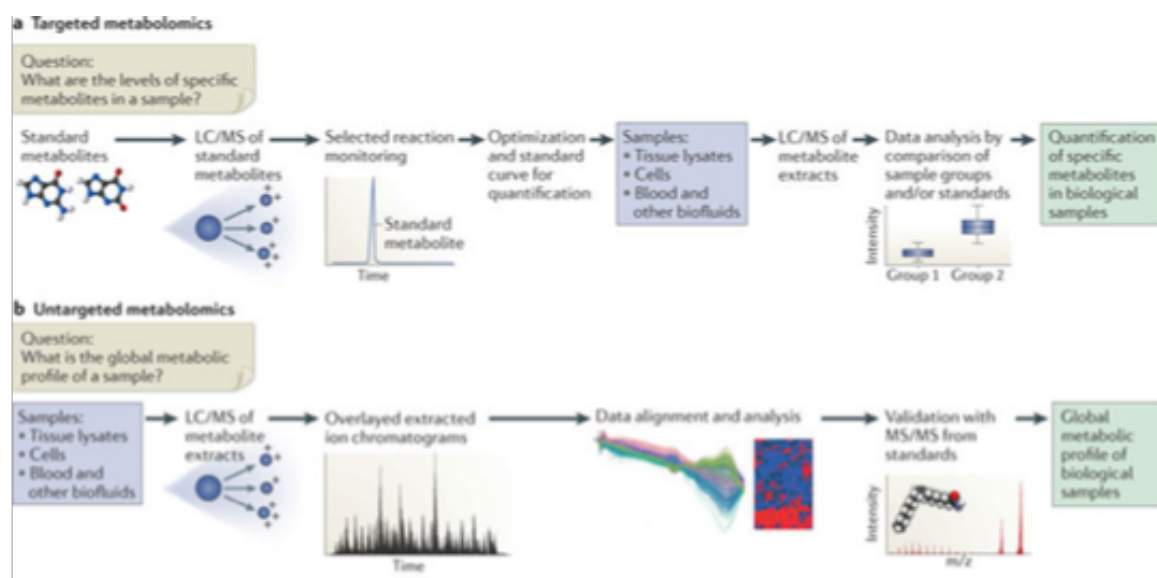


Figure 6 Metabolomics science. Major themes of targeted and untargeted metabolomics investigations are presented. Reprinted from [60].

Metabolomics science leverages analytical chemistry technology where Ultra high performance Liquid Chromatography (UPLC) coupled to Mass Spectrometry (MS) are the preferred data generating instruments for untargeted experimentation[59]. Using these technologies, samples are separated across an affinity matrix (a bead matrix with a fixed hydrocarbon backbone) and subject to a ramping solvent gradient (UPLC aspect), which at some solvent composition most molecules reach a point of solvency and are liberated from the matrix. Liberated samples acquire a charge through an ionization process (e.g. Electrospray Ionization) and are subsequently detected (MS aspect). Individual resultant data points therefore possess a unique mass-to-charge ratio (m/z) and a retention time (rt), which can be identified in multiple samples and the intensity of each molecule compared. Group-to-group samples can be aggregated and spectral intensity data subject to multidimensional statistical analyses. The most efficient metabolomics software for pairwise and multi-group analysis is XCMS Online[61, 62] which can be applied to multiple lines of research from plant physiology[63], stem cell biology[64], to toxicity studies[65]. Metabolomics has particular promise when applied to infectious parasitic diseases by virtue of the introduction of a novel metabolome unique from the host. A notable example is the detection of differential features of Onchocerciasis[66] with subsequent discovery of a unique small molecule with definitive differential utility[67]. Multiple disease states leverage the power of metabolomics to characterize endogenous metabolites, including diseases caused by viruses.

As obligate intracellular parasites, the salient feature of viral infection is the subversion of host cellular processes to the benefit of the virus and the detriment of the

host. Viral infection and/or active replication lead to disease, which manifests as a disruption of cellular homeostasis. This disruption can be phenotypically detected at the metabolic analyte level. A clear example is illustrated in a non-human primate (NHP) model of human Acquired Immunodeficiency Syndrome (AIDS) where infection and neuropathology caused by Simian Immunodeficiency Virus (SIV) causes changes in multiple cerebral spinal fluid (CSF) lipid species. These lipids can act as biomarkers of AIDS neuropathology. Differential activity of a lipid-modifying enzyme was verified transcriptionally and via ELISA, explaining aberrant CSF lipid profiles and reinforcing the validity of the small molecule finding[68]. Subsequent proteomics studies of the CSF of SIV infected NHPs reveals the increase of biomarker proteins and corollary increase in transcription of said proteins[69].

Blood and its fluid fractions—serum and plasma—present a routinely and easily accessible clinical specimen. Viral infection of cells in the circulatory system such as leukocytes and endothelial cells results in phenotypic changes in the blood, i.e. dynamic levels of serum small molecules. In a murine model of arenaviral disease the dynamics of lipid metabolism, energy substrates including the citric acid cycle, and the presence of unique protein breakdown products where the result of LCMV infection and dynamic over the time course of active replication and immune clearance[70]. Capturing the metabolome through a controlled, time-course experiment of arenaviral disease is important as the animals return to pre-infection serum metabolome status, suggesting that the changes registered in serum small molecules may possess prognostic character. A caveat in attempting to extrapolate too much from these data is the lack of terminal

disease phenotype in the mice as a result of LCMV infection compared to arenaviruses causing VHF in primates.

The clinical utility of metabolomics technology is evidenced in the screening of all newborns for inborn errors of metabolism[71, 72]. In viral disease and biomarker discovery metabolomics reveals populations of small molecules correlating with disease status. With the implementation of the Milwaukee Protocol, rabies—historically a uniformly fatal disease if not immediately vaccinated post-exposure—has seen a handful of patients survive[73, 74] but with failures as well[75, 76]. A metabolomics investigation of the CSF from human rabies patients revealed metabolites capable of differentiating the stages of rabies, including differentiating survivors from those whom perish[77]. Extrapolating these findings and applying a metabolomics investigation to a murine model of rabies led to the identification of host pathways activated by the virus resulting in protection of the virus from host inflammatory responses[78]. The findings suggest that the mechanism of the MP may be slowing/inhibiting brain metabolism, leading to the maintenance of brain homeostasis and mitigating rabies disease mortality. Metabolomics studies can therefore identify pathways targetable for pharmacological therapeutics to mitigate the impact of viral disease[78]. Additional utility in viral metabolomics is found in the identification of features with the ability to delineate stages of viral disease.

Dengue fever, a disease akin to LF with respect to annual cases of VHF, offers a rigorous example of the application of metabolomics to viral disease. There are 4 Dengue virus serotypes[79] and disease is noted to be more serious upon secondary/reinfection with a serotype different from the primary virus[79, 80]. Metabolomics interrogation of

DF has revealed small molecule and lipid signatures correlating with disease status[81]. Importantly, this study employed sampling of patients through the course of disease. The serial sampling strategy procured data of metabolite populations with the ability to differentiate patient groups as a function of the time-status continuum, potentiating the identification of prognostic biomarkers[81]. Further, small molecule biomarkers differentiate the transition of DF to the severe forms Dengue hemorrhagic fever (DHF) or Dengue shock syndrome (DSS)[82]. The utility of global biomolecule characterization can be illustrated in the noteworthy identification of elevated corticosterone in mice infected with rabies[78] and the elevation of cortisol in early human dengue cases[81]. Both steroidal hormones are known to mitigate inflammatory responses, the presence of which may result in a protective effect against active immune (over)-responses. Alternatively, virally mediated expression of anti-inflammatory molecules may be a mechanism to protect the virus from host antagonism. In this view, acquiring and multiplexing data such as comprehensive serum metabolomes of multiple viral diseases—particularly viral hemorrhagic fever—will generate information leading to refinement of the understanding of viral pathogenesis. Further, conserved pathways dysregulated by diverse viral species may present as targets for broad-spectrum therapeutics. Finally, metabolites with biomarker character will be discovered and be particularly useful as markers correlating with the trajectory of disease progression.

Proteomics and Enveloped Viral Particles

Proteomics is the systems biology discipline seeking to isolate, detect, characterize, and identify all proteins from a given sample. As with metabolomics, proteomics leverages modern analytical chemistry technology where LCMS is employed to

characterize hundreds-to-thousands of peptides from a given matrix[83]. A key advantage of proteomics compared to other Systems level research is material readout: whereas genomics or transcriptomics yields *potential*, the proteome reveals *phenotype*. Proteomics has been applied to a multitude of biological applications including fundamental scientific and clinical research[84]. The acquisition of host proteins in extracellular viral particles has historically been well characterized[85] and is one example of the utility of global protein characterization[86]. Global proteome characterization of extra-viral constituents may therefore reveal cellular interacting proteins pertinent to viral replication.

As a Health and Human Services (HHS) Select Agent and National Institutes of Allergy and Infectious Disease (NIAID) Category A pathogen LV requires Biosafety Level 4 (BSL-4) handling[87]. Lassa virus-like particles (VLPs) are infectious, non-replicative biophysical entities akin to native virus but devoid of RNA genome and L, alleviating the need for BSL-4 containment[18]. Lassa VLPs are generated by ectopic expression of LV GPC, NP and Z. The resulting VLPs possess all the major antigenic determinants of LV, induce specific immune responses and visually resemble native virions[18]. With the antigenic properties of Lassa VLPs these entities present as a vaccine candidate[18, 88]. Characterizing the totality of the protein content of VLPs would be valuable information to review prior testing vaccine potential to mitigate possible off-target effects[88].

Proteomic analysis of Lassa VLPs offers the potential to ascertain comprehensive protein profiles of the VLPs. For agents of VHF both LV and Ebola viruses have been shown to acquire multiple host proteins such as TSG101[89, 90] or HSP90AA1[91]

which can trace the replicative events of viral synthesis (**Figure 4**). Some host protein incorporations are likely incidental and expected such as cytoskeletal proteins given the nature of viral biogenesis and egress. Alternatively, the role of virally incorporated host proteins may be specific where the native function of the protein is necessary for viral replication. This can be shown with genetic knockdown in parental cells with adverse effects on viral replication and/or infectivity[91]. The phenomena of host protein incorporation into extracellular viral particles are well appreciated for multiple viral species, dating back to early studies with HIV[85]. Expanding proteomics data to multiple enveloped viral particles will reveal host-virus interactions conserved across multiple viral families.

Robust LCMS examination of multiple viral families has been performed but not compiled into a single database focusing on *host protein* representation. Aggregation of available viral proteomics data serves as a nucleation for the development of a comprehensive understanding of pan-viral proteomes. This type of metaanalysis will yield a dataset of virion proteomes focused on the constitution of host proteins in extracellular enveloped viral particles. Patterns spanning viral species may present as targets for antiviral inhibition or useful for investigating fundamental molecular biology of viral synthesis.

The systems biology disciplines of metabolomics and proteomics offer new views of an old disease—LF—and the virus that causes it, LV.

Specific Aims

The purpose of the forthcoming research is to characterize Lassa fever and Lassa virus using Systems Biology approaches. Complimentary technologies will assess the

small molecule profiles representative of the time course of LF and proteome of Lassa virus-like particles. Lassa fever has an annual incidence of 300,000 to 500,000 cases with a case fatality rate of 30% for acute disease and is one of a handful of viral hemorrhagic fevers originating in Africa to be exported through globalization. The disease presents primarily as a nondescript febrile illness where acutely ill patients can advance to fulminant, terminal disease in rapid succession. Where modern technologies have proven invaluable in the diagnosis of Lassa fever there remains a pressing need for differential clinical characteristics and specific treatment options.

The known pathology of Lassa fever exhibits disruption of a handful of clinically significant and well-characterized metabolic markers. However no study has examined the global characterization of all small molecules in Lassa fever patients. Given the cellular tropism of Lassa virus coupled with endothelial dysregulation it is hypothesized that virally mediated perturbations to cellular processes at the metabolite level would be observed with significant difference through the course of disease. Analytes found significantly dysregulated at specific stages of disease are candidate biomarkers indicative of disease trajectory.

As Lassa fever disease progresses no approved, specific treatment options are available beyond supportive therapies, hastening the need for novel antiviral therapeutics. Most antivirals target specific viral proteins or are non-coding nucleotide analogues requiring early administration. Some antivirals target host factors pertinent to a specific step in the viral replication cycle. These targeted interactions are mediated through virus and host proteins. Therefore, a complimentary Systems-level discovery to host-pathogen interactions is characterizing the totality of the virion proteome through a proteomics

analysis of extracellular viral particles. Multiple studies assessing virion proteomes have been published which, when synergized with our primary analyses of Lassa virus-like particles reveal patterns of host protein incorporation into enveloped viral particles and propose targetable interactions for antiviral interventions.

Therefore the studies detailed below utilize Systems Biology methodologies to glean information from human clinical samples and Lassa virus-like particles. The issues raised above are addressed by querying two main points. First, is the endemic serum metabolome of Lassa fever differential in a patient population triaged to a Lassa fever treatment ward? Metabolomics performed on clinical specimens has generated biomarkers for infectious diseases such as Onchocerciasis and Dengue fever. It is hypothesized that the phenotypes of acute and terminal Lassa fever sera are significantly dysregulated given the severity of Lassa fever disease. Second, does the virion proteome reveal host-protein incorporation patterns into lipid viral particles? Evidence suggests the presence of multiple host proteins—including those pertinent to Arenaviral biogenesis—may incorporate into the viral particle. Aggregating and synergizing this information in a meta-analysis compiled from enveloped viruses will reveal conserved but unappreciated patterns of host protein incorporation into viral particles. These issues are addressed by two specific aims:

Aim 1: Characterize the serum metabolomes of Sierra Leonean Lassa fever patients across disease status for features with biomarker property.

Aim 2: Analyze the total protein profile of Lassa virus-like particles and enveloped viruses with primary proteomics data synergized with a proteomics meta-analysis.

MATERIALS AND METHODS

Clinical sample derivation

Suspected LF patients or those reporting to Kenema Government Hospital (KGH), Kenema, Sierra Leone with febrile illness were eligible to participate in these studies. Blood from human volunteers was collected as previously described[44]. Briefly, small blood volumes (typically five milliliters [mL]), for serum separation were collected from study subjects with consent from the attending physician. Blood from patients presenting to KGH with febrile illness that met preclinical criteria of suspected LF patients. Patient samples received a coded designation and were collected in serum vacutainer tubes and allowed to coagulate for 20 minutes at room temperature. Serum was separated from coagulated blood by centrifugation (200 x g, 20 minutes at room temperature). The serum fraction was collected for further metabolite processing with aliquots stored in cryovials at -20C.

Preparation of small molecules from patient sera

All sera collected at the Lassa Fever Ward (LFW) KGH were treated as if containing live Lassa Virus. Sera samples were prepared via a validated metabolomics preparation method utilizing ice-cold methanol (CHROMASOLV, Sigma) for extraction[59, 92]. Separated serum samples were depleted of protein by addition to one part sera (100 μ L) 4 parts ice-cold Methanol (400 μ L), the mixture was vortexed vigorously for 10 seconds, and incubated 1 hour at -20C followed by centrifugation at 14,000 x g, 15 minutes, 4C. The supernatant was collected and transferred to a new,

sterile vial and dried under vacuum (SpeedVac). The resultant small-molecule containing pellets were stored in desiccated, sealed containers and shipped to Tulane University. Small molecule containing pellets were dissolved in a solution of 95:5 water:acetonitrile (TJ Baker; CHROMASOLV, Sigma) transferred to autosampler vials, and held at -20C or 4C immediately prior to analysis[70]. All reagents utilized were of HPLC grade.

Liquid Chromatography Mass Spectrometry for small molecule detection

Detection of metabolites was performed via HPLC separation with ESI-MS (electrospray mass spectrometry) detection. HPLC was performed with an aqueous normal-phase, hydrophilic interaction chromatography (ANP/HILIC) HPLC column: an Cogent Diamond Hydride Type-C column with 4 μm particles and dimensions of 150 mm length and 2.1 mm diameter was used with an Agilent 1290 HPLC system. Two identical Diamond Hydride columns were connected in series to obtain better separations. The column were maintained at 60 °C with a flow rate of 900 $\mu\text{L}/\text{min}$. Chromatography was as follows: solvent consisted of H_2O with 0.1% (v/v) formic acid for channel “A” and acetonitrile with 0.1% formic acid for channel “B”. Following column equilibration at 98% B, the sample was injected via autosampler, and the column was flushed for 2.0 min to waste. From 2.0 min to 14.5 min, the gradient was linearly ramped from 98% to 65% B. From 14.5 min to 16.0 min, the gradient was ramped from 65% to 25% B. From 14.5 to 18.0 min the column was held at 25% B, and from 18.0 to 18.2 minutes the gradient was ramped from 25% to 98% B. From 18.2 to 20.0 minutes the column was re-equilibrated with 98% B. This method was optimized based upon meta-sample consisting of an equal-volume mix of all 50 samples.

The mass spectrometer used as an Agilent 6538 Q-TOF with dual-ESI source: resolution is approximately 20,000 and accuracy is 1 ppm. Source parameters were as follows: drying gas 12 L/min, nebulizer 60 psi, capillary voltage 3500V, capillary exit 100V. Spectra were collected in positive mode from 50 to 1700 m/z at a rate of 1 Hz.

Data analysis

Raw spectral data in .d format were uploaded to XCMS Online (Version 1.0.42) and processed as pairwise comparisons using parameters optimized for data acquired with UPLC on an Agilent 6538 MS[61]. When putative identifications were not returned, manual screening for appropriate spectral character (above signal:noise), conserved retention time, and valid ionization profiles were performed manually searching METLIN[93], KEGG[94], and Human Metabolome Databases[95] where appropriate. For computational analyses including logistic regression, cluster analysis, value plotting, statistical comparison, and machine learning raw spectral intensity values were extrapolated from XCMS output files (.xlsx or .csv) and new files created (.csv). All statistical analyses were carried out using the R statistical software package[96], Microsoft Excel, or Graphpad Prism[97]. Raw mass spectral intensity values and a unique identifier for specific spectral features were extrapolated from XCMS output and compiled into .csv files for machine learning analysis with predefined outcome. Machine learning algorithms built into the R package were utilized with outputs quantifying the sensitivity, specificity, accuracy, and/or receiver operating character computed depending on diagnostic (binary) or prognostic (multi-outcome) analysis. The Random Forest algorithm was employed for all analysis reported[98]. The Random Forest algorithm was set to select features through permutation of the data set yielding the strongest indicators

of the input features. The datasets were run with 10-fold cross validation ensuring that ranked output features were selected on importance after predefined, multiple rounds of random training and testing. FacomineR and R were utilized to generate PCA's[99].

Mammalian Cell Culture and Antibodies

The Human Endothelial Kidney 293 cell line (HEK293/T-17,) African Green Monkey Kidney cell (Vero-E6), and Madin Darby Canine Kidney cell (MDCK) were obtained from the American Type Culture Collection. Cells were cultivated in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) amended with 10% heat inactivated Fetal Bovine Serum, 1X Non-Essential Amino Acids (Invitrogen), and 1X L-Glutamine (Invitrogen) maintained in a humidified tissue culture incubator at 37°C, 5% CO₂.

Polyclonal goat antibodies raised against Lassa Josiah Glycoprotein Complex (GenBank: HQ688672.1), Nucleoprotein (GenBank: HQ688673.1) and Ring Finger Protein Z (GenBank: HQ688674.1), were obtained from AutoImmune Technologies and used at a working concentration of 1 µg/mL. Antibodies raised in rabbits against human ribosomal proteins were purchased from Thermo reacting to ribosomal protein S3 (PA5-17214), RPL4 (PA5-29453), RPL3 (PA5-27550), S3a (PA5-21632), and RPS8 (PA5-31676) and Cell Signaling S6 ribosomal protein (2217S) were used at the manufacturers recommended concentration for western blot analysis. Fluorescently conjugated secondary antibodies including donkey anti-goat, goat anti-mouse, goat anti-human, and goat anti-mouse were purchased (Licor Technologies) and used at the manufacturers recommendation (1:10,000-20,000) for Western Blot analysis.

Generation of Lassa virus-like particles

Lassa and Ebola virus-like particles (VLPs) were generated as previously described [18]. Briefly, HEK293/T-17 cells were grown to 90% confluence in sterile 10 cm polystyrene tissue culture dishes functionalized with poly-Lysine (50 mg/mL, Sigma). Transient expression vectors were constructed and have been described for the expression of Lassa GPC, NP, and Z [18]. Transfection of 10 µg DNA/dish (5 µg each vector) was mediated with JetPrime (PolyPlus) following the manufacturer's protocol. Virus-Like Particles were generated by cotransfection with Lassa GPC, NP, and Z (LAS VLP). Control material was generated by transfecting equal amounts of empty pcDNA3.1 + intA (VEC) and preparing materials in parallel. Following transfection, cells were incubated for 48 to 72 hours, the supernatant collected and clarified by centrifugation at 1000 x g, 4°C, 15 minutes. Clarified supernatant was centrifuged at 15,000 x g, 4°C, 1 hour (SW28 rotor) and the supernatant decanted. Pelleted material was resuspended in sterile Phosphate Buffered Saline (PBS) amended with protease inhibitor cocktail (Sigma) at a volume of 250 µL/dish and incubated at 4°C for a minimum of 30 minutes to 12 hours. Suspended material was overlaid on a discontinuous sucrose gradient consisting of 20, 30, 40, and 60% sucrose (w/v) and subject to ultracentrifugation at 50,000 RPM, 4°C, 2 hours (SW60-Ti rotor). Sucrose layers were aspirated and the pellet was washed once in PBS. Concentrated materials were resuspended in aqueous buffer compatible for subsequent analysis.

Purifying Lassa virus-like particles, protein isolation, and LCMS analysis

Protein precipitates were each dissolved in 50 µl 25 mM ammonium bicarbonate in preparation for trypsin digestion. To the two membrane samples were added 45 µl

Laemmli buffer with 2-mercaptoethanol. The two membrane samples were allowed to sit at room temperature for at least 2 hours, then all spun and the supernatants were assayed for protein content. Samples were briefly run on 1-D gel electrophoresis at 100 V (half of usual running voltage) for 7 minutes allowing prestained protein standards to clear the well, indicating protein entered the SDS-PAGE matrix. The SDS-PAGE gel was stained overnight with LCMS-grade Coomassie. Three gel plugs each from samples staining positively for the presence of protein were cut in a diagonal pattern and pooled.

Typically plugs were prepared for MS analysis as follows. Fifty μL 25 mM ammonium Bicarbonate (ABC) was added to gel plugs and incubated at room temperature (RT) for 5 minutes. Discard solution and add 50 μL Acetonitrile (ACN), incubate at RT for 5 minutes. Discard solution and add 40 μL DTT, heat to 60°C and incubate for 10 minutes. Let cool 15 minutes. Discard solution and add 30 μL 100 mM iodoacetamide (IA) in 25 mM ABC. Incubate for 35 minutes at RT, discard solution and wash two times with ABC and ACN as previous. Prepare trypsin solution immediately prior to addition. Dissolve LCMS-grade trypsin in 0.3% formic acid to give a 0.1 $\mu\text{g}/\mu\text{L}$ solution and dilute to 5 ng/ μL in 25 mM ABC. Add 15 μL of trypsin solution, incubate for 10 minutes. Add 10 μL ABC and incubate for 4 hours at 37°C. Add 15 μL 0.3% formic acid. Protein identification from excised 1D PAGE gel plugs is performed by manual injection and using an LC gradient of 60- 90 minutes on a Thermo Scientific Easy nano Liquid Chromatography 1000 coupled to a Triple Quadrupole MS equipped with a Thermo Flex Flex Nanospray source. Raw LCMS files were analyzed with the Scaffold Proteomics software package[100].

Western blot

Concentrated and purified Lassa VLP and vector control samples in sterile PBS where dissociated in NuPAGE LDS loading buffer (LB) and sample reducing agent (RA) (Thermo NP0004 & NP0007) at a ratio of 13:6:1 (μL) sample:LB:RA and proteins denatured for 10 minutes at 65°C. Samples where centrifuged briefly, loaded into individual wells of 10% NuPAGE Bis-Tris gels (Thermo NP0302) with pre-stained (NOVEX) and unstained protein standards (Magic Mark). Proteins where separated across gel matrix by electrophoresis at 200V for 40 minutes in the presence of antioxidant (Thermo NP0005). Proteins where transferred to nitrocellulose membranes via low-voltage electrophoresis in NuPAGE transfer buffer (composed with an additional 100 mL methanol) (Thermo NP0006) following manufacturers operation instructions (Thermo EI0002). Membranes where blocked with western blot buffer (WBB) composed of 5% non-fat dry milk, 1% bovine serum albumin, 0.05% Tween-20, and 0.1% thymersol in 1X PBS, pH 7.4 for 1 hour, RT. Antibodies reacting to LV GPC, NP, and Z (above) where diluted to 1 $\mu\text{g}/\text{mL}$ in WBB and blots stained either 1 hour, RT or 4°C overnight, washed 2 times in western blot wash buffer (WBWB) composed of 0.05% Tween-20 in 1X PBS, pH 7.4. Fluorescently labeled secondary antibodies reacting to the isotype of the primary antibody (above) where diluted in WBB and incubated as for primary and blots washed 3 times in WBWB prior to processing. Blot images where captured with a Licor fluorescent imaging system using standard presents for 680 (red) and 800 (green) nm wavelength channels.

Generation of the Virion Proteome Database

The NCBI PubMed database was queried for combinations of Title/Key Word/Abstract containing “Proteomics, Proteome, Virus, or Virion.” Primary research articles were manually filtered for those purifying and characterizing extracellular enveloped viral particles. Initially the dataset consisted of proteins of host origin detected in extracellular (purified) viral particles. Positive identifications were assumed accurate when a minimum of 2 unique peptide spectra were reported. Secondly, data were sourced from primary research articles characterizing protein-protein interaction studies utilizing viral proteins as bait to catch host factors where the findings explicitly corroborate/include proteins detected in extracellular virus. Interaction studies were included when host protein-virus interaction was confirmed molecularly (typically via western blot analysis) or spectrally. A dataset was compiled from identified host proteins recorded as NCBI Gene Symbol, NCBI Gene Number, UniProt Recommended Name, and UniProt identifier.

Primary search strategy starts w/ (24) papers describing total viral proteomes; secondary strategy augments dataset with primary research data (articles) where proteins identified in the primary search are found in or affiliated with viral proteins.

Information is compiled into a single master dataset and presented with multiple options including open access scripts plotting an interactive world cloud and chord plot, a “venn compare,” and the ability list protein by virus or virus by protein.

Taxonomic Tree Generation

Taxonomic identifiers for previously characterized extracellular virions were obtained from the NCBI Virus Taxonomy Browser for the following viruses: Ebola virus

- Mayinga, Zaire, 1976 (Taxonomy ID: 128952), Marburg virus - Musoke, Kenya, 1980 (Taxonomy ID: 33727), Human immunodeficiency virus 1 (Taxonomy ID: 11676), Moloney murine leukemia virus (Taxonomy ID: 11801), Epstein-Barr virus (strain B95-8) (Taxonomy ID: 10377), Herpes simplex virus type 1 (Taxonomy ID: 10298), Pseudorabies virus (Taxonomy ID: 10345), Human cytomegalovirus (Taxonomy ID: 10359), Kaposi's sarcoma-associated herpesvirus - Human herpesvirus 8 (Taxonomy ID: 37296), Murine herpesvirus 68 (Taxonomy ID: 33708), Influenza A virus (A/NWS/1933(H1N1)) (Taxonomy ID: 382819), Hepatitis C virus (Taxonomy ID: 11103), Vesicular stomatitis Indiana virus (Taxonomy ID: 11277), Rift Valley fever virus (Taxonomy ID: 11588), Newcastle disease virus (Taxonomy ID: 11176), Vaccinia virus (Taxonomy ID: 10245), Porcine reproductive and respiratory syndrome virus (Taxonomy ID: 28344), and Human respiratory syncytial virus A2 (Taxonomy ID: 11259). Taxonomic Identifiers were uploaded to phyloT and visualized through the Interactive Tree of Life using default settings[101].

GO Annotation of the Virion Proteome Database

A list of all characterized host proteins was compiled as UniProt Accession Entries. The list was then applied as a QuickGO filter[102]; dataset results were analyzed with Microsoft Excel.

Compounds

Withaferin A was purchased from Cayman Chemical (Item № 11352), TMN 355 (Cat. № 4152) and GP3466B maleate (Cat. № 2966) were purchased from Torcis Bioscience, and T56 (SML1305) and Pifithrin- μ (P0122) were purchased from Sigma

Aldrich. Stock concentrations of small molecules were rehydrated in molecular biology grade Dimethyl sulfoxide (DMSO, Sigma D8418) at 10 or 100 mM and stored at -20°C.

Screening candidate compounds

A HIV pseudoparticle (PP) with a luciferase reporter was utilized to screen candidate compounds as novel inhibitors of viral synthesis[103]. The NL4-3 Human Immunodeficiency Virus (HIV) plasmid deficient for the HIV Envelope protein (*Env*-) containing a Firefly luciferase gene (*luc*)[104] was obtained from the AIDS reagent source. The pCAGGS-G-Kan expression vector encoding the Vesicular Stomatitis Virus glycoprotein (VSV-G) was obtained from Kerafast (Cat. # EH1017). Inhibitors were screened with the NL4-3/VSV-G PP generated by cotransfection of HEK293/T-17 cells with NL4-3 and VSV-G plasmids in complete media. Four hours post-transfection the media is replaced with 200 µL media containing inhibitor and incubated for 72 hours. PP containing supernatants are collected and clarified by centrifugation at 1000 x g, 4°C, 15 minutes, aliquoted and stored at -80°C. Inhibition was assessed by infecting uniform volume of viral supernatant to MDCKs at 10,000 cells/well in black walled, clear bottom 96-well trays. After a 72-hour incubation 100 uL/well luciferin containing Beta-Glo (B-Glo, Promega Cat #E4780) is added, plates incubated 5 minutes at RT, and luminescence read with a Biotek Synergy 2 plate reader.

Toxicity

Cellular toxicity was assessed by Alamar Blue (AB, Invitrogen Cat # DAL1100) metabolic assay[105]. Over night cultures of HEK293/T-17, Vero-E6, or MDCK cells plated at 10,000 cells/well in black walled, clear bottom 96-well trays were established. The media is replaced with 150 µL media containing inhibitor and incubated. Following a

72 hour incubation 100 μ L complete media with 5% AB is added per well yielding 1% final concentration and incubated at 37°C, 5% CO₂. Following a 4-hour incubation fluorescence is read using an excitation wavelength of 540–570 nm (peak excitation is 570 nm) and emission at 580–610 nm (peak emission is 585 nm) with a Biotek Synergy 2 plate reader.

Viruses

The recommended strain for antiviral testing of Influenza virus H3N2 (ATCC VR-1679) and Dengue virus serotype 2 (DEN2, ATCC VR-1584) were obtained from the American Type Culture Collection. Viral stocks were generated by infection of H3N2 in 80-90% confluent MDCKs and DEN2 in 80-90% confluent Vero-E6 in minimal volume serum free Eagles Modified Essential Media (sfEMEM) (1 mL/T75) for 1 hour rocking every 15 minutes. Following infection EMEM plus 2% FBS was added and incubated 2 days for H3N2 and 6 days DEN2, monitoring for cytopathic effects (CPE). Once profuse CPE was apparent culture supernatants were collected and clarified by centrifugation at 1,000 x g, 15 minutes, at 4°C. Clarified viral supernatants were passed through a 0.45 μ m filter, and stored at -80°C in 500 μ L aliquots.

Viral titration via plaque assay

A 10-fold dilution series of H3N2 or DEN2 stocks were established by diluting 100 μ L stock into 900 μ L sfEMEM (10^{-1}) and 100 μ L into 900 μ L dilutions repeated from 10^{-1} down to 10^{-8} in sfEMEM. Monolayers of MDCK or Vero-E6 cells at 90% confluence in 6 or 12-well tissue culture trays were washed 2 times with warm sterile Phosphate Buffered Saline pH. 7.4 (PBS, Invitrogen). In duplicate (6-well) or triplicate (12-well) cells were infected with viral dilutions in minimal volume (250 μ L/well in 6-

well format or 125 μ L/well in 12-well format) and incubated at 34°C or 37°C for H3N2 or DEN2 for 1 hour, rocking every 15 minutes. For H3N2 a 2X overlay media composed of 2X MEM (10X MEM, Invitrogen), 1.2% BSA (7.5% in PBS, Invitrogen), 2X L-Glutamine (100 mM, Invitrogen), 2X Sodium Pyruvate (100X, Invitrogen), 2X Non Essential Amino Acids (NEAA, 100X, Invitrogen), and 1 μ g/mL TPCK-treated Trypsin (Worthington Biochemical) is mixed with equal volume 2X Avicel (2.4 g/100 mL H₂O). For DEN2 2X MEM with 4% FBS is mixed with equal volume 2X Avicel. Following the infection incubation 2 mL or 1 mL/well Overlay media are added per 6 or 12-well tray. Trays are incubated at 34°C for 2 to 3 days for H3N2 and 37°C for 6 days for DEN2. Assays are terminated by adding equal volume PBS and gently removing overlays, and monolayers fixed by addition of equal volume 10% Formalin (Sigma, HT501128) for at least 30 minutes. Plaques are visualized by staining with 1X crystal violet.

p24 ELISA

An Enzyme Linked Immunosorbent Assay (ELISA) detecting to the HIV p24 protein located in the lentiviral capsid was executed according to the manufacturers recommendation (Clonetechn Cat # 632200). Briefly, 100 μ L PP supernatant was processed in duplicate for lysis and anti 24 staining. Following addition of horseradish peroxidase (HRP) conjugated secondary, samples where processed with TMB and stop solution. Plates where immediately read on a Biotek Plate reader at Absorbance 450 (A_{450}).

Pre-incubation

Two hours prior to cotransfection the media of HEK293/T-17 cells was changed to media plus inhibitor or equal concentration vehicle control. Cotransfection was

executed in the presence of inhibitor and media exchanged 2 hours after with fresh media plus inhibitor or equal concentration vehicle control and incubated at 37°C, 5% CO₂. Following a 72-hour incubation relative viral quantity was assessed via luminescence as above.

Residual/Carryover toxicity assay

Cotransfection was carried as previous. Four hours post-transfection media was exchanged for fresh media plus 2.5 µM WA, equal concentration DMSO, or complete media. Un-transfected HEK293/T-17 cells were cultivated with 2.5 µM WA in parallel to generate spent media (SM). All plates were incubated at 37°C, 5% CO₂. Following a 72-hour incubation, PP supernatants were collected. Half of each PP was concentrated via centrifugation at 21,000 x g, 30 minutes, 4°C. The supernatants were aspirated and resuspended in equal volume fresh complete media (FM) or SM. Infection of MDCK cells was performed as previous, incubated for 72 hours and assessed via luminescence as above.

Plaque inhibition assay

For H3N2 virus 1 x 10⁶ MDCK cells per well were plated in 6-well trays and allowed to attach overnight. Cells were washed 2 times with warm sfMEM and 50 plaque forming units (PFU) applied in 250 µL sfMEM and plates incubated 34°C, 5% CO₂ with gentle rocking every 15 minutes for 1 hour. During the incubation MEM-Avicel overlay media is prepared and inhibitor added and a dilution series established. At the end of the incubation 2 mL per well overlay media is added, with biological replicates and no inhibitor controls. Plates are incubated at 34°C, 5% CO₂ and processed as described previously for PFU enumeration.

For DEN2 virus 5×10^5 Vero-E6 cells per well were plated in 12-well trays and allowed to attach overnight. Cells were washed 2 times with warm PBS and 25 plaque forming units (PFU) applied in 125 μ L sfEMEM and plates incubated 37°C , 5% CO_2 with gentle rocking every 15 minutes for 1 hour. Media is prepared as described above. At the end of the incubation 1 mL per well overlay media is added, with biological replicates and no inhibitor controls. Plates are incubated at 37°C , 5% CO_2 and processed as described previously for PFU enumeration.

Plaque-forming unit reduction assay

For DEN2 virus 5×10^5 Vero-E6 cells per well were plated in 12-well trays and allowed to attach overnight. Cells were washed 2 times with warm PBS and 25 PFU DEN2 applied in 125 μ L sfEMEM and plates incubated at 37°C , 5% CO_2 with gentle rocking every 15 minutes for 1 hour. During the incubation a 2-fold dilution series of complete media plus inhibitor was established. Vehicle controls were also established with highest concentration vehicle utilized in assay in complete media. After incubation, 1 mL media per well was added and plates incubated at 37°C , 5% CO_2 . Following a 6-day incubation, DEN2 supernatants are collected, clarified of cellular debris by centrifugation at 1,000 x g, 15 minutes, 4°C . Control material was quantified by plaque assay as described to determine countable dilution. Subsequently supernatants from all conditions were diluted down to 10^{-3} and enumerated via plaque assay as described above.

RESULTS

Aim 1

The non-descript presentation of febrile illness

The overall process, goals, and application of a clinically derived serum metabolomics investigation are presented in **Figure 7**. A panel of 50 serum samples of patients presenting to Kenema Government Hospital (KGH), Sierra Leone, Africa and triaged to the Lassa Fever Ward (LFW) with febrile illness were analyzed with LCMS. Specimens approximate an acute disease-to-convalescence time-spectrum of the pathogenesis of LF, encompassing a comprehensive immunological status of LF patients (Tables 1 & 2). The diagnostic status of the samples used in for this study and aggregate positive LF sampled during the time period these samples where generated reflect a common trend with positive diagnostics appearing at 7 days since symptom onset (**Figure 8**).

Table 1

Gnumber	Positive	Outcome	Ribavirin	Days Since Onset	Gender	Age
G-2585-1	Ag	Not admitted	N	3	F	38
G-2319-1	Ag & IgG	Discharged	Y	4	F	24
G-2557-1	Ag & IgG	Died	Y	5	F	38
G-2740-1	Ag & IgM	Discharged	Y	8	F	20
G-2349-1	Ag & PCR	Died	N	NA	F	20
G-2612-1	Ag & PCR	Died	Y	14	M	24
G-2770-1	Ag & PCR	Died	N	NA	F	30
G-2511-1	Ag, IgM, &	Discharged	Y	6	F	22
G-2554-1	Ag, IgM, &	Discharged	Y	10	F	39
G-2587-1	Ag, IgM, &	Discharged	Y	4	M	6
G-2615-2	Ag, IgM, &	Died	Y	17	M	32
G-2429-1	IgG	Not admitted	N	4	M	20
G-2432-1	IgG	Not admitted	N	2	F	21
G-2433-1	IgG	Not admitted	N	24	M	32
G-2446-1	IgG	Not admitted	N	1	M	65
G-2466-1	IgG	Discharged	N	11	M	56
G-2617-1	IgG	Not admitted	N	5	F	28
G-2318-1	IgG & PCR	Not admitted	N	NA	M	4
G-2444-1	IgM	Discharged	Y	9	M	20
G-2541-1	IgM	Not admitted	N	3	M	38
G-2559-1	IgM	Not admitted	N	7	F	20
G-2607-1	IgM	Not admitted	N	3	M	10
G-2619-1	IgM	Not admitted	N	3	M	20
G-2724-1	IgM	Not admitted	N	14	F	25
G-2731-1	IgM	Discharged	Y	6	F	19
G-2755-1	IgM	Not admitted	N	NA	M	43
G-2325-1	IgM & IgG	Not Admitted	N	7	F	4
G-2326-1	IgM & IgG	Not admitted	N	7	F	21
G-2437-1	IgM & IgG	Not admitted	N	11	F	18
G-2552-1	IgM & IgG	Not admitted	N	5	F	35
G-2660-1	IgM & IgG	Not admitted	N	NA	NA	NA
G-2725-1	IgM & IgG	Not admitted	N	30	M	33
G-2733-1	IgM & IgG	Discharged	Y	13	F	19
G-2380-1	IgM & PCR	Not admitted	N	21	F	43
G-2387-1	IgM & PCR	Discharged	Y	14	M	23
G-2424-1	IgM & PCR	Discharged	N	8	F	7
G-2299-1	None	Died	N	14	F	9
G-2308-1	None	Not admitted	N	9	M	30
G-2314-1	None	Not admitted	N	3	F	11
G-2370-1	None	Died	N	7	F	19
G-2402-1	None	Not admitted	N	7	F	40
G-2449-1	None	Not admitted	N	4	M	40
G-2561-1	None	Not admitted	N	7	F	18
G-2726-1	None	Not admitted	N	NA	F	19
G-2746-1	None	Not admitted	N	NA	M	3
G-2423-1	PCR	Discharged	Y	7	M	9
G-2560-1	PCR	Not admitted	N	NA	M	35
G-2614-1 ¹	PCR	Discharged	Y	10	F	43
G-2622-1	PCR	Not admitted	N	7	M	42
G-2727-1	PCR	Not admitted	N	3	F	40

¹ Discharged against medical advice

Clinical data (where available) for the panel illustrate the relative homogeneity of febrile illness (Table 2). Forty-two of the 50 (84%) patient specimens utilized reported a time since symptom onset (in Days) allowing for additional grouping of populations based on time in addition to serostatus. There were 28 female and 21 male patient sera screened (57% female) with gender information unavailable for one sample; mean age for the 49/50 samples is 26.0 years (SD +/- 13.94 years) with age information excluded for one patient. Patients presenting with febrile illness but testing negative by all molecular assays for presence of Lassa virus (Negative) were utilized as endemic non-Lassa controls (Table 2).

Table 2

Characteristic	All Samples	No Lassa	Lassa	p-value
Age (range in years)	26.02 (3-65)	21 (3-40)	26.94 (4-	0.2489
Sex, n (%)				
Male	21 (42.8)	3 (33)	18 (45)	0.7137
Female	28 (57.1)	6 (66)	22 (55)	
Days Since Symptom Onset	8.5 (1-30)	7.28 (3-14)	8.74 (1-30)	0.4165
Mortality, n (%)	7 (14)	2 (4)	5 (10)	0.5949
Symptoms (n Yes,No)				
Headache	46,2	6,1	40,1	0.273
Weakness	7,36	1,4	6,32	1
Dizziness	35,5	6,0	29,5	1
Cough	24,18	1,6	23,12	0.03052
Runny Nose	NA	NA	NA	NA
Conjunctivitis	5,36	1,6	4,30	1
Jaundice	5,35	1,6	4,29	1
Vomiting	31,11	5,2	26,9	1
Diarrhea	14,24	0,4	14,20	0.2759
Rash	1,37	0,6	1,31	1
Deafness	0,39	0,5	0,34	1
Confusion	9,32	2,5	7,27	0.6372
Bleeding	17,25	2,4	15,21	1
Sore Throat	19,22	3,4	16,18	1
Breathing Difficulty	NA	NA	NA	NA
Head Swelling	10,32	2,5	8,27	1
Inflammation	3,34	0,5	3,29	1
Convulsions	3,38	1,6	2,32	0.4386
Vital Means (SD, n)				
Height (cm)	148.19 (33.33, 22)	149.5 (24.75, 2)	148 (34.58,	0.949
Weight (kg)	51.59 (21.70, 21)	55.9 (43.98, 2)	51.208	0.9048
Temperature (degrees C)	37.52 (1.51, 40)	37.26 (0.62, 5)	37.56	0.4567
Systolic BP	102.06 (12.31, 35)	98.40 (19.31, 5)	102.67	0.6543
Diastolic BP	61.71 (10.71, 35)	62.00 (17.89, 5)	61.33	0.9387
Heart Rate	90.03 (13.22, 39)	91.50 (6.19, 4)	89.86	0.6844
Respiratory Rate	25.88 (12.82, 40)	24.00 (4.69,5)	26.14	0.5008
Interventions				
Ribavirin Treatment, n (%)	14, (28)	0 (0)	14 (28)	
Ribavirin Mortality, n (%)	3, (21)			
Received Antimalarials	29,12	4,3	25,9	0.3978
Received Antibiotics	32,9	4,3	28,6	0.1651
Received Other Medications	32,8	4,2	28,6	0.5803
Diagnostic, n (%)				
Ag	11	0	11	
IgM	22	0	22	
IgG	14	0	14	
PCR	17	0	17	

From the sample set chosen, 7 patients perished in KGH (14% total sample group) with a mean terminal time point (time in days since the onset of symptoms) of 11.4 days (SD +/- 5.13 days) with time since symptom onset data excluded for 2 patients. The mean age for patients whom perished is 24.57 years (SD +/- 9.66 year). Fourteen patients in this panel with ≥ 1 positive Lassa diagnostic received Ribavirin and with 21% (3/14) of treated patients perishing (Table 2). Spectral features of protonated, sodiated, and potassiated ($m/z = 245.0852, 267.0775, \& 283.0408$; $rt = 4.30, 4.32, 4.34$, respectively) ribavirin adducts were detected in these samples (not shown).

Figure 7

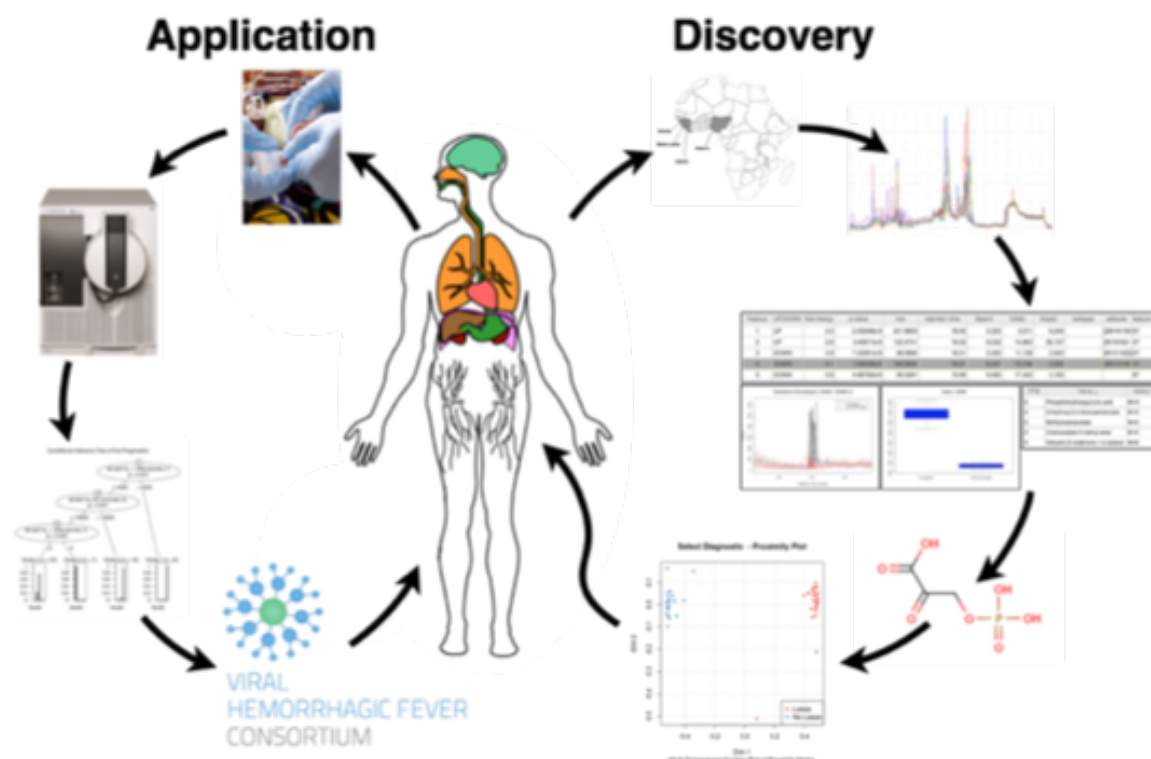


Figure 7 Discovery objectives for untargeted serum metabolomics investigation to characterize and validate small molecule biomarkers with diagnostic or prognostic utility in Lassa fever patients. Analytical chemistry and computational methods will yield small molecule signatures of Lassa fever. These signatures can potentially be translated to the clinical scenario from patients presenting with febrile illness.

Figure 8

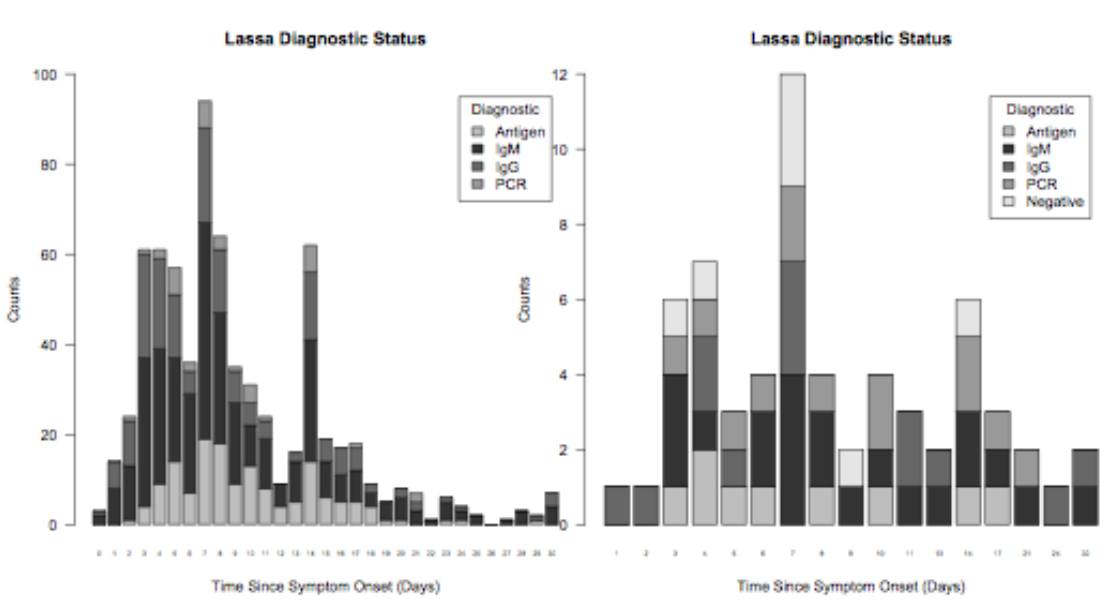


Figure 8 Diagnostics histograms. The histogram on left is a summation of all positive Lassa diagnostics obtained at the LFW during the time the samples analyzed in this investigation were acquired. The histogram on the right is a summation of the diagnostic status of the samples analyzed. In both cases the number of positive diagnostics is highest on Day 7 since symptom onset.

When available clinical data are accounted for in this dataset only one symptom (parameter), cough, is significantly different between Febrile Non Lassa (FNL) and Lassa Positive (LP) samples where LP patients experience cough more frequently than NL ($p=0.03$, Table 2). No other symptoms are significantly different between FNL and LP patients sampled for this study (Table 2). Of LP samples, 11 tested positive for the presence of viral antigen (viral nucleoprotein), 22 tested positive for the presence of anti-Lassa Immunoglobulin M (IgM), and 14 tested positive for anti-Lassa IgG by ELISA. Seventeen samples tested positive for viral genome by RT-PCR. Mean age for all patients was 26.02 years old (range from 3 to 65) where the mean age for FNL was 21.0 and LP was 26.94 ($p=0.2489$). Patient gender was 58% female and 42% male for all patients where there was no difference between gender FNL and LP populations ($p=0.7137$). Mortality rate for all patients was 14% where 4% of FNL and 10% of LP patients perished. Serum samples were drawn upon admittance and all ELISA diagnostic tests for patients whom perished were performed the same day the sample was drawn. Two FNL patients perish <1 day after admittance while 5 LP patients perish with a mean 1.25 days after admittance.

The Sierra Leonean LFW patient serum metabolome is dynamic

Individual samples produced between 3,100 and 6,900 *differential* spectral features compared to FNL patients. For primary analyses samples were grouped based on the patient serostatus. Pairwise principal components analyses (PCA) executed in XCMS analyzing the FNL (Negative) group to the indicated LF positive diagnostic group illustrated differences between Negative and LP *groups* (Table 3 and **Figure 9**). Similarly

pairwise cloud plots illustrate that gross, conserved spectral features are differentially present in Negative and LP samples (**Figure 10**).

Table 3

Pairwise Comparison	Total Aligned Features	Features ($p \leq 0.05$ & fold change ≥ 1.5)
Negative vs. IgM+	3,216	55
Negative vs. IgG+	3,333	3
Negative vs. IgG/IgM+	3,197	17
Negative vs. IgM/PCR+	3,020	215
Negative vs. Ag/IgM/PCR+	3,244	454
Negative vs. Deceased	3,410	360

Figure 9

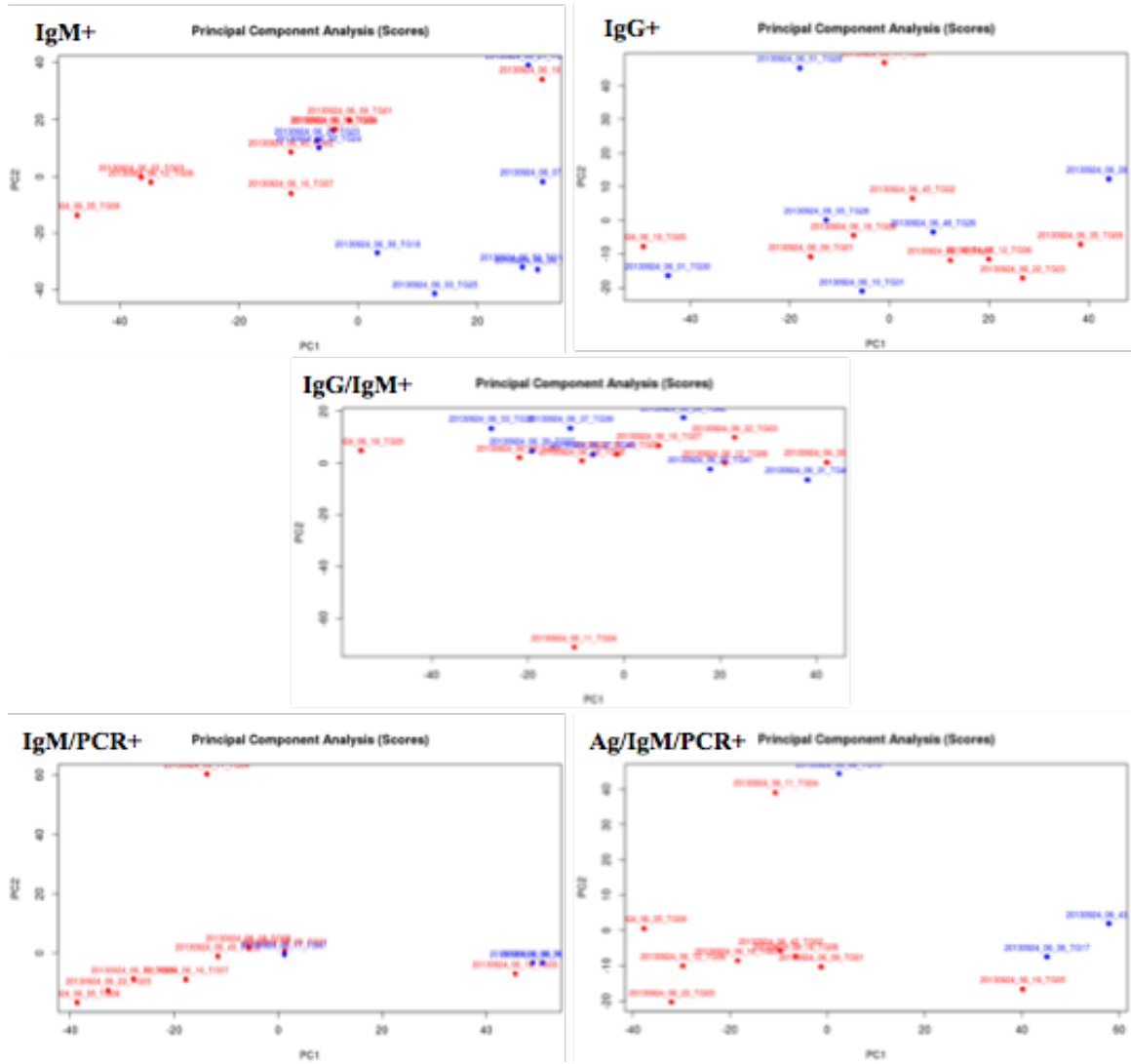


Figure 9 Serostatus PCAs. Principle components analysis of all samples analyzed grouped by the serological Lassa status of the patients. Red data points are from febrile non Lassa patients, blue data points are indicated in the top left of each plot.

Figure 10

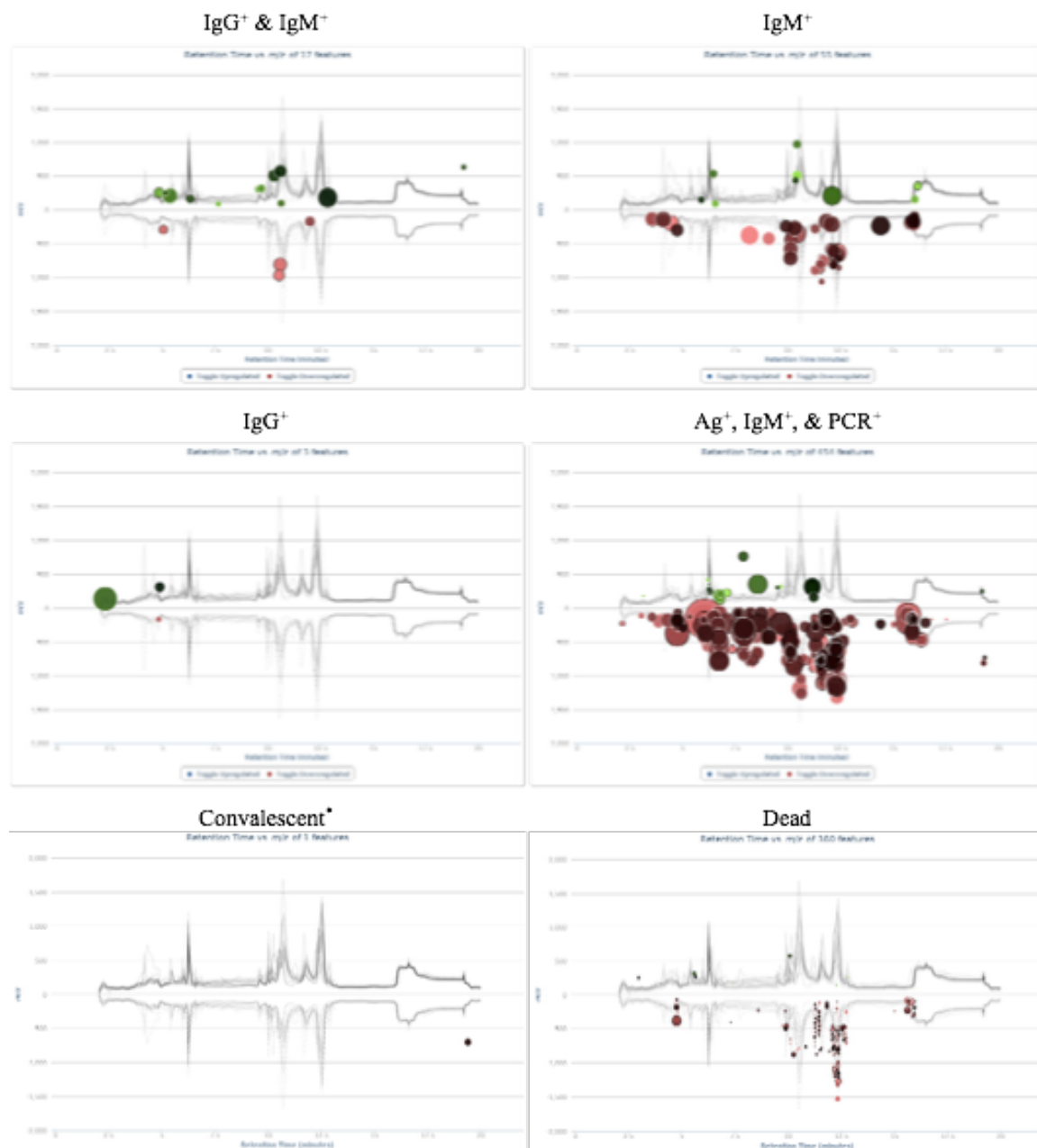


Figure 10 Serum metabolome is dynamic during LASF. Cloudplots (XCMS) illustrate the dynamic flux of the serum metabolome of Lassa fever patients whereby acute and terminal samples exhibit drastic differences while late and convalescent sera are virtually indistinguishable when compared to Non-Lassa control samples at the small molecule level. *In order to generate a cloudplot for convalescent sera graphing parameters were adjusted to plot features of less than 1.5-fold change, as there are no feature differences with $p = 0.05$ and ≥ 1.5 -fold change between Non-Lassa and convalescent sera metabolomes.

Secondarily, all seropositive samples were grouped into Early (Days 1-5), Mid (Days 6-10), and Late (Days 11-30) based on time since symptom onset and pairwise analyses repeated where PCAs offer a mixed grouping (Table 4 and **Figure 11**) and cloud plots reveal almost no significant differences between populations (**Figure 12**). Greater, though not statistically significant dissimilarity is observed when serum samples are grouped according to time since symptom onset. Observing a unique grouping within the Early group resulted in a refined grouping via manual screening of spectral intensities for selected markers shared amongst early samples yielding a metabolite profile hypothesized to represent a patient experiencing acute LF and termed “Acute.” This group observed a dichotomous sorting upon PCA and predominantly shows early humoral immune responses consistent with anticipated adaptive immunity (71% IgM⁺) and the presence of viral antigen. Finally, all samples in the Late grouping whom are discharged from KGH (walk out alive) yields a group of admitted and discharged and non-admitted patient samples representing metabolic profiles of those patients presenting with febrile illness, testing positive for LF via at least one diagnostic and surviving; these samples were assembled to represent convalescent LF and where termed “Survive.”

Table 4

Pairwise Comparison	Total Aligned Features	Features Detected $p \leq 0.05$ & fold change ≥ 1.5
Negative vs. Early	2,938	49
Negative vs. Mid	2,721	55
Negative vs. Late	2,926	2
Negative vs. Acute	3,636	573
Early vs. Dead	3,047	185
Survive vs. Dead	3,528	207
Acute vs. Dead	3,624	224
Survive vs. Acute	3,517	505

Figure 11

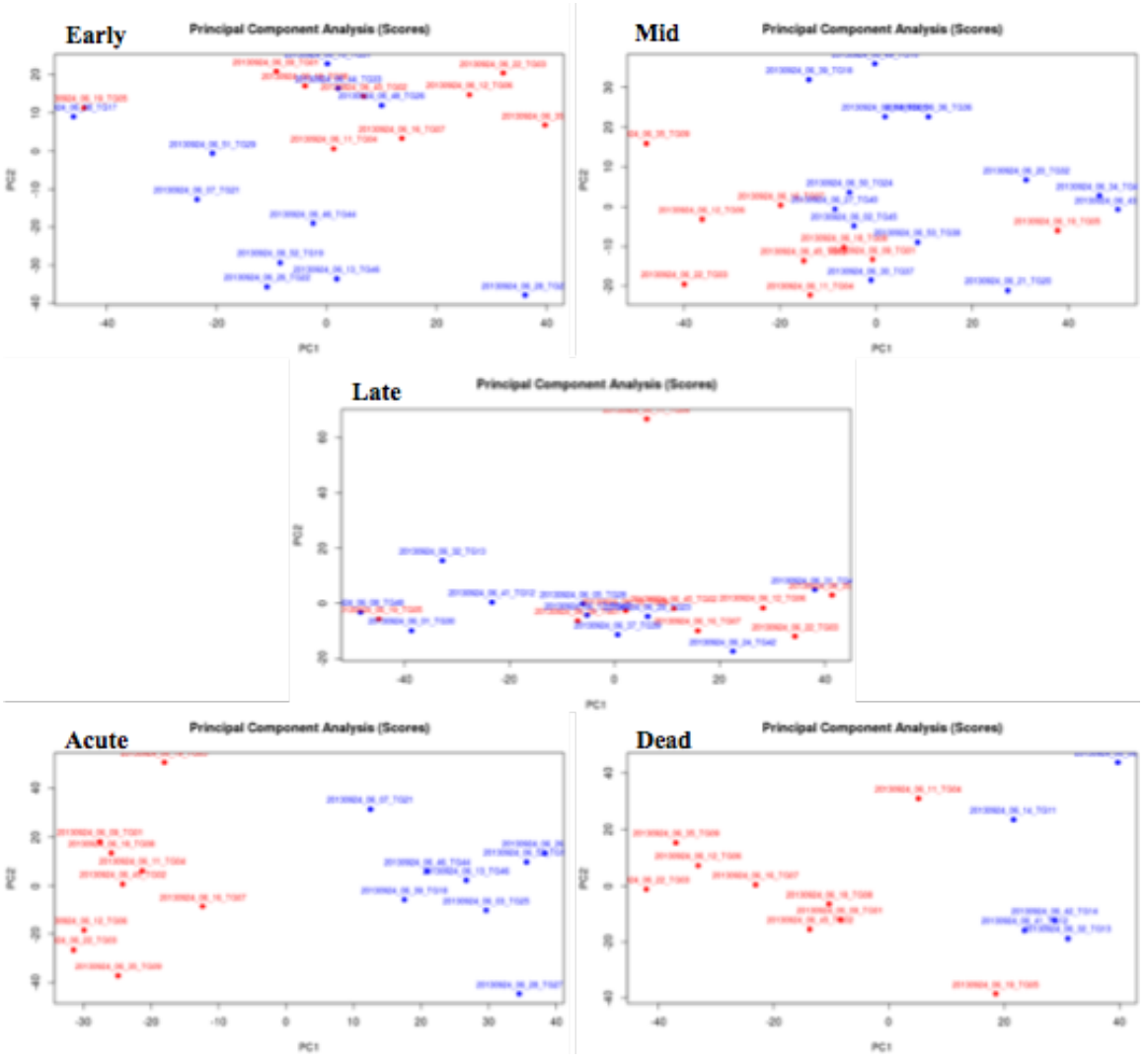


Figure 11 Time since symptom onset PCAs. Principle components analysis of all samples analyzed grouped by the time since symptom onset of the Lassa positive patients. Red data points are from febrile non Lassa patients, blue data points are indicated in the top left of each plot.

Figure 12

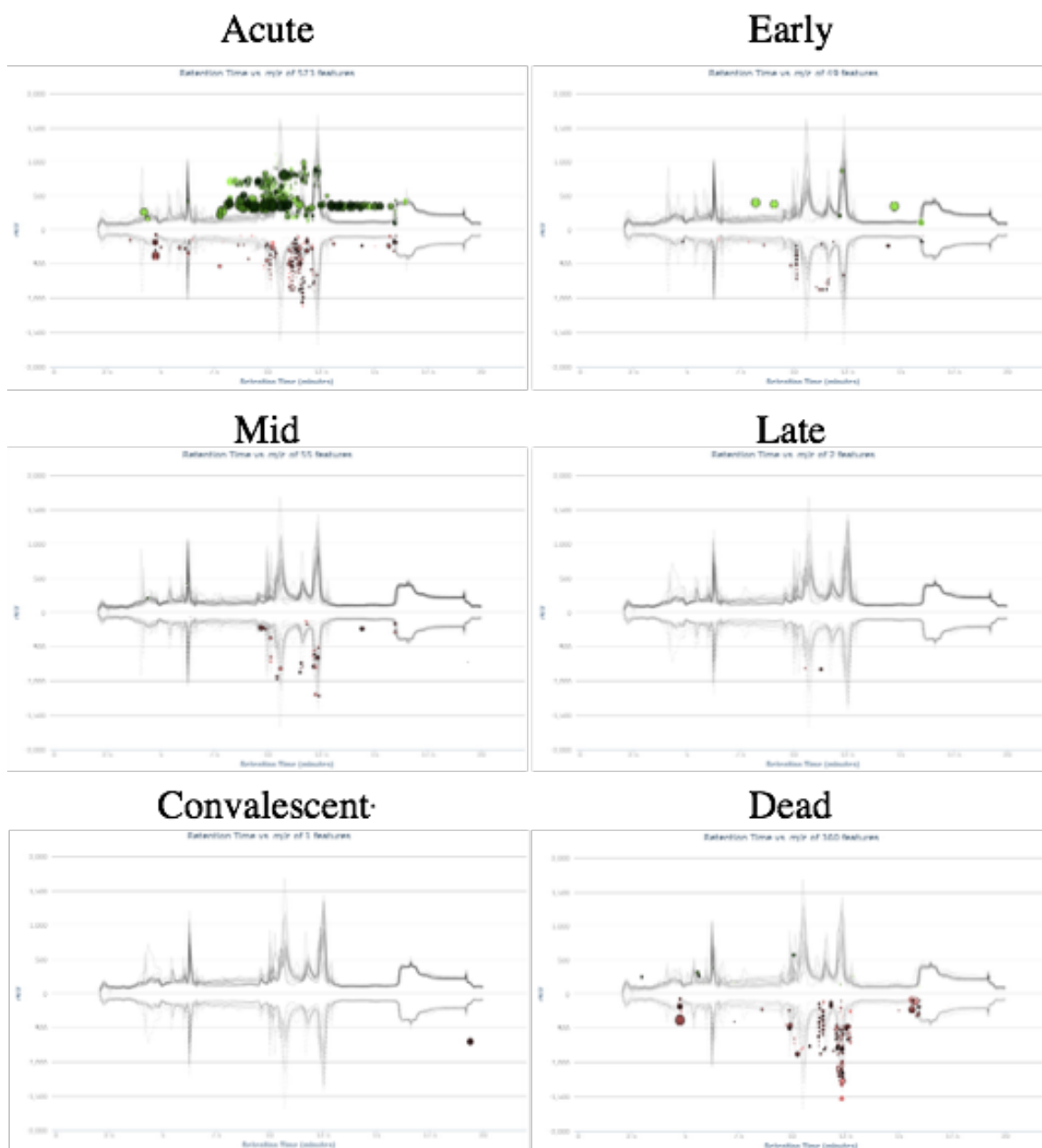


Figure 12 Serum metabolome is dynamic during LASF. Cloudplots (XCMS) illustrate the dynamic flux of the serum metabolome of Lassa fever patients whereby acute and terminal samples exhibit drastic differences while late and convalescent sera are virtually indistinguishable when compared to Non-Lassa control samples at the small molecule level. *In order to generate a cloudplot for convalescent sera graphing parameters were adjusted to plot features of less than 1.5-fold change, as there are no feature differences with $p = 0.05$ and ≥ 1.5 -fold change between Non-Lassa and convalescent sera metabolomes.

To understand global metabolite patterns as a cross section of the complete febrile population principle components analyses comparing the entire LCMS dataset where performed with FactoMineR. Globally, PCA discriminates FNL verse any LP specimens. Additionally, FNL and Terminal LP (“Dead”) samples form distinct groups but sero-positive samples form an indiscriminate mix (**Figure 13**). When grouped by time since symptom onset, global PCA reveals a more distinct metabolic picture where FNL and Terminal samples reserve the distinct grouping and acute samples form a more distinct group than when assigned by sero-status. Thus patients hypothesized to be suffering the active phases of LF (acute & terminal) present a pattern of dysregulation featuring greater serum metabolite flux (disparity) highlighted by the heterogeneity of endemic serostatus verse disease time (**Figure 13**).

Figure 13

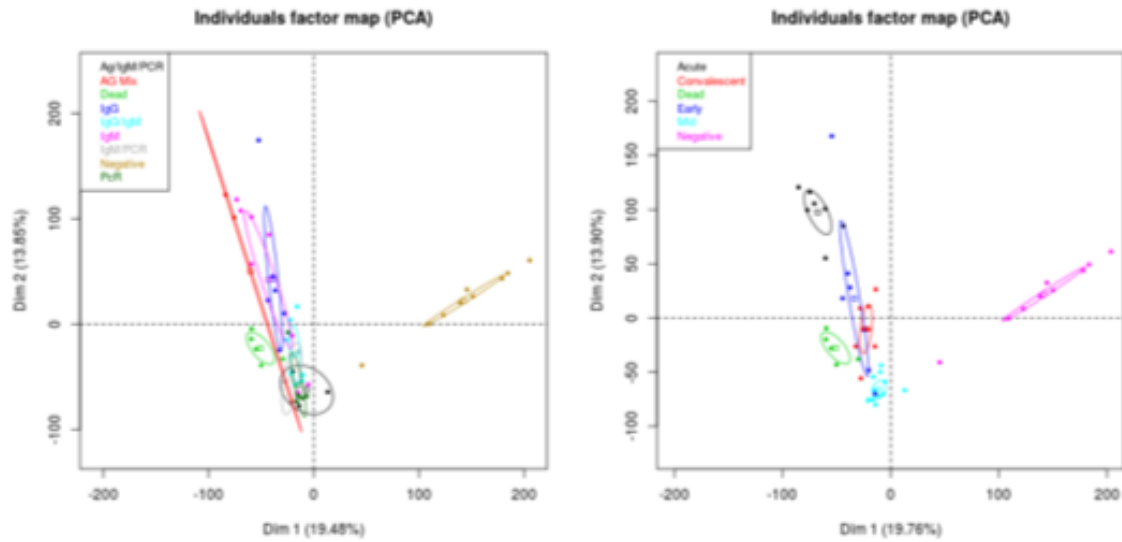


Figure 13 Principle Component Analysis. PCA for each sample in the panel was performed in R on total post-XCMS raw mass spectral data for each patient sample using the FactoMineR package. PCA analysis was performed with data grouped by sero-status (left) and by disease state (right). Sorting by sero-status indicator yields a more heterogeneous clustering while spectral features delineate time since symptom onset more clearly.

Putatively identified molecular classes

Lipid species constituted the bimolecular class most frequently assigned putative identification. Substituents of the primary lipid classes including fatty acids and conjugates, fatty esters, glycerophosphocholines, glycerolipids, diacylglycerols, glycerophospholipids, prenol, sterol, sphingolipids, vitamin D3 and derivative species where amongst those putatively identified and significantly dysregulated in one LP condition compared to FNL sera. Cluster analyses where performed computing divergence of standard deviation from the mean. As a first approximation, a dataset representative of the total putative lipid profile was assembled and analyzed (**Figure 14**). Following cluster analysis (represented via heatmap) total lipid data are utilized for machine learning (ML) analyses with the RF algorithm. In one analyses samples representing the stages of febrile illness at KGH train RF. In a second, only those samples from FNL and those from hypothesized acute sera are used. The RF analysis ranks features on they're computed predictive power, which can be used to remove features poorly performing before reanalyzing with cluster analysis (**Figure 14**). Putative and lipid and peptide datasets with highest RF accuracy and gini importance (typically ≥ 0.5 , and ≥ 0.25 , respectively) where chosen to analyze the most relevant sample populations (**Figure 15**). These datasets reveal conserved results across lipid or peptide species. Deceased patient samples constituted one of two major groupings where dysregulation of lipids and peptides manifests as lower signal intensity observed compared with acute phase or FNL. Note figure legends are opposite of each other. Within the second dominant grouping, acute samples are most unique where increased serum

glycerophospholipids and carnitine species are observed from these patient sera. Conversely, serum lipid profiles of non-Lassa and convalescent patients are most closely related suggesting a progression towards a non-Lassa state in convalescence. Patients with a diagnostic profile positive for the 3 indicators viral antigen (Ag), viral genome (rtPC), and anti-Lassa IgM where included, clustering between acute and convalescence samples where loss of most lipid signals is observed but lacking elevated species observed in more acute phases of disease, potentially illustrating the lipid profile of a transition phase of LF. Acute sera additionally contain elevated groupings of glycerophosphoserines, glycerphosphoglycerols, and glycerophosphocholines.

Figure 14

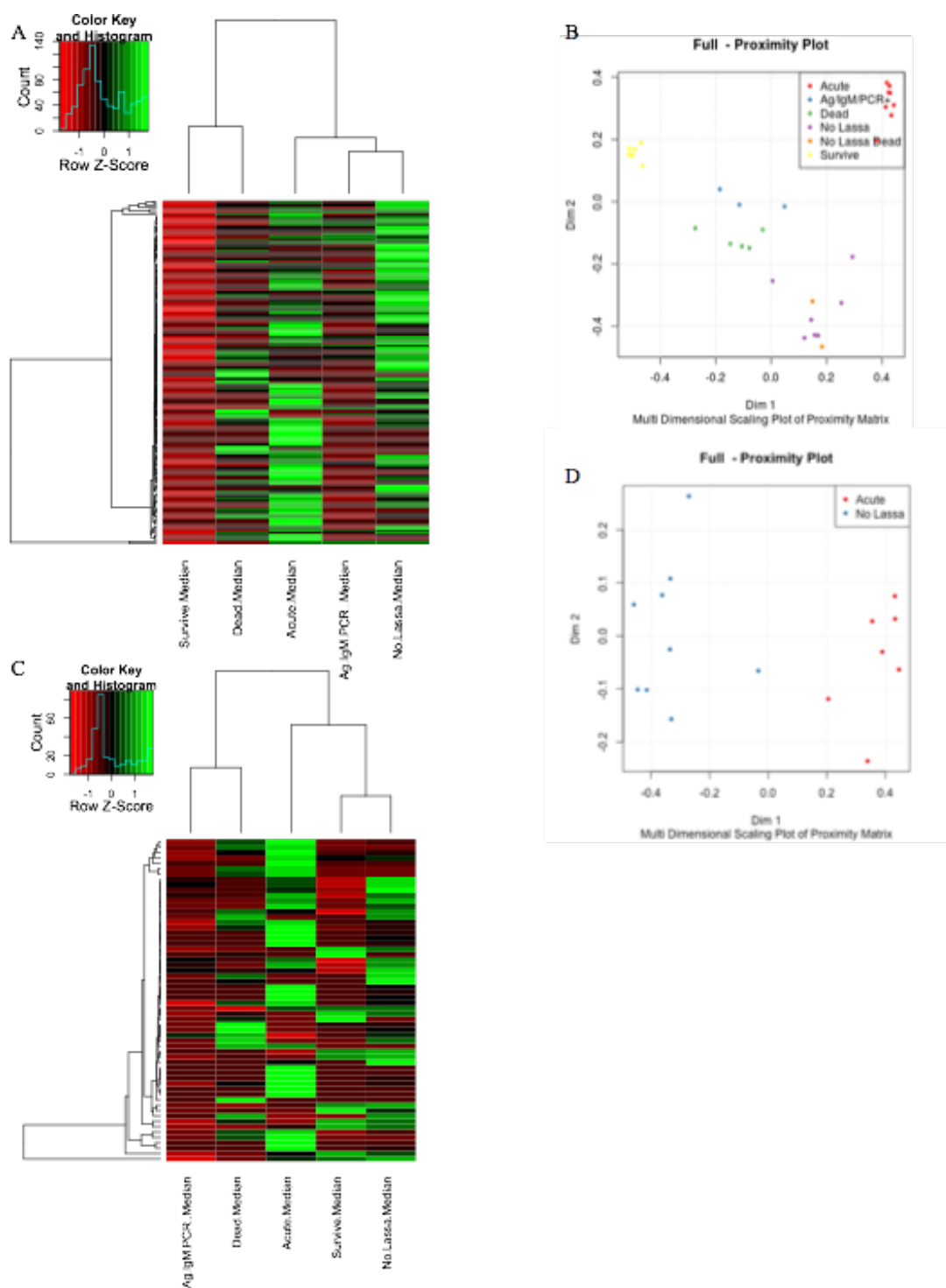


Figure 14 Iterative analysis and training. An initial lipids dataset featured 160 putative lipids (top heatmap). The dataset was analyzed by the Random Forests machine learning algorithm for disease status (top, right) and outcome (binary). Using the feature rank, features with little predictive value in RF were removed and cluster analysis performed again (left, bottom). The refined dataset illustrates how to train, trim, and retrain with features to identify serum small molecules as the strongest biomarkers of disease status.

Figure 15

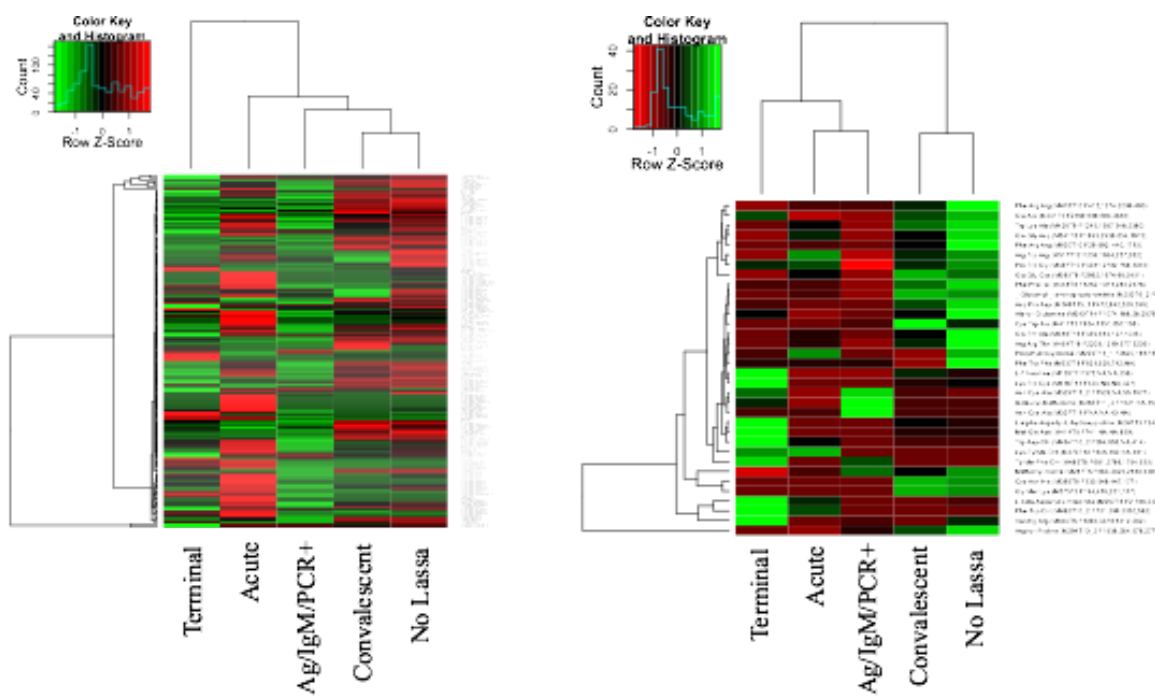


Figure 15 Cluster analysis of molecular class. Left heat map shows the top 60 most powerful lipids based on machine learning analysis. Right heat map shows putatively identified peptide species.

Metabolites with biomarker property

In addition to computing the statistical significance of all comparable, differential small molecules in the samples analyzed XMCS outputs include putative metabolite identifications. Multiple features are of particular interest with respect to LF pathology. The phospholipid Platelet-activating factor (PAF), its metabolic precursors and breakdown products, and other PAF-like lipids are putatively identified in multiple samples as statistically significant features as computed by Welch's t test (p values) and confirmed by computing false discovery rate (q value) (Table 5). Protonated and sodiated adducts of the phosphatidylcholine, platelet-activating factor (PAF) C-16, its metabolic precursors Lyso-PAF C-16 and Arachidonoyl PAF C-16 (**Figure 16**), and 9 additional PAF-like lipids are putatively identified or identified via manual *m/z* screening (**Figure 17**).

Table 5

Name	Adduct	rt	Observed <i>m/z</i>	Theoretical <i>m/z</i> (METLIN)	% Same <i>m/z</i>	p	q
PAF C-16	H+	12.16	524.3659	524.3711	1.000010	0.01	0.05
	Na+	12.16	546.3475	546.353	1.000010	0.001	0.04
Lyso-PAF C-16	H+	12.68	482.3557	482.3605	1.000010	0.006	0.04
	Na+	12.68	504.3373	504.3424	1.000010	0.006	0.04
Arachidonoyl PAF C-16	H+	10.44	768.579	768.5901	1.000014	0.07	0.1
	Na+	10.5	790.5632	790.5721	1.000011	0.01	0.06
PAF C-18	H+	12.03	552.3945	552.4024	1.000014	0.01	0.05
	Na+	12.03	574.3759	574.3843	1.000015	0.01	0.06
Lyso-PAF C-18	H+	12.49	510.3858	510.3918	1.000012	0.01	0.06
	Na+	12.5	532.3672	532.3737	1.000012	0.01	0.06
PC(O- 10:1(9E)/2:0)	H+	9.7	438.2934	438.2615	0.999927	0.02	0.08
	Na+	9.73	460.2754	460.2434	0.999930	0.01	0.06
PC(O- 16:1(11Z)/2:0)	H+	12.21	522.3504	522.3554	1.000010	0.001	0.04
	Na+	12.21	544.3327	544.3373	1.000008	0.001	0.04
PC(O- 18:1(10E)/2:0)	H+	12.08	550.3808	550.3867	1.000011	0.002	0.04
	Na+	12.08	572.3629	572.3686	1.000010	0.0002	0.04
PC(O- 18:2(9Z,12Z)/2:0)	H+	12.15	548.3552	548.371	1.000029	0.0007	0.04
	Na+	12.13	570.3463	570.353	1.000012	0.0005	0.04
PC(O-12:0/2:0)	H+	12.58	468.3034	468.3085	1.000011	0.002	0.04
	Na+	12.58	490.2852	490.2904	1.000011	0.004	0.04
PC(O-14:0/2:0)	H+	12.32	496.3352	496.3398	1.000009	0.04	0.12
	Na+	12.32	518.3167	518.3217	1.000010	0.003	0.04
PC(O-15:0/2:0)	H+	12.24	510.35	510.3554	1.000011	0.009	0.05
	Na+	12.24	532.3317	532.3374	1.000011	0.008	0.05

Figure 16

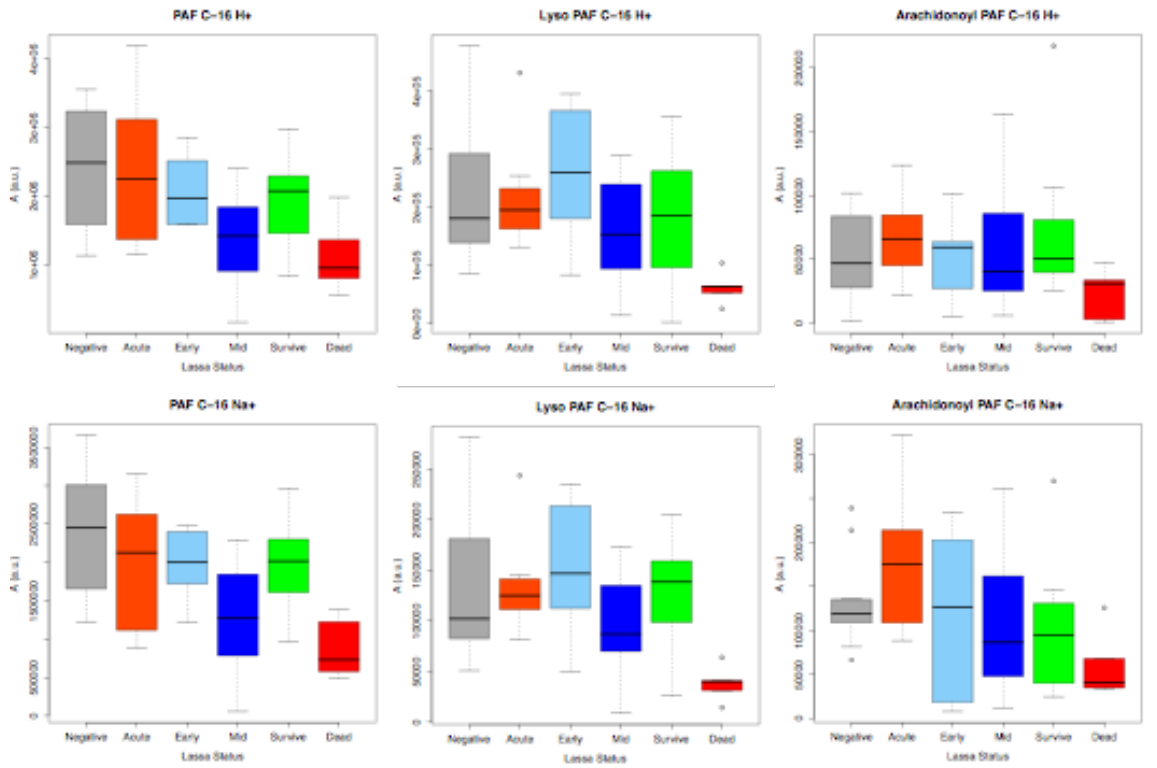
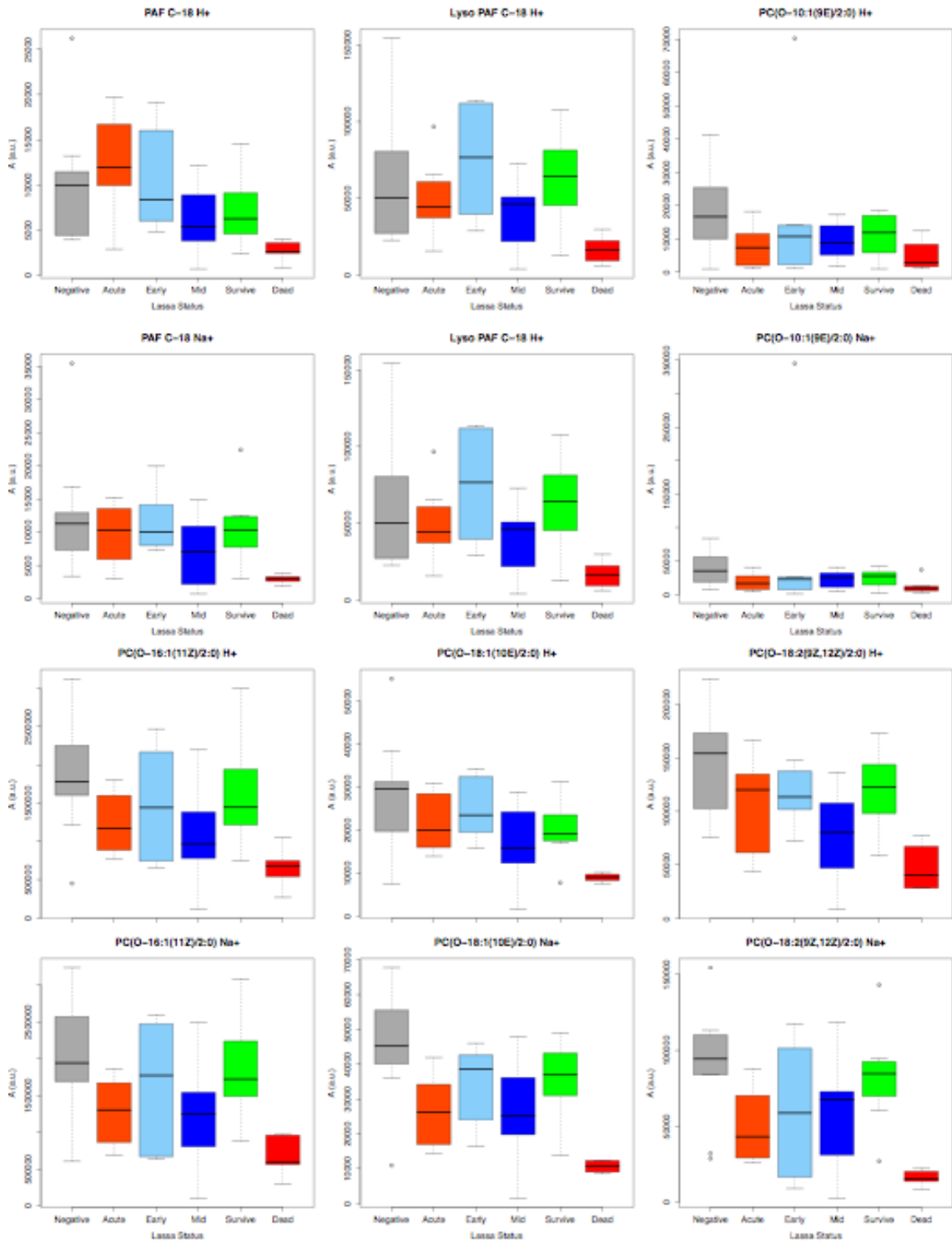


Figure 16 Platelet-activating factor. Spectral intensity for Platelet-activating factor C-16, Lyso Platelet-activating factor C-16, and Arachidonoyl Platelet-activating factor C-16 showing uniform intensity loss in terminal patients.

Figure 17



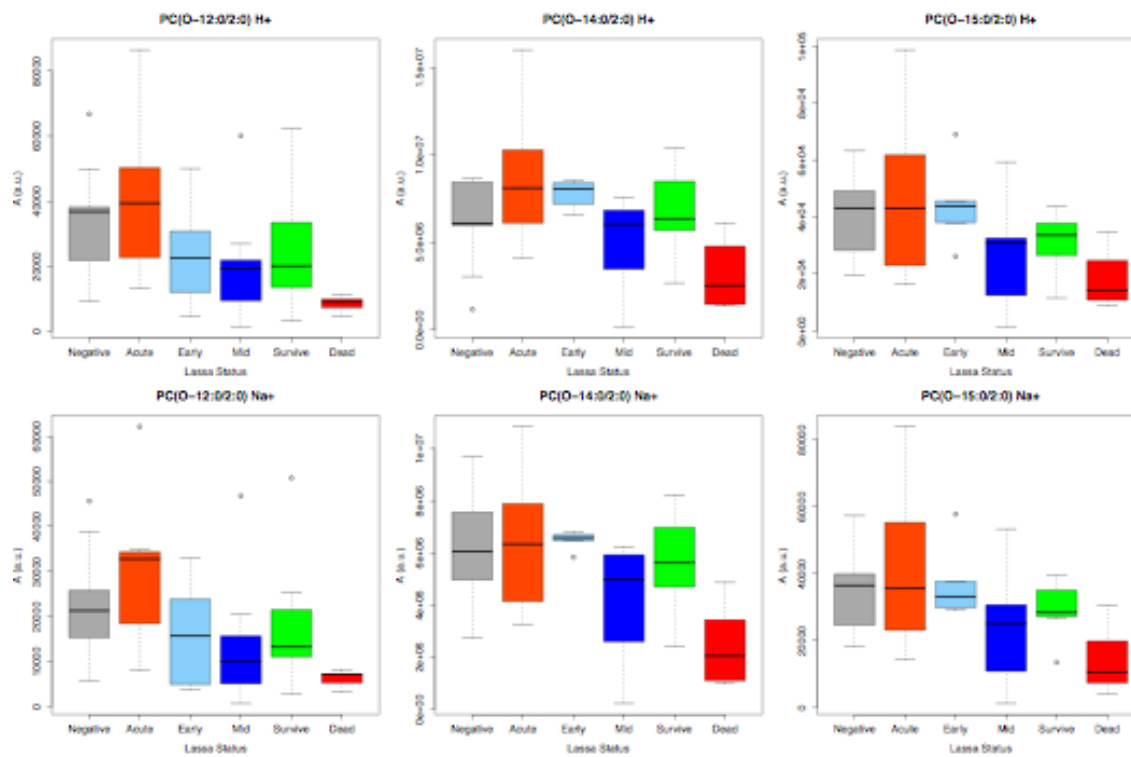


Figure 17 Platelet-activating factor-like molecules. In addition to the most well characterized PAF molecules 9 additional lipids with PAF-like structure are identified with conserved pattern of signal intensity loss in terminal versus febrile non Lassa patient sera.

Representative extracted ion chromatograms (EICs) and mass spectrums from PAF C-16 and Lyso PAF C-16 are presented (**Figure 18**). A general pattern of signal intensity loss between Lassa Negative and terminal patient sera is observed and consistent among all PAFs. Compared to Lassa Negative patients PAF C-16 is significantly lower in Mid disease patients ($p = 0.01$ & 0.006) and terminal patients ($p = 0.01$ and 0.001) for H^+ and Na^+ adducts, respectively. Lyso PAF C-16 ($p = 0.006$ H^+ and Na^+) and Arachidonoyl PAF C-16 ($p = 0.01$ for Na^+) are significantly lower in terminal patient sera compared to FNL patients. Two PAF-like lipid species register H^+/Na^+ adducts at m/z 550.3867/572.3686 ($p = 0.002/0.0002$) and 548.3710/570.3530 ($p = 0.0007/0.0005$) exhibit marked reduction in signal intensity in terminal patients compared to non-Lassa controls.

Figure 18

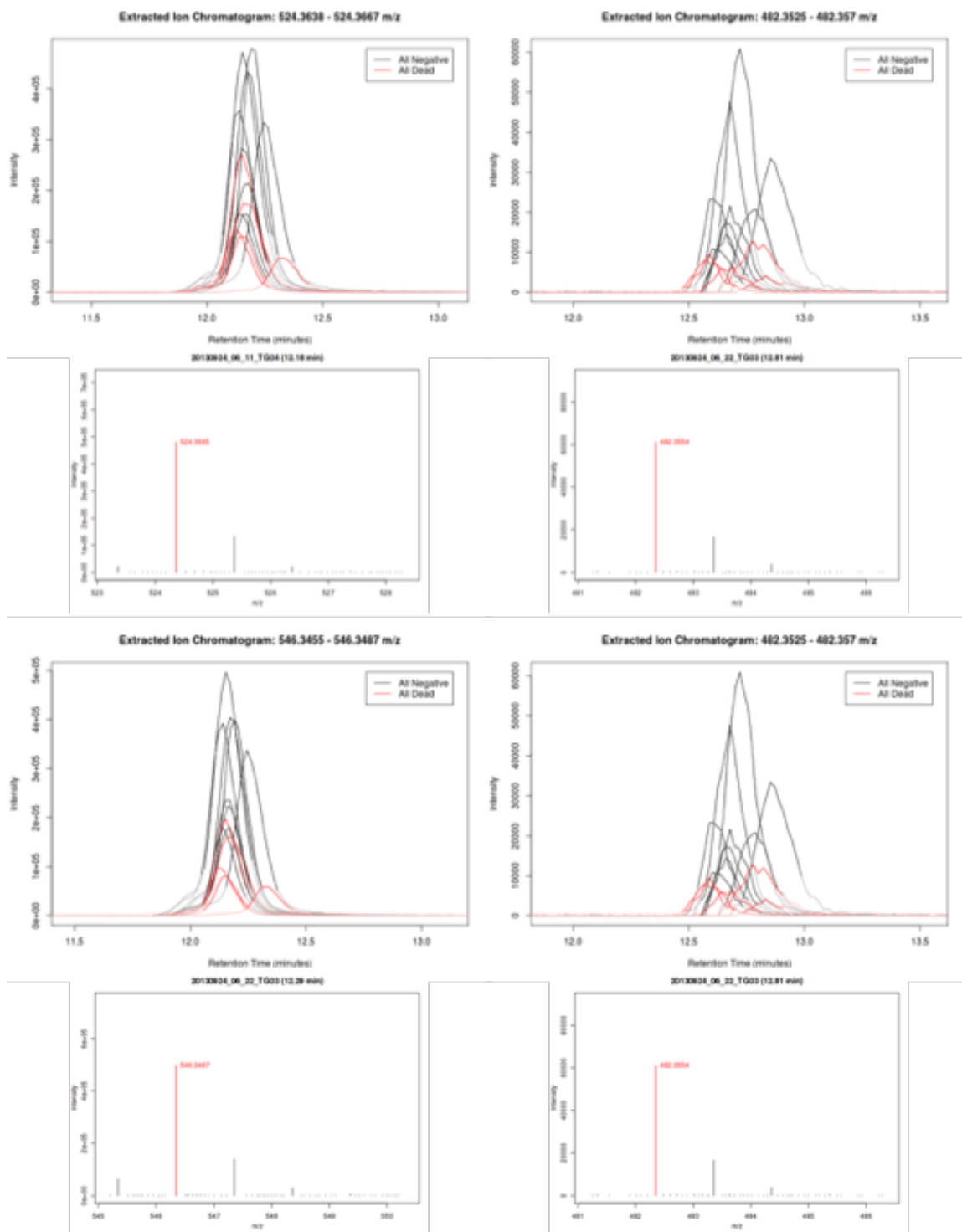


Figure 18 Platelet-activating factor analysis. Extract Ion Chromatograms (EICs) and mass spectrometry for PAF C-16 (left) and Lyso PAF C-16 (right), H⁺ and Na⁺ adducts top and bottom.

When spectral intensity for all PAF molecules is normalized (max/column equal to 1.0) patterns start to emerge where Arachidonoyl PAF C-16 H⁺, Arachidonoyl PAF C-16 Na⁺, PC(O-14:0/2:0) H⁺, PC(O-15:0/2:0) Na⁺, PC(O-12:0/2:0) H⁺, PC(O-12:0/2:0) Na⁺, and PAF C-18 H⁺ all show increased spectral intensity compared to FNL while PAF C-16 Na⁺, PAF C-18 Na⁺, Lyso-PAF C-18 H⁺, Lyso-PAF C-18 Na⁺, PC(O-18:1(10E)/2:0) H⁺, PC(O-18:1(10E)/2:0) Na⁺, PC(O-18:2(9Z,12Z)/2:0) H⁺, PC(O-18:2(9Z,12Z)/2:0) Na⁺, PC(O-16:1(11Z)/2:0) H⁺, PC(O-10:1(9E)/2:0) H⁺, and PC(O-10:1(9E)/2:0) Na⁺. While the protonated adduct of PAF C-16 remains unchanged between Acute and FNL, it and all PAF features have a precipitous loss of signal intensity in terminal patient sera (**Figure 19**).

Figure 19

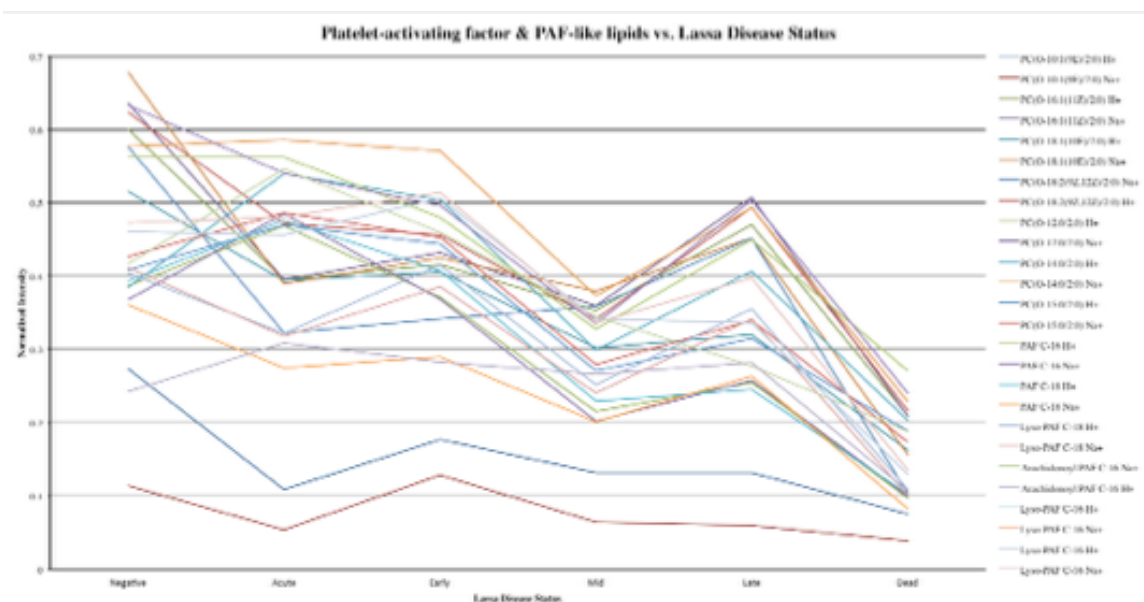


Figure 19 Normalized spectral intensities for Platelet-activating factor (PAF) and PAF-like lipids. Serum glycerophosphocholine intensities are higher in febrile non-Lassa (Negative) patient sera while precipitously lower in terminal Lassa patient sera compared to all other Lassa time points.

Several products of coagulation and hemoglobin breakdown are putatively identified and dysregulated in acute and terminal Lassa fever patient sera. The sodiated adduct of soluble fibrin monomer is observed at m/z 168.0750 with a significant drop in acute ($p = 0.0001$) patient samples compared to FNL (**Figure 20**). Fibrin remains significantly lower through the course of disease compared to control sera where median values gradually increase yet remain lower in terminal samples ($p = 0.03$). Also shown in **Figure 20** are protonated adducts of the hemoglobin breakdown product(s) D-Urobilinogen and/or I-Urobilin with identical protonated m/z 591.3195 and $rt = 12.10$ are putatively identified as significantly reduced signal intensity in terminal patients versus Lassa Negative samples ($p = 0.0007$). Expectedly, a spectral feature ($rt = 12.20$) with m/z 613.3223 consistent with the sodiated adduct of D-Urobilinogen and/or I-Urobilin is also present and similarly significantly reduced in terminal sera versus FNL ($p = 0.0008$). A feature with H^+ m/z 593.3344 (rt 12.13) and Na^+ m/z 615.3373 (rt 12.17) is consistent with Mesobilirubinogen and significantly reduced in terminal samples compared to Lassa Negative ($p = 0.0005$ and $p = 0.0003$ for H^+ and Na^+ , respectively).

Figure 20

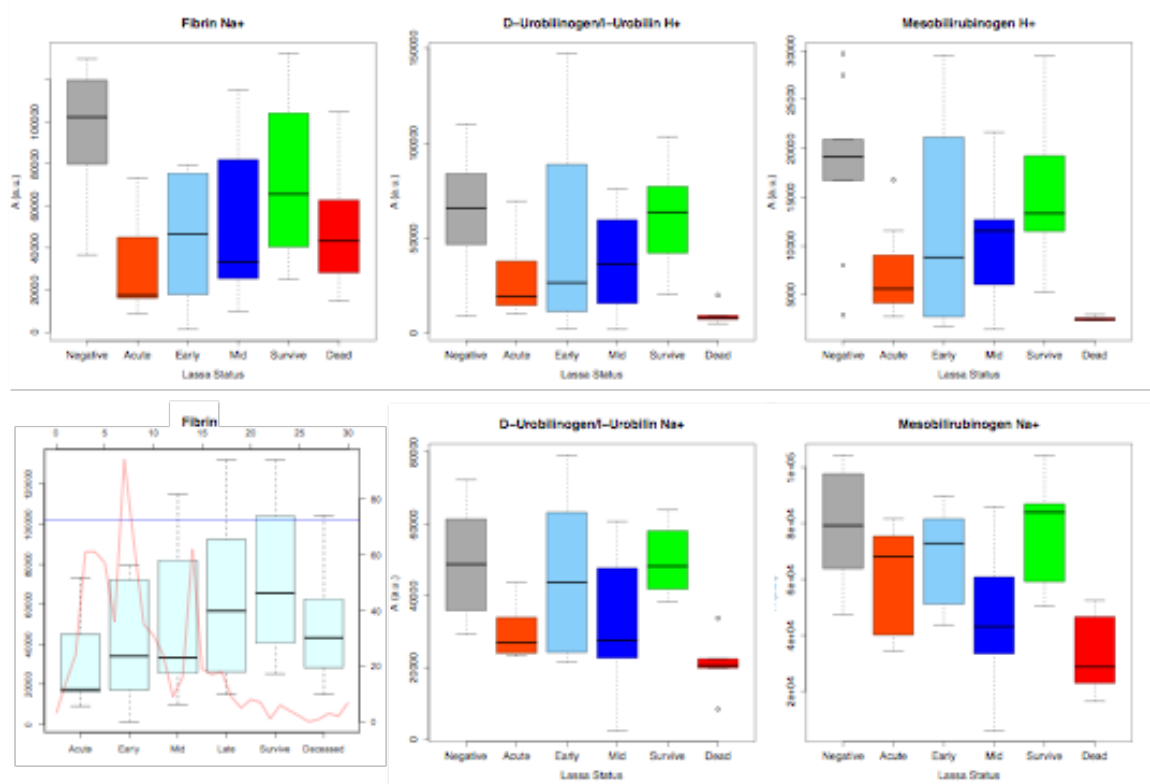


Figure 20 Blood product abnormalities. Soluble fibrin monomer (left top and bottom) and hemoglobin breakdown products show loss in acute patient sera versus febrile non-Lassa. In terminal Lassa fever patients, soluble fibrin monomer intensity rises slightly though is still significantly lower than febrile non-Lassa, while heme breakdown products feature lower intensity in terminal patients.

A spectral feature putatively identified as the sodiated and potassiated adducts of the methylated nucleotide, 1-methylinosine with m/z 305.0856 and 321.0596 is significantly elevated in Terminal patients ($p = 0.004$ and $p = 0.01$, Na^+ and K^+ , respectively). Additionally, a spectral feature consistent with the protonated adduct of 1-methylinosine is detected with m/z 283.1016 (theoretical $m/z = 283.1037$) also with retention time at 5.58 (**Figure 21**).

Figure 21

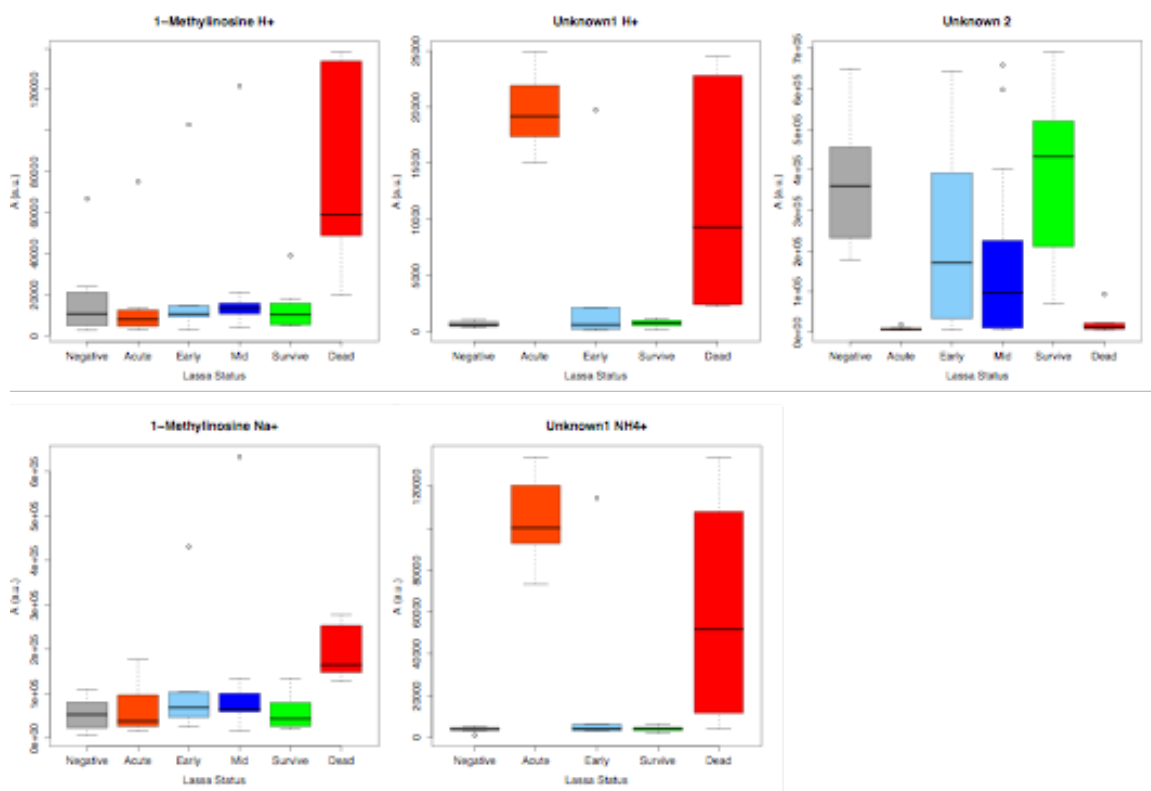


Figure 21 Methylinosine and Unknown molecules. The methylated nucleoside 1-methylInosine is observed with significantly elevated signal intensity only in terminal patient sera. Two Unknown molecules are presented where Unknown 1 is significantly elevated in acute sera and remaining elevated in terminal patient sera. Unknown 2 is significantly lower in acute samples, returns to the identical serum representation in convalescent patients while remaining significantly lower in terminal patient sera.

Multiple unknown features are present in these samples registering significant disparity between Acute and/or Terminal sera and FNL controls. However, exhaustive characterization of all unknowns presents a limiting step in all metabolomics studies, as is the case at present. It is however useful to illustrate unknowns. Two unknown features are particularly pronounced in the current dataset. Unknown 1 is predicted to have a molecular weight of 101.046 and detected with H^+ m/z 102.0537 and NH_4^+ m/z 119.08 (rt = 15.95) significantly elevated in Acute samples compared to Lassa Negative ($p = 7.42 \times 10^{-7}$ and $p = 1.87 \times 10^{-6}$ for H^+ and NH_4^+ adducts, respectively) (**Figure 21**). A second highly significant unknown spectral feature is detected at m/z 187.0693 (rt = 4.73) with significant reduction in Acute and Terminal sera compared to Lassa Negative ($p = 0.0001$ and $p = 0.0002$, respectively). The validity of unknown features is illustrated by EICs and spectrums of protonated and ammoniated adducts of Unknown 1 in acute verse FNL or the detection of Unknown 2 in acute verse FNL and terminal verse FNL (**Figure 22**). Further, these data recapitulate in an additional sample set where the Unknown feature with m/z 119.08 is observed in LP samples (**Figure 23**) with a EIC and spectrum similar to the primary data (Figure 17). The presence of the feature with m/z 119.08 in 2 FNL may represent a non-specific nature of the metabolite. Alternatively this data may represent a feature detectable in patients with LF before conventional diagnostics detect LF antigen or antibodies.

Figure 22

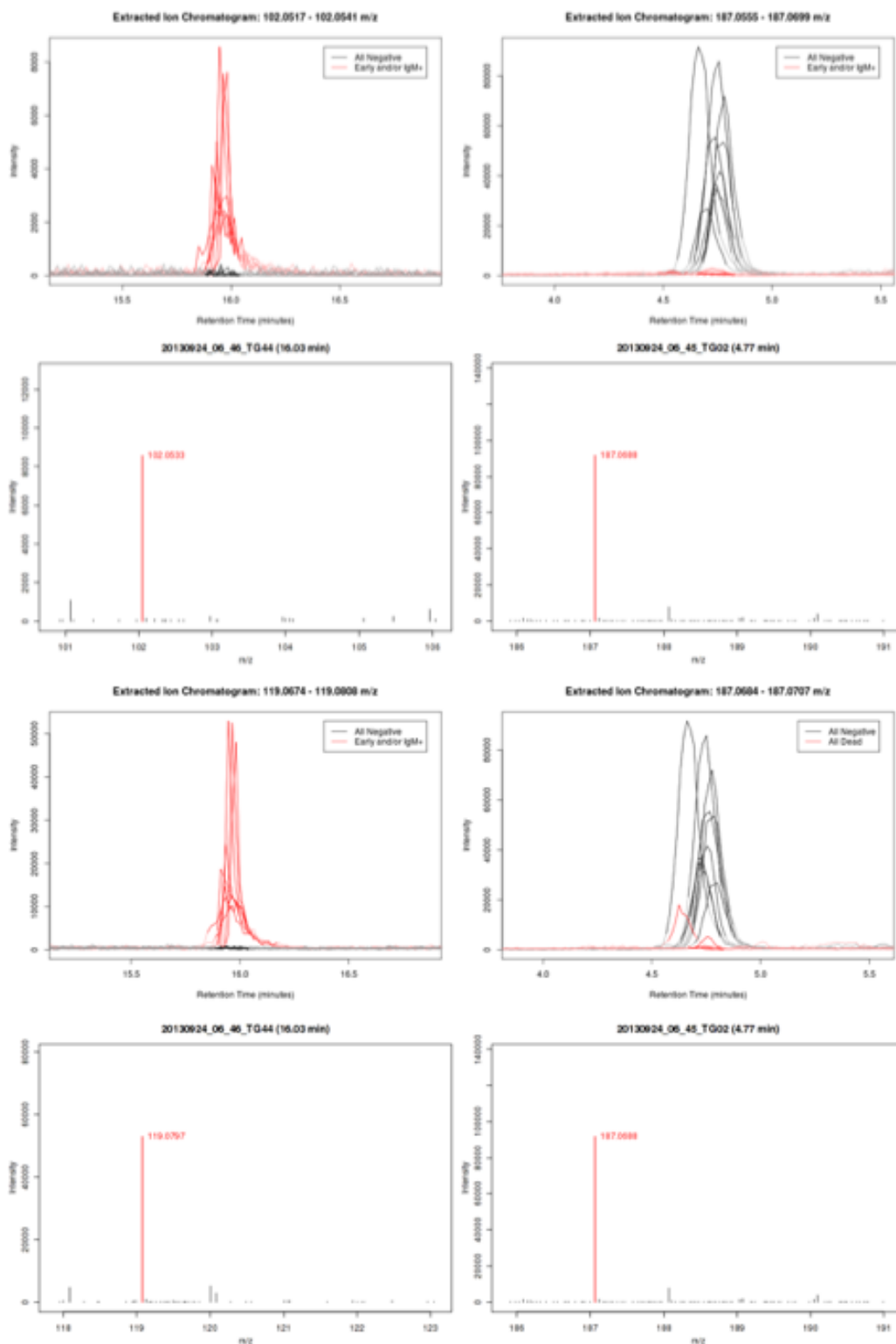


Figure 22 Unknowns spectral data. EICs and spectrums of Unknown 1 and 2. U1 is only present at significant quantity in Acute patients while U2 is lost in Acute patients and remains down in Terminal patients.

Figure 23

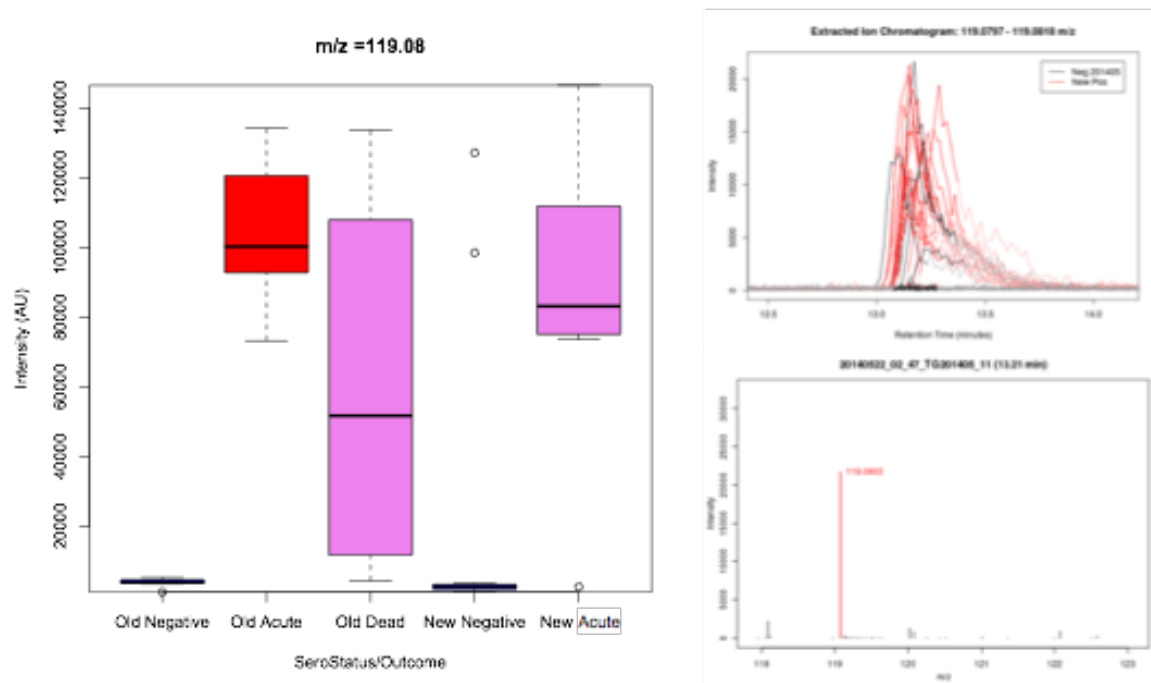


Figure 23 Unknown 1 recapitulates. The ammoniated adduct of Unknown 1 is detected in a second sample set of Lassa positive samples compared to febrile non Lassa samples corroborating its correlation as a function of Lassa virus.

Machine learning identifying metabolites with biomarker property

Machine learning provides a quantitative assessment of the differential utility of input data. In addition to computing a classification tree and multidimensional scaling plots the random forests machine learning algorithm ranks the power of the input variables in predicting outcome which was determined to the optimal machine learning algorithm to identify serum metabolites harboring predictive quality in Sierra Leonian LFW patients. Firstly, metabolites were identified for further analysis based on being significantly different in one state of the LF spectrum (identified by ranking differential features by p value) and analyzed for the ability to delineate positive LF versus FNL. Seventy-three of the most significant spectral features were computed through the RF algorithm with 10-fold cross validation and visualized with proximity plots (**Figure 24A**). Of the 73 features, only 14 spectral features register a Mean Decrease Accuracy (MDA) of ≥ 1.0 suggesting weakly predictive features and evidenced by the homogeneity of samples derived from LP (Lassa) and terminal patients (dead). Refinement of feature selection by observing valid spectral character, acceptable signal/noise, and presence in multiple groups similarly from primary data yields homogenous groupings (**Figure 24B**). This analysis yields 9 features with MDA of ≥ 5.0 .

The diagnostic power of metabolites was similarly assessed via RF registering binary outcome (Lassa Positive or No Lassa). To a 22 diagnostic biomarker (DB) dataset cluster analysis was applied where data were grouped as either sero-status or time since symptom onset (**Figure 24C**). Four DBs predict LP versus LN samples with sensitivity and specificity of ≥ 0.95 . Proximity plot of the DB dataset (**Figure 24D**) reveal the distinct sorting corroborating the groupings observed previously (**Figure 24C**).

Figure 24

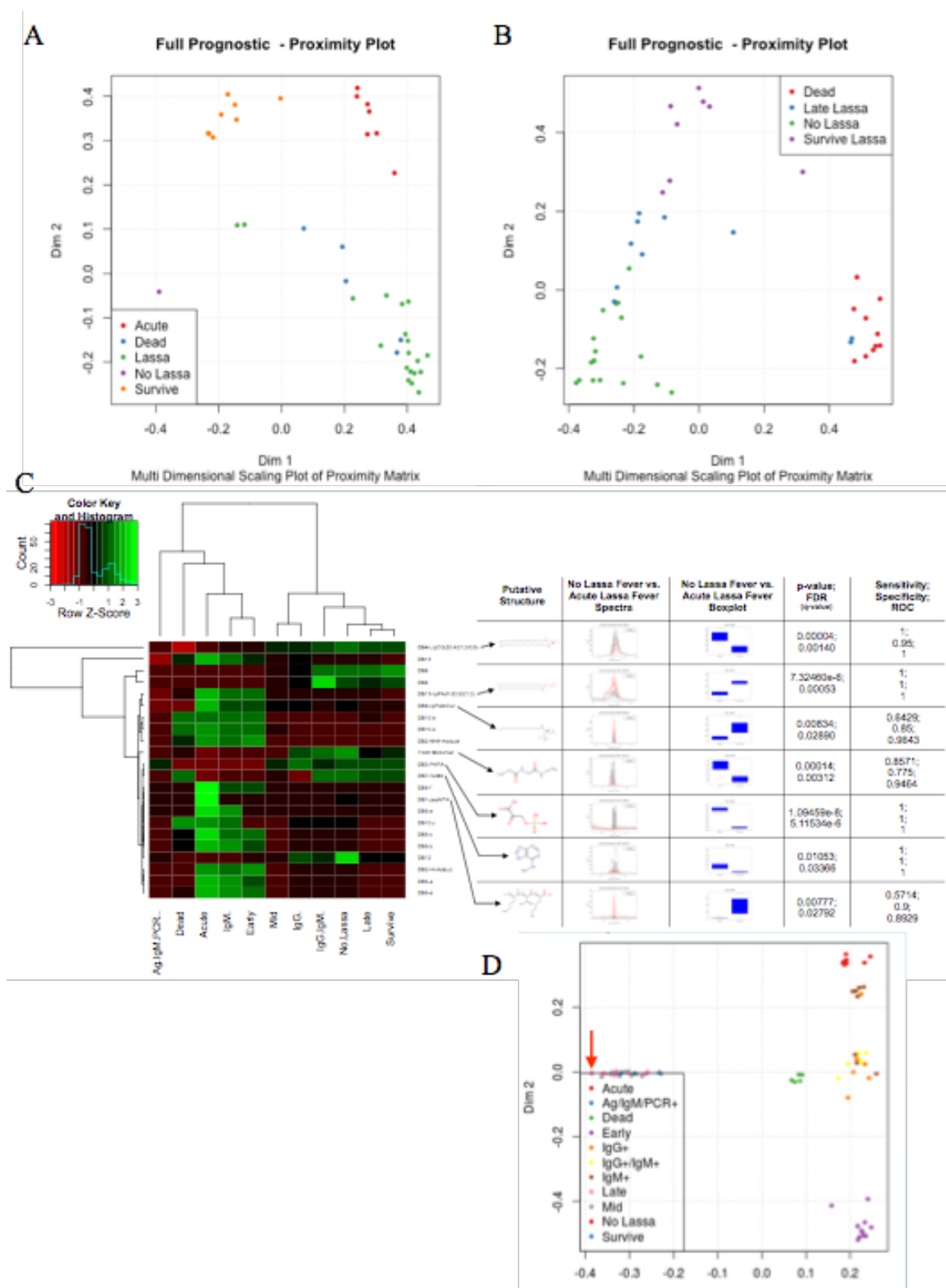


Figure 24 Diagnostic biomarkers. The discovery and validation of spectral features with correlation to Lassa positive samples.

Prognostic quality of serum biomarkers was similarly assessed with a particular focus on features with probable putative identification. In addition to the two Unknown features detailed above (m/z 102.05/119.08 and m/z 187.06) **Figure 25** details name, mass, formula, and structure of putative features assessed for prognostic character in LF.

Figure 25

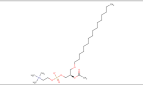
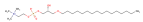
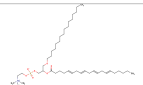
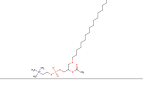

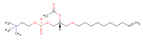

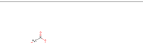




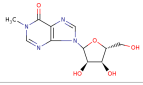
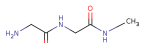
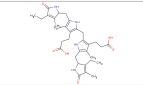
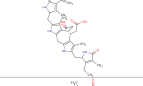
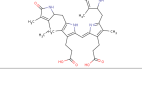
Name	Mass	Formula	Structure
PAF C-16	523.3637704	C ₂₆	
Lyso PAF C-16	481.3532062	C ₂₄	
Arachidonoyl PAF C-16	767.5828592	C ₄₄	
PAF C-18	551.3950688	C ₂₈	
Lyso PAF C-18	509.3845046	C ₂₆	
PC(O-10:1(9E)/2:0)	437.254226	C ₂₀	
PC(O-16:1(11Z)/2:0)	521.3481212	C ₂₆	
PC(O-18:1(10E)/2:0)	549.3794196	C ₂₈	
PC(O-18:2(9Z,12Z)/2:0)	547.3637704	C ₂₈	
PC(O-12:0/2:0)	467.301189343	C ₂₂	
PC(O-14:0/2:0)	495.332489471	C ₂₄	
PC(O-15:0/2:0)	509.348139535	C ₂₅	
1-Methylinosine	282.096421	C ₁₁	
Fibrin	145.085126611	C ₅	
Mesobilirubinogen	592.32608516	C ₃₃	
D-Urobilinogen	590.310435096	C ₃₃	
I-Urobilin	590.310435096	C ₃₃	

Figure 25 Putative prognostic molecules. Identifying information and structure for known molecules used in prognostic machine learning analysis.

When analyzed with binary outcome (Alive vs. Dead), four spectral features perform admirably with ROC, Sensitivity (true positive rate), and Specificity (true negative rate) at or above 0.75 (Table 6) and these prognostic biomarkers are sufficiently unique to group terminal LF patient samples from other LP and FNL (**Figure 26**). The protonated Mesobilirubinogen adduct performs best with Sensitivity of 0.97, and Specificity and ROC of 1.0. The sodiated adduct of Lyso-PAF C16 results in a Sensitivity of 0.97, Specificity of 0.92, and ROC of 0.99. The protonated adduct of the PAF glycerophosphocholine PC(O-14:0/2:0) has a Sensitivity of 0.88, Specificity of 0.83, and ROC of 0.91. Finally, the protonated adduct of methylated nucleoside 1-Methylinosine has a Sensitivity of 0.93, Specificity of only 0.75, and ROC of 0.90. Though non-exhaustive in nature given the thousands of differential features present in the complete dataset, panning the serum metabolome for discrete features has produced biomarkers with strong indication in specific stages of LF.

Table 6

	ROC	Sensitivity	Specificity
Mesobilirubinogen H+	1	0.9666667	1
Lyso-PAF C-16 Na+	0.9861111	0.9666667	0.9166667
PC(O-14:0/2:0) H+	0.9166667	0.875	0.8333333
1-Methylinosine H+	0.8993056	0.9291667	0.75
Unknown 1 H+	0.9305556	0.9416667	0.6666667
D-Urobilinogen/I-Urobilin H+	0.9027778	0.9375	0.6666667
PC(O-18:2(9Z,12Z)/2:0) H+	0.9305556	0.9208333	0.6666667
Lyso-PAF C-16 H+	0.8611111	0.9708333	0.5
PC(O-14:0/2:0) Na+	0.7673611	0.8208333	0.5
PC(O-18:2(9Z,12Z)/2:0) Na+	0.8159722	0.8125	0.5
Arachidonoyl PAF C-16 Na+	0.8958333	0.8833333	0.4166667
Fibrin Monomer	0.8993056	0.8541667	0.4166667
PC(O-18:1(10E)/2:0) H+	0.8298611	0.9416667	0.3333333
Lyso-PAF C-18 H+	0.8611111	0.8958333	0.3333333
Unknown 1 NH4+	0.6805556	0.875	0.3333333
1-Methylinosine Na+	0.8715278	0.8708333	0.3333333
D-Urobilinogen/I-Urobilin Na+	0.8645833	0.8666667	0.3333333
PAF C-16 Na+	0.7951389	0.8583333	0.3333333
PC(O-18:1(10E)/2:0) Na+	0.8090278	0.8458333	0.3333333
Lyso-PAF C-18 Na+	0.7743056	0.7958333	0.3333333
PC(O-15:0/2:0) Na+	0.8298611	0.8958333	0.1666667
PAF C-18 H+	0.8159722	0.8958333	0.1666667
PC(O-12:0/2:0) Na+	0.7986111	0.8666667	0.1666667
PC(O-16:1(11Z)/2:0) Na+	0.7465278	0.8583333	0.1666667
PC(O-15:0/2:0) H+	0.9201389	0.8083333	0.1666667
PAF C-18 Na+	0.6840278	0.8708333	0.0833333
Arachidonoyl PAF C-16 H+	0.8854167	0.8375	0.0833333
Mesobilirubinogen Na+	0.8368056	0.9333333	0
Unknown 2	0.7361111	0.9166667	0
PC(O-12:0/2:0) H+	0.7916667	0.8791667	0
PC(O-10:1(9E)/2:0) H+	0.7569444	0.8791667	0
PC(O-16:1(11Z)/2:0) H+	0.65625	0.8541667	0
PAF C-16 H+	0.7673611	0.8458333	0
PC(O-10:1(9E)/2:0) Na+	0.7986111	0.8291667	0

Figure 26

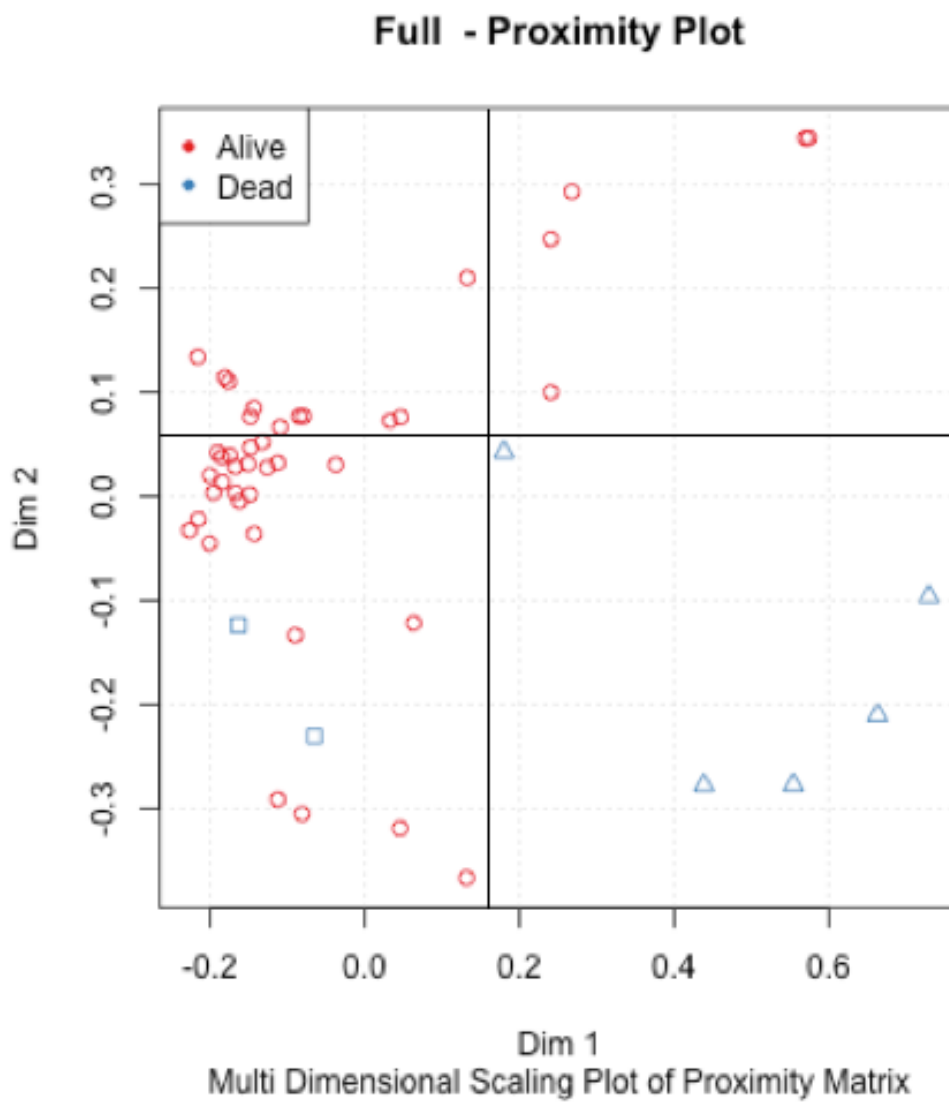


Figure 26 Machine learning with binary outcome. Random forests analysis of candidate prognostic biomarkers reveals terminal Lassa fever patients (blue triangle) sort into a unique group separate from any other Lassa positive samples or febrile non Lassa (red circle) or terminal febrile non Lassa (blue square). Segments added manually to aid in visualization.

Aim 2

Host Proteins Exclusively From Extracellular Viral Particles

The LCMS-characterized global protein composition of lipid virions, virion core, or replication-deficient vectors (gene therapy vectors) has been carried out from 9 families of enveloped viruses: Retroviridae including Human Immunodeficiency Virus-1[106-111] and Moloney murine leukemia virus[112]; Poxviridae featuring Vaccinia virus[113]; Herpesviridae including Human Cytomegalovirus[114], Human Herpes simplex virus type 1[115], Epstein-Barr virus[116], Kaposi's sarcoma-associated herpesvirus[117, 118], Murid herpesvirus 68[119, 120], and Pseudorabies virus[121]; Arteriviridae analyzing Porcine respiratory and reproductive syndrome virus[122], Flaviviridae analyzing Hepatitis C virus[123], Bunyaviridae analyzing Rift Valley fever virus[124], Orthomyxoviridae analyzing Influenza A virus[125], Paramyxoviridae analyzing Newcastle disease virus[126] and Respiratory Syncytial virus[127], Rhabdoviridae analyzing Vesicular stomatitis virus[128], and the Filoviridae analyzing Zaire Ebola virus and Marburg virus[91]. Genetic divergence of representative viral species are analyzed and illustrated by viral genome (**Figure 27A**). From 24 studies covering 18 distinct viral species 1,157 proteins of host origin are identified in isolated extracellular virus. Of all proteins identified there are 713 unique proteins (718 individual proteins when accounting for interspecies differences, 713 trace to unique Gene identifiers). Nine of the most frequently identified proteins are found in at least 50% of all virions examined by LCMS regardless of spectrometric conditions and study design. Cytoskeletal proteins and scaffold affiliated proteins make up the most frequently

identified host proteins in enveloped virions. Cytoplasmic actin (ACTB) is identified in 72.2% of virions; Annexins A2, A1, and A5 were identified in 77.7, 61.1, and 55.5 % of virions respectively; Cofilin 1 was also in 55.5% of virions. Protein binding proteins are also frequently identified including Cyclophilin A (PPIA) an enzyme mediating protein folding that is present in 61% while Heat shock 70 kDa protein 8 (HSPA8) was in 55.5%. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is in 66% while Enolase is in 50% of virions. Additional cytoskeletal and protein binding proteins are found most frequently in extracellular virions though exhaustive detail is beyond the scope of this manuscript; all features detected and the percentage of virions they appear in are available in Table 2. Gene Ontology (GO) classification partitions proteins into 3 categories: Biological process, Molecular function, and Cellular component [129, 130]. The host protein composition of extracellular lipid virions yields 4,058 GO identifications. The top 5 GO classifications by annotation of host proteins in virions shows 14.9% of all annotations (35,775) mediate protein binding, 7.36% derive from cytosol, 3.7% from plasma membrane, 3.5% to extracellular vesicular exosome, and 2.9% from cytoplasm (**Figure 27B**). Similarly, when classified by protein function in extracellular vesicular exosome, 67.7% mediate protein binding, 66.2% derive from cytoplasm, 58.4% derive from membrane, and 39.9% from nucleus (**Figure 27B**). Of 4058 unique Annotations, 10% (445/4058) receive ≥ 10 annotations times while 8% (325/4058) of Proteins are annotated ≥ 10 functions (i.e. 90% are unique).

Figure 27

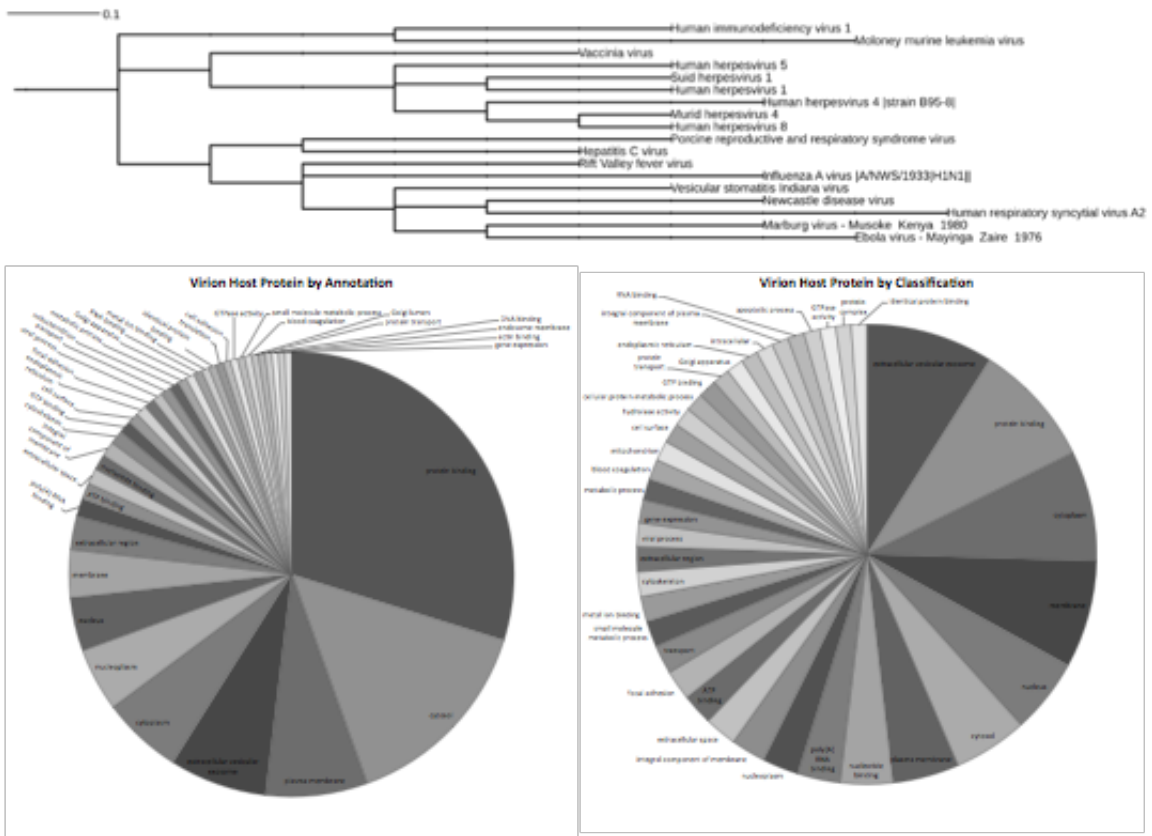


Figure 27 A. Taxonomic representation of extracellular enveloped viral particles whose global protein composition has previously been characterized. The data represent Genera spanning DNA viruses, RNA viruses with DNA intermediates, positive-, and negative-sense RNA lipid enveloped viruses. B. Function and derivation of host protein composition of extracellular virions by QuickGO analysis.

Augmenting the dataset

The global findings from LCMS-based proteomic analysis of extracellular virus are complemented by numerous host-virus studies taking direct measure of specific cellular protein(s). These observations were utilized to augment the virion proteome dataset when the identified protein(s) corroborated the observations of extracellular virus and included the specific identification of host protein(s) incorporated into viral particles or binding to viral proteins verified molecularly by one or a combination of techniques comprised of (co)immunoprecipitation, western blot, or mass spectrometry. At present, the curated dataset consists of 2,021 host proteins (virionproteome.org). A majority 65.0% (1315/2021) of proteins are identified singularly in one viral species where 46.9% (617/1315) of single ID's derive from HIV datasets. Conversely, proteins identified in ≥ 3 instances are found in 18.4% (372/2021) of the dataset and comprise diverse molecular functions. Of the 4,150 GO protein Annotations, no differences are observed in the top 44%. The majority (3,553) of protein Annotations in the dataset are identical. Of the 597 differences in protein Annotation when virion associated host proteins are added, 73% (440/597) are singular. Ten Protein Annotations gain ≥ 10 more Annotations in the complete dataset including those with roles in programmed cell death (32, 3593), pathogenesis (30, 61), small GTPase mediated signal transduction (24, 55), stimulatory C-type lectin signaling pathway (19, not in virion), cytoplasm (15, 5), innate immune response (13, 57), canonical glycolysis (12 not in virion), transcription initiation from RNA polymerase II promoter (11, 640), cytosol (10, 2), and membrane (10, 8). GO Annotation of the full dataset parallels that of the extracellular virus dataset alone,

revealing the continuity of host proteins incorporated into viral particles (**Figure 28**). From the entire dataset 340 proteins are different between virion only while 3,553 proteins are the same when GO annotations are compared. Amongst the virion only proteins that are different from the entire dataset proteins with the most differences include 32 with roles in programmed cell death, 30 in pathogenesis, 24 are small GTPase mediated signal transducers, 19 with roles in stimulatory C-type lectin receptor signaling pathways, 15 are cytoplasmic, 13 involved in innate immune response, 12 in canonical glycolysis, and 11 in transcription initiation form RNA polymerase II promoter (**Figure 28**).

Figure 28

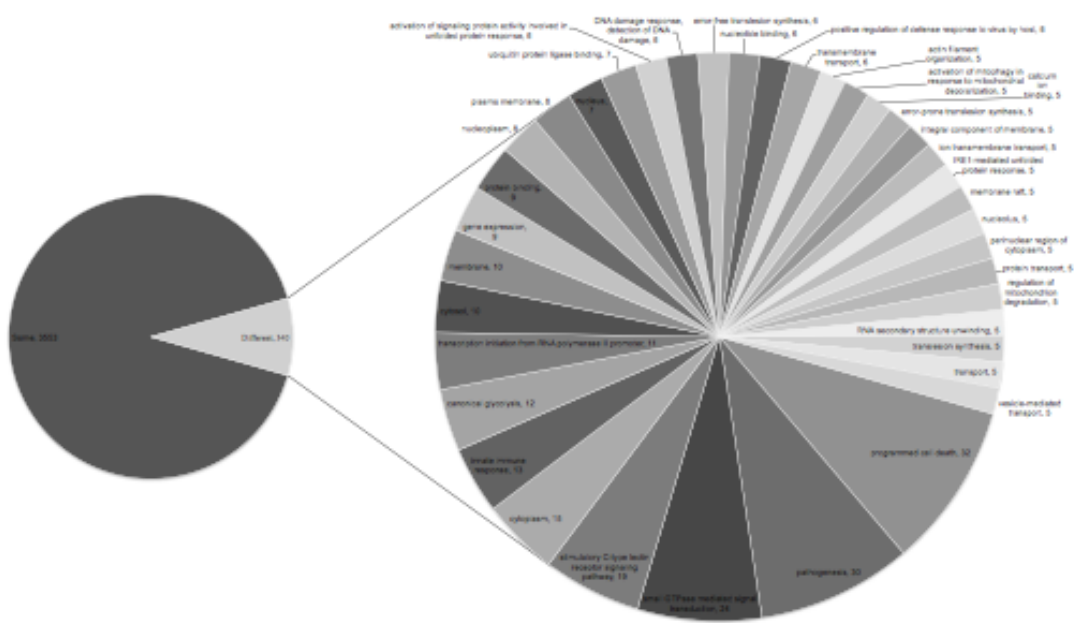


Figure 28 Pie-of-Pie. Difference in GO Protein Annotation between virion-only and VPDB. The majority of Annotations remain unchanged. An additional 593 Protein Annotations are registered from the addition of all proteins identified in or associated with extracellular virus; GO Identifiers with 5 or more additions are shown as the difference in datasets (right).

Lassa Virus-like particles and LCMS analysis

Lassa VLPs prepared in HEK293/T-17 cells were analyzed to confirm the presence of viral protein via western blot (**Figure 29**). In addition to the expected Z product (12 kDa), both the secreted and full-length GPC products and full NP (~65 kDa) were detected (~50 kDa). Equal volumetric quantity VLP and VEC material was prepared and analyzed for protein presence prior to LCMS analysis. In control material there was so little protein that they were below the detection limit of assay detection of 0.02 µg/µl. However, the VLP material contains large quantities of protein, e.g. one analysis registering 0.94 µg/µL, corresponding to a total protein content of 42 µg. Peptide characterization was executed to exclude non-human proteins to remove probable contaminants (e.g. bovine proteins deriving from FBS or trypsin solutions). Proteins of *Homo sapiens* origin with a minimum of 2 positively identified peptides meeting a protein threshold of 95.0% were considered positive hits in VLPs. Representative peptide fragments, spectra, and detected ion species are presented in **Figure 29** for Heat Shock Protein 90α. In total, 22 host proteins were identified in VLPs, 45% of which are ribosomal proteins (Table 7). To confirm the protein observations of LCMS VLP and VEC material were analyzed via WB with antibodies raised against detected host proteins. Ribosomal proteins identified in VLPs via LCMS are detected via WB in VLPs (**Figure 29**). The detection was not exclusive, in the case of 60S ribosomal L3 in the supernatant of VEC material as well but was shown to be specific as antibodies directed against ribosomal proteins not identified with LCMS are similarly not detected via WB, as seen with 40S ribosomal S3 (**Figure 29**). The list of proteins in Table 7 incorporated into Lassa VLPs was added to the dataset of host proteins in extracellular viral particles.

Table 7

Protein Identified	Species	Swiss Prot Number	Molecular Weight (kDa)
Poly [ADP-ribose] polymerase 1	Human	P09874	113
60S ribosomal protein L4	Human	P36578	48
60S ribosomal protein L3	Human	P39023	46
40S ribosomal protein S8	Human	P62241	24
Heterogeneous nuclear ribonucleoprotein U	Human	Q00839	91
Myosin-9	Human	P35579	227
Tubulin alpha-1B chain	Human	P68363	30
Myosin-10	Human	P35580	229
40S ribosomal protein S3a (fragment)	Human	P61246	30
Signal recognition particle subunit SRP72	Human	O76094	75
60 S Ribosomal protein L18	Human	Q07020	22
60S ribosomal protein L14	Human	P50914	23
Heat shock protein HSP 90-alpha	Human	P07900	85
60S ribosomal protein L5	Human	P46777	34
60S ribosomal protein L7a	Human	P62424	30
Elongation factor 2	Human	P13639	16
40S ribosomal protein S4, X isoform	Human	P62701	30
T-complex protein 1 subunit beta	Human	P78371	57
X ray repair cross-complementing protein 5	Human	P13010	83
60S ribosomal protein L18a	Human	Q02543	21
Desmoplakin	Human	P15924	332
DNA topoisomerase 2-alpha	Human	P11388	174

Figure 29

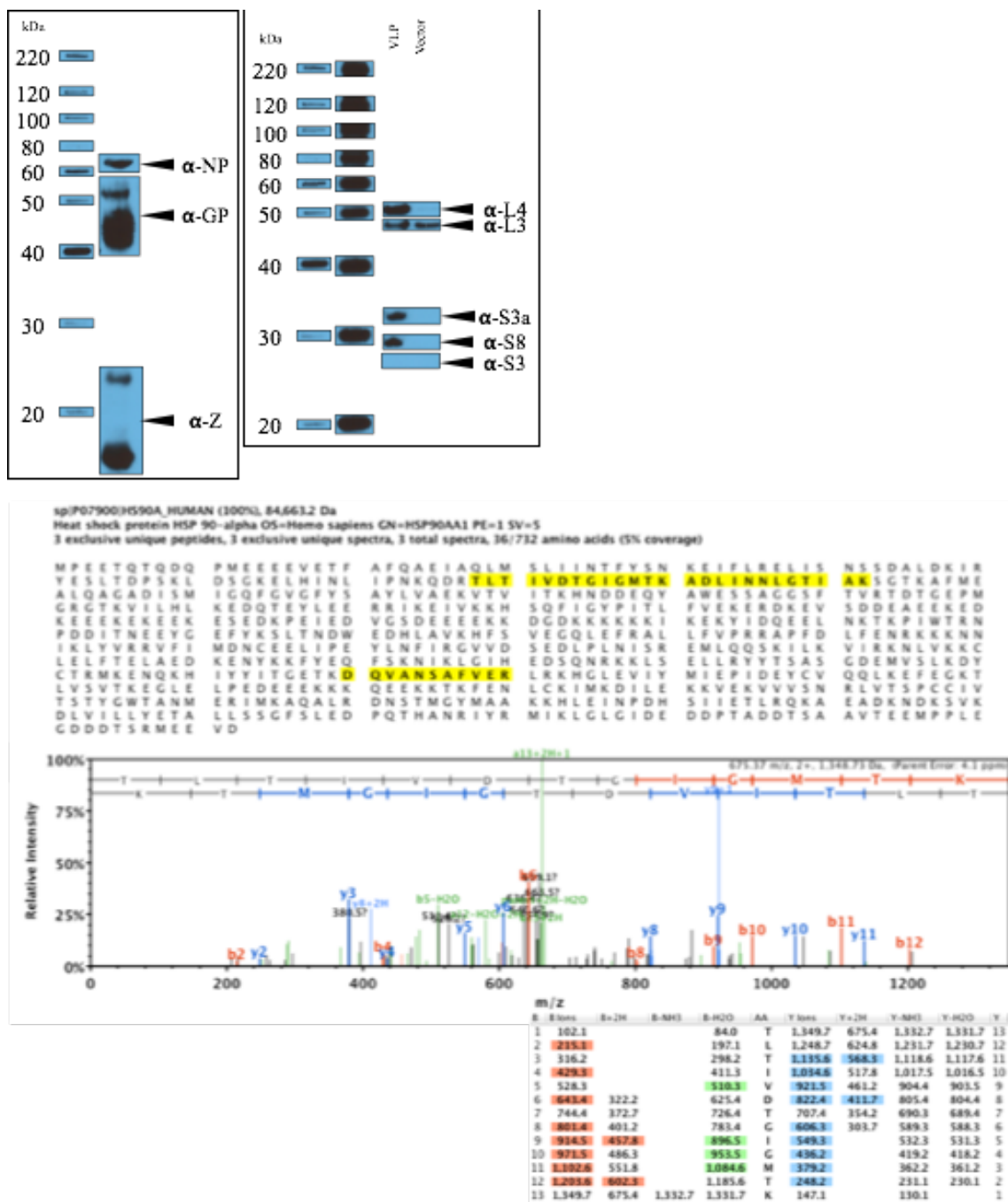


Figure 29 Host Proteins in Lassa Virus-like particles. Western blot analysis of Lassa VLPs showing all 3 expected Lassa proteins are present as are ribosomal proteins detected by LCMS. Ribosomal protein S3 is not detected via LCMS and is not detected via western blot. Representative amino acid sequence, peptide spectra, and ionization profiles are presented for heat shock protein 90-alpha, one of 22 host proteins found incorporated into Lassa VLPs.

Purified Lassa VLPs or specific fractions of singularly expressed Lassa genes where analyzed via western blot targeting ribosomal proteins. In HEK293/T-17 cells transfected with expression vectors for LV GPC, NP, or Z the ribosomal protein S3a staining with molecular weight of approximately 34 kDa is observed most pronounced in Z transfected cells (**Figure 30A, column 6**). In a separate experiment the S3a protein is observed in the extracellular media of GPC in addition to Z transfected cells (**Figure 30B, columns 3, 5**) but is not detected in the media of untransfected control cells or Lassa VLPs (**Figure 30B, columns 6 and 7**). The ribosomal protein S6 staining with a molecular weight of ~32 kDa is detected in an HIV-Lassa GPC pseudoparticle (**Figure 30C, column 6**) and only faintly in the HIV core alone or Lassa VLP (**Figure 30C**). Ribosomal S3a is detected in Lassa VLP but not in pcDNA3.1+intA control transfected material (**Figure 30D**) or in Lassa VLP but less pronounced in Z alone (**Figure 30E**).

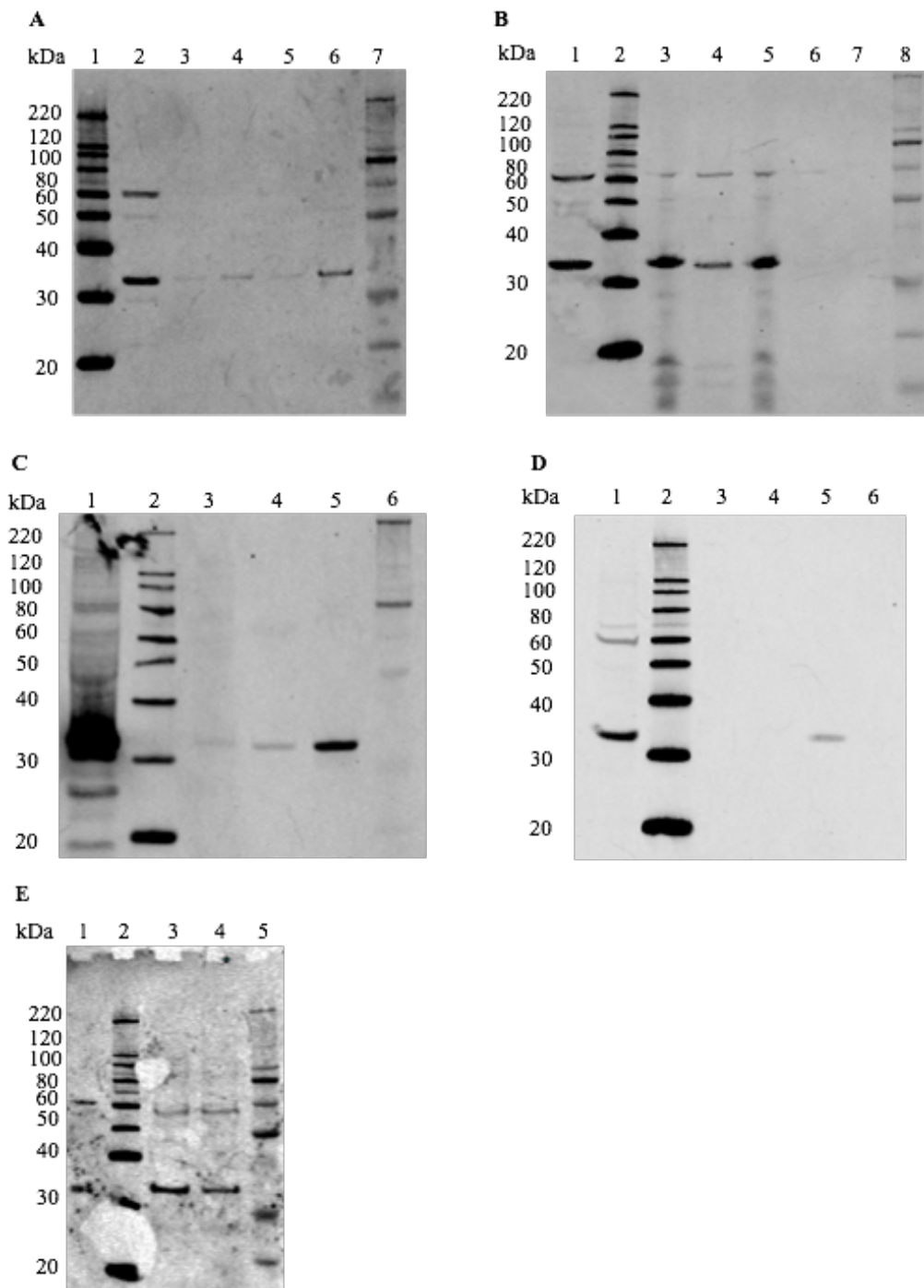


Figure 30

Figure 30 Host Proteins in Lassa VLPS. A WB against Ribosomal protein L4 (~50 kDa) and S3a (~34 kDa); 1. Magic mark ladder, 2. HEK lysate, 3. Untransfected control, 4. Lassa GPC, 5. Lassa NP, 6. Lassa Z, 7. Novex Prestain ladder. **B** WB against Ribosomal protein S3a (~34 kDa); 1. HEK lysate, 2. Magic mark ladder, 3. Lassa GPC, 4. Lassa NP, 5. Lassa Z, 6. Untransfected control, 7. Lassa VLP, 8. Novex Prestain ladder. **C** WB against Ribosomal protein S6 (~32 kDa); 1. HEK Lysate, 2. Magic mark ladder, 3. Lassa VLP, 4. HIV SG3 Core, 5. Lassa GPC pseudoparticle, 6. Novex Prestain ladder. **D** WB against Ribosomal protein S3a; 1. HEK lysate, 2. Magic mark ladder, 3. Untransfected control, 4. pcDNA3.1 + intA transfected control, 5. Lassa VLP, 6. Lassa Z. **E** WB against Ribosomal protein S3a; 1. HEK lysate, 2. Magic mark ladder, 3. Lassa VLP, 4. Lassa Z, 5. Novex Prestain ladder.

The Virion Proteome Database

The aggregated data generated from the observations of virally incorporated host proteins was compiled into a single dataset, denoted The Virion Proteome Database (VPDB). Cluster analysis of host protein representation in extracellular viral particles suggests patterns of conserved incorporations and suggests proteins highly likely to be incorporated into viral particles they have not been previously screened for (**Figure 31**). A web-based platform was implemented to offer multiple analysis and data visualization options including an interactive world cloud based on protein summation, a chord plot illustrating percentage shared host protein from a source virus to other viruses in the database, a Venn diagram comparison tool generating interactive Venn diagrams with lists of proteins found in the selected viruses, and the ability list protein by virus or virus by protein (**Figure 31**). Interactive comparative tools are available at www.virionproteome.org for the entirety of the current dataset.

The relatedness of enveloped viral particles as a function of host protein composition reinforces the patterns observed by summation of total protein profiles when taken as a percentage of shared proteins between 2 viral species. Host protein incorporation and interaction with HIV is the most comprehensively characterized virus at present and on average 62% of proteins found in all other enveloped viruses referenced are also found in or associated with HIV (**Figure 31**). Similarly, genetically related viruses have shared host protein incorporation, e.g. 85% of host proteins found in extracellular Marburg viruses are also found in Ebola virus but constitute only 9% of the total protein content/interaction of Ebola virus (Chord Plot). In addition to the aforementioned analysis tools, summation of the proteins found in the VPDB based on 3-letter gene prefix reveals protein families most frequently incorporated into extracellular viral particles (**Figure 32**). Unsurprisingly histone (50 unique prefixes) and ribosomal proteins (69 unique prefixes) are found with the highest frequency though multiple protein families are of interest, including heat shock proteins (22 unique prefixes). In total, the VPDB contains information spanning diverse viral species, which can be used to develop novel hypotheses. Testing these hypotheses may produce results in one species yet translatable to other distantly related virus(es), such as the identification of new antiviral compounds.

Figure 32

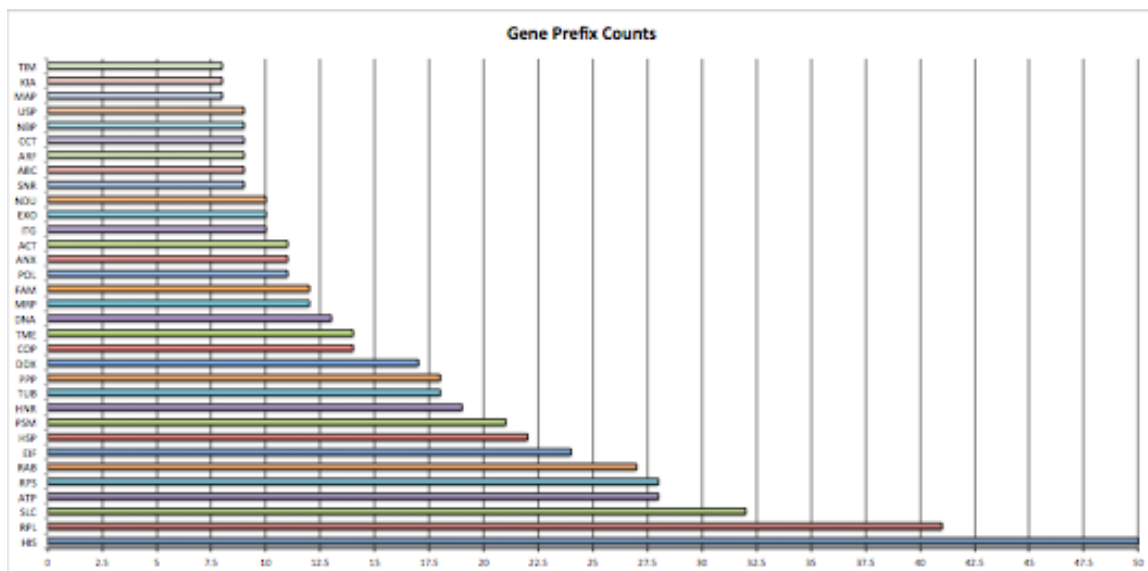


Figure 32 Proteins by group. Proteins as a group are represented by summed gene prefix.

High(er) throughput screening for novel antiviral activity

Molecules identified to have a defined molecular activity on proteins observed with high frequency in the VPDB were selected to screen for differential activity on viral synthesis. A scalable HIV PP encoding a luciferase gene expressed upon infection of a susceptible cell was utilized to screen compounds for the ability to modulate viral synthesis. Inhibition of viral replication was assessed via transfection with vectors to generate infectious PP followed by addition of candidate inhibitor in dilution series. Pseudoparticle infectivity was subsequently assessed and luminescence quantified. Cellular metabolism of alamar blue was used to assess endpoint toxicity of compounds. Withaferin A (**Figure 33**), TMN355 (**Figure 34**), T-56 (**Figure 35**), Pifithrin- μ (**Figure 36**), and CGP 3466B maleate (**Figure 37**) were screened against the PP system. Withaferin A showed significant synthesis inhibition with a $p = 0.001$ at the lowest concentration of $0.3 \mu\text{M}$ but cellular toxicity in the PP producing cell (HEK293/T-17) and assay cell (MDCK) closely mirrored antiviral profile and the vehicle was significantly lower with a $p = 0.03$ where DMSO concentration is the same as in the highest compound concentration. TMN355 significantly inhibited infectious PP at the second lowest concentration of 0.6 nM with a $p = 0.001$ but no significant reduction or toxicity at any other concentration. T-56 significantly inhibited infectious PP only at $0.7 \mu\text{M}$ with a $p = 0.04$ and was not significantly toxic at any concentration. Pifithrin- μ showed significant reduction of infectious PP at $25 \mu\text{M}$ with $p = 0.002$ but at any lowest dilutions no inhibitory activity was present while at higher concentrations significant toxicity is observed. CGP 3466B maleate showed increasing inhibition of infectious PP

starting at 6.25 μM with a $p = 0.02$ down to 100 μM with a $p = 0.0006$ with only mild toxicity in the PP producing cell with 48% viability at 100 μM but no toxicity in assay cells. Modulating host protein dynamics with characterized small molecules exerts an antiviral activity on a HIV PP and in select cases does so without overt toxicity.

Figure 33

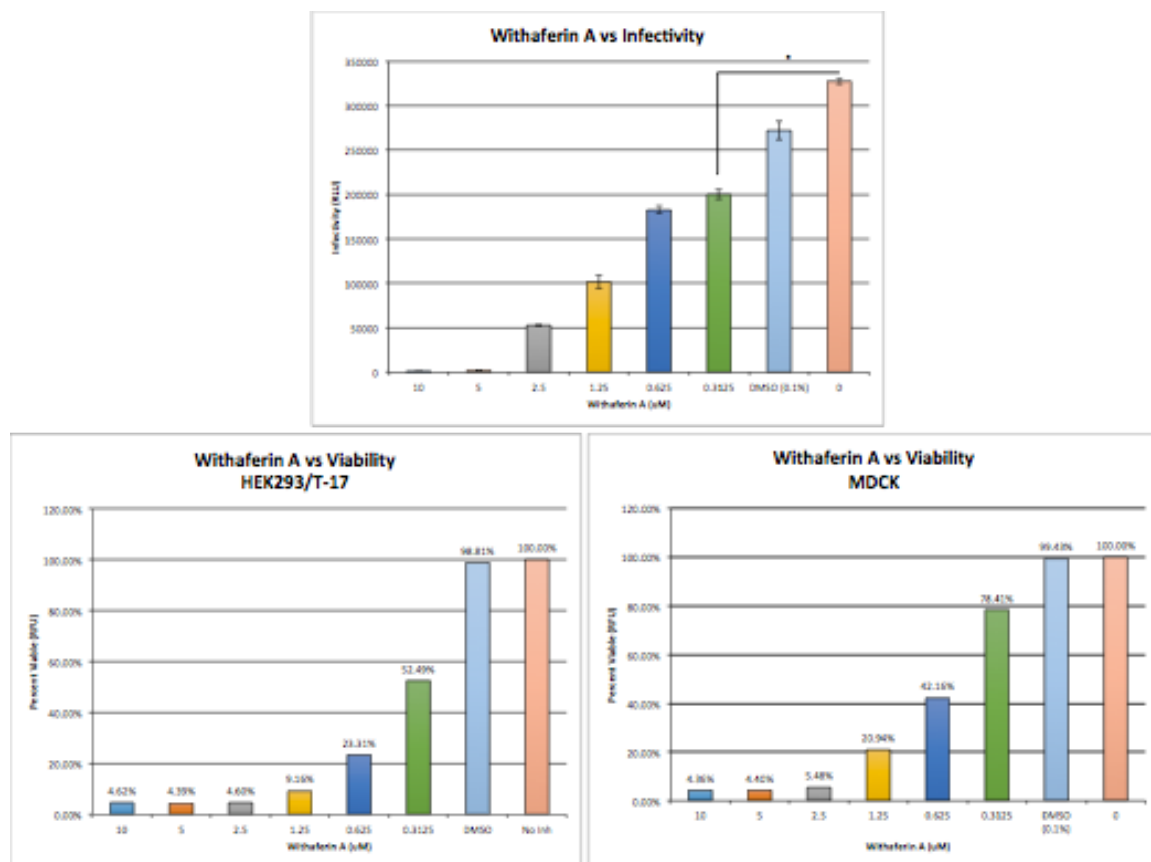


Figure 33 Withaferin A screen. Withaferin A shows dose-dependent and significant inhibition of an HIV pseudoparticle. Withaferin A was also shown to be a potent cytotoxin. N = 3, *p = 0.001, error bars = SEM.

Figure 34

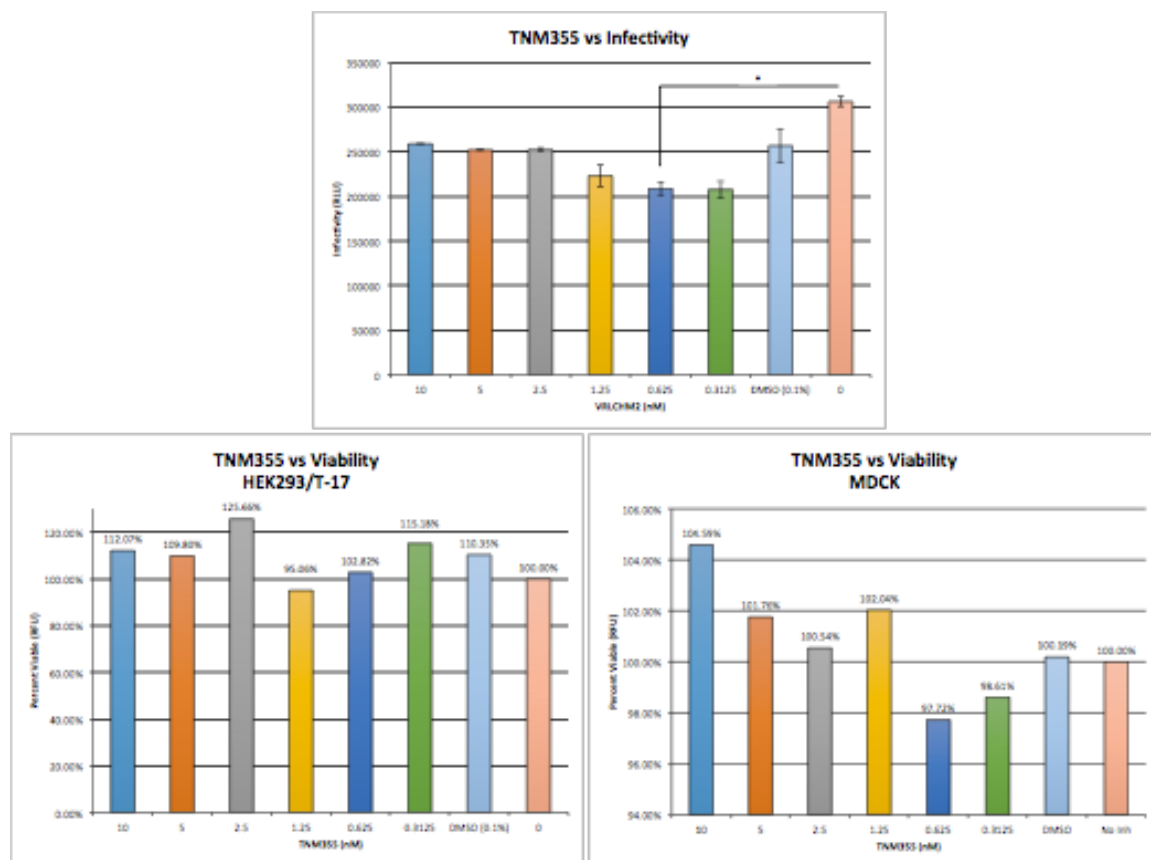


Figure 34 TNM355 screen. TNM355 shows significant inhibition of an HIV pseudoparticle at only one concentration. TNM355 was not toxic to cells. N = 3, *p = 0.001, error bars = SEM.

Figure 35

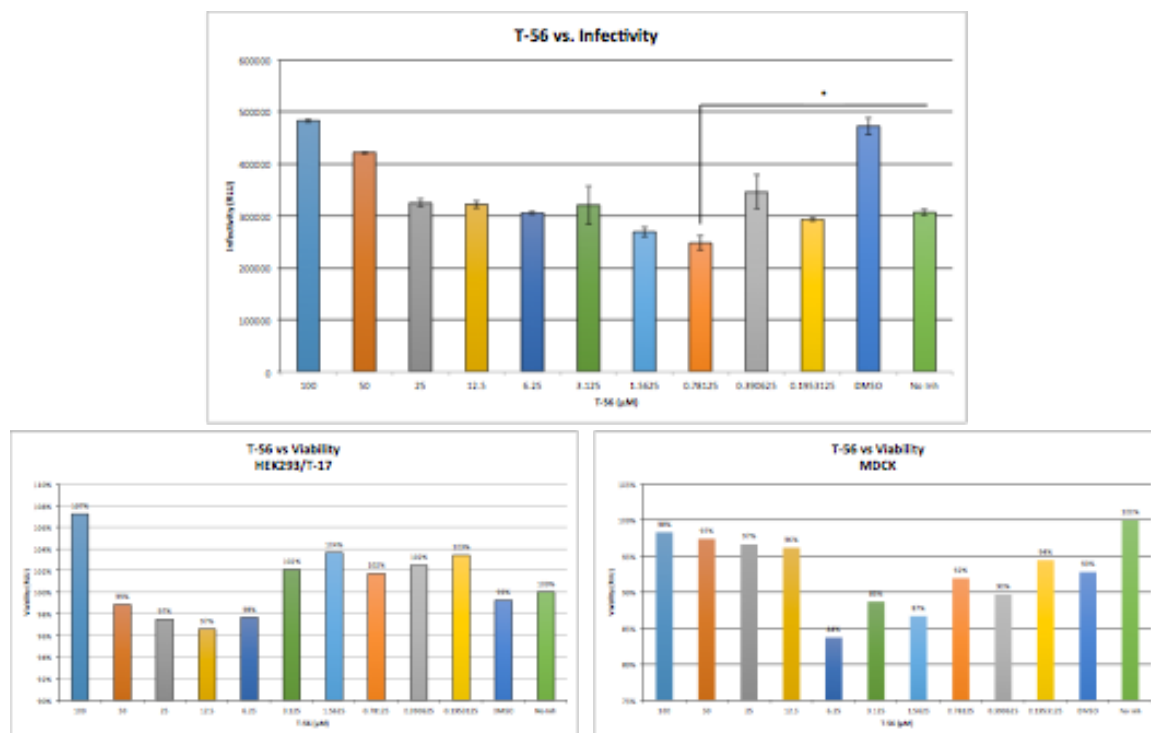


Figure 35 T-56 screen. T56 shows significant inhibition of an HIV pseudoparticle at only one concentration. T56 was not toxic to cells. N = 3, *p = 0.04, error bars = SEM.

Figure 36

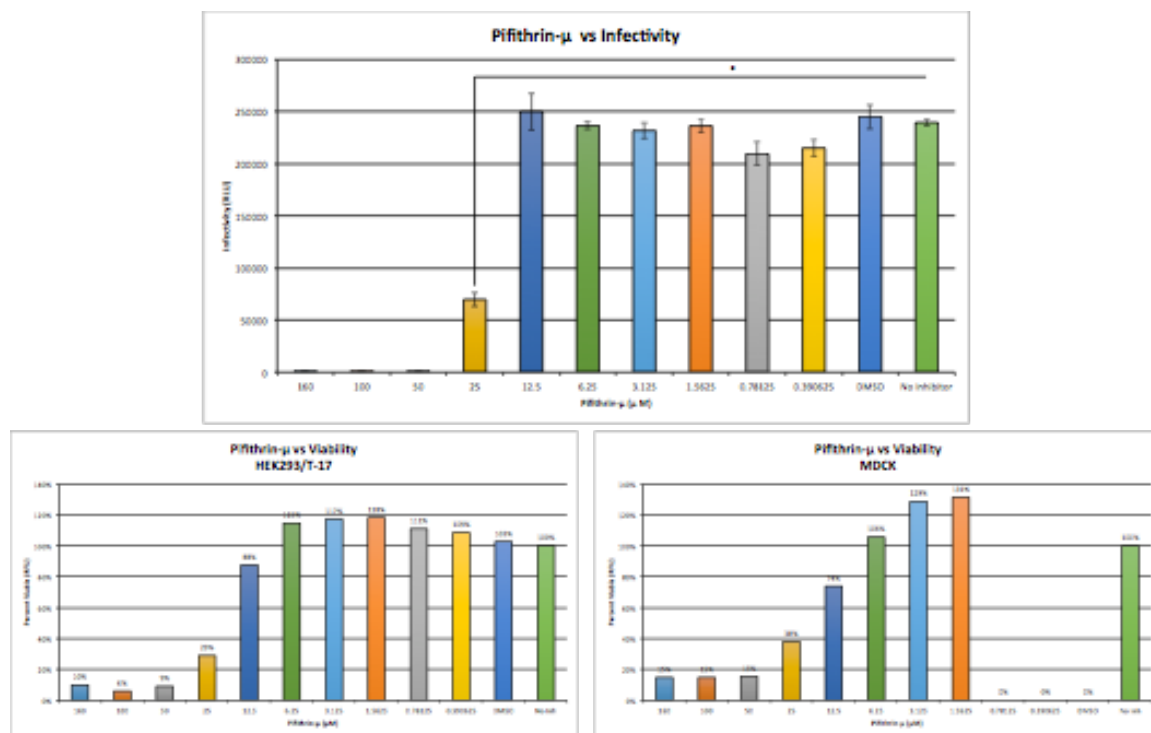


Figure 36 Pifithrin- μ screen. Pifithrin- μ shows significant inhibition of an HIV pseudoparticle at and below 25 μ M. Pifithrin- μ was also toxic to cells at the same concentrations. N = 3, *p = 0.002, error bars = SEM.

Figure 37

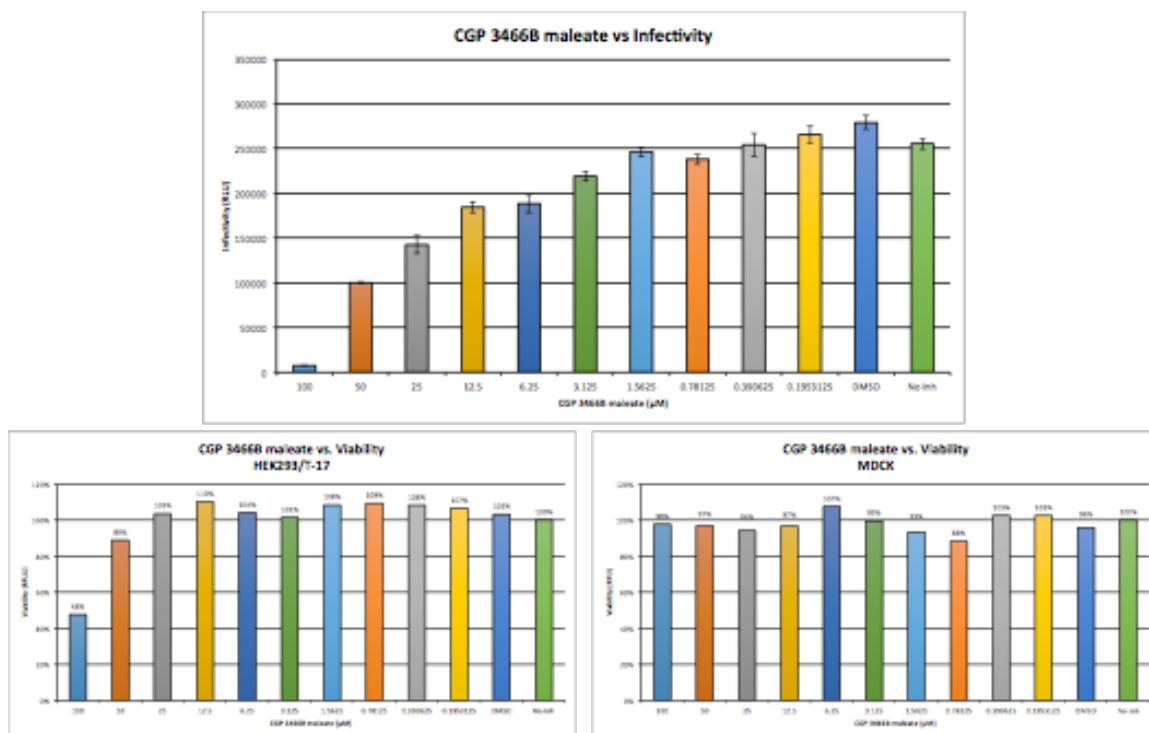


Figure 37 GP 3466B maleate screen. CGP 3466B maleate shows dose-dependent and significant inhibition of an HIV pseudoparticle at at 3.25 [μ]M and all higher concentrations. CGP 3466B was only toxic to HEK293/T-17 cells at the highest concentration of 100 [μ]M. $N = 3$, $p \leq 0.02$ from 6.25 to 100 [μ]M, $p=0.08$ @ 3.125 [μ]M but $p=0.01$ @ 1.5 [μ]M, error bars = SEM.

p24 ELISA

The production of PP was assessed via quantification of the viral p24 protein by Enzyme linked immunosorbent assay (ELISA). Between Withaferin A, TMN355, T-56, and Pifithrin- μ the only significant reduction in extracellular PP was in virus produced at the highest concentration of Withaferin A at 10 μ M with a $p = 0.05$ and Pifithrin- μ at 100 μ M with a $p = 0.02$ (**Figure 38**). The observed antiviral activities therefore may not be a result of gross PP reduction.

Figure 38

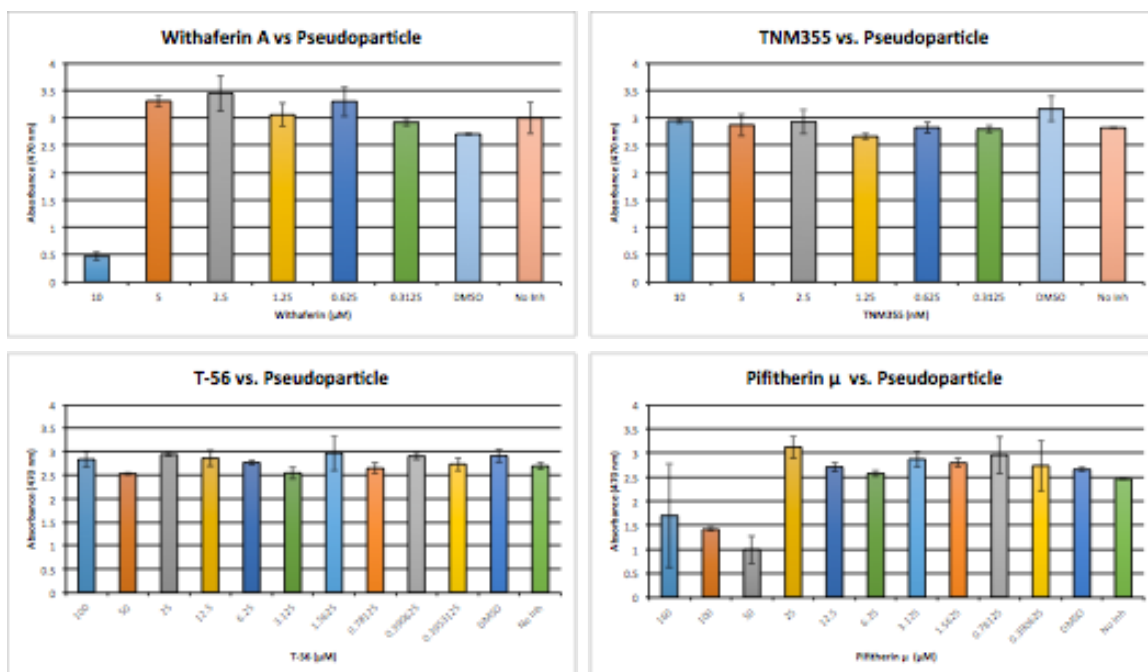


Figure 38 p24 ELISA. Total extracellular virus was quantified as absolute concentration of viral p24 protein. In only two instances is there significantly less extracellular virus, at 10 [μM] Withaferin A and 50 [μM] Pifitherin-μ. N=2, error bars = SEM.

Pre-incubation and carryover toxicity

Three molecules with antiviral activity in the PP infectivity screen were selected to assay for differential activity on PP activity when the cells are preincubated with candidate inhibitor, opposed to addition of inhibitor 4 hours post-transfection.

Preincubation with Withaferin A at 0.3 μM or TNM355 at 0.3 nM had no significant effect on the infectivity of resultant PP while preincubation with 0.7 μM Pifithrin- μ resulted in a significantly less infectious PP with $p = 0.04$ (**Figure 38**). All pre-incubations judged statistically against equal concentration vehicle (DMSO) control.

Withaferin A was selected to control for carryover toxicity. Pseudoparticle was produced in the presence of 2.5 μM Withaferin A, equal concentration vehicle, or complete media. In parallel producer cells were cultivated in the presence of 2.5 μM Withaferin A to generate a metabolically processed spent media (S.M.). Prior to infectivity assay PP was concentrated via centrifugation and resuspended in fresh complete media (F.M.) or S.M. No matter what the media, PP produced in the presence of Withaferin A were uniformly significantly lower (**Figure 39**). At 2.5 μM Withaferin A infectivity is significantly lower than PP produced with equal concentration DMSO with $p = 0.006$ or PP produced in complete media with $p = 0.002$. Concentrated PP grown in 2.5 μM Withaferin A and resuspended in F.M. is significantly lower than PP produced in complete media with $p = 0.0005$ and compared to PP produced with equal concentration DMSO with $p = 0.0007$. Finally, PP concentrated and resuspended in S.M. is equally significantly less infectious than PP produced in complete media with $p = 0.006$ or PP produced with equal concentration DMSO with $p = 0.003$. In two instances

PP produced in the presence of DMSO is significantly less infectious than PP produced in complete media when concentrated and resuspended in F.M. $p = 0.03$ and S.M. $p = 0.02$, possibly a function of concentration as there is no significant difference between DMSO control and PP produced in complete media un-concentrated with $p = 0.06$.

Figure 39

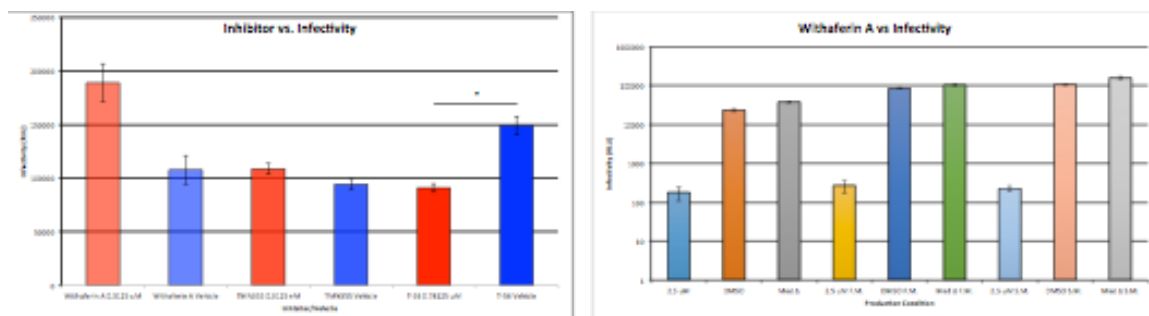


Figure 39 Preincubation and carry over toxicity control. Cells were incubated with inhibitor prior to transfection, with only a significant effect observed with T-56 at the screening concentration showing inhibition. There was no observed toxicity carry over effect from Withaferin A when virus was concentrated and media replenished with fresh media or spent media at 2.5 [μ]M Withaferin A. N=3, $p = 0.04$, error bars = SEM.

Screening hits in competent viral systems

Molecules identified as having antiviral activity in the PP screen were assayed at a concentration range consistent with observed activity in PP with competent viral systems in plaque reduction assay or plaque forming unit reduction assay. For plaque reduction assays, permissive cell types were infected with a uniform quantity of pre-titered virus, incubated and an overlay containing molecule or no molecule as a positive control was added. Plaques were quantified after an incubation appropriate for the virus. Withaferin A inhibited the influenza H3N2 virus where 2.5 μM molecule resulted in a significant reduction in resultant plaque formation with a $p = 0.04$ (**Figure 40**).

Figure 40

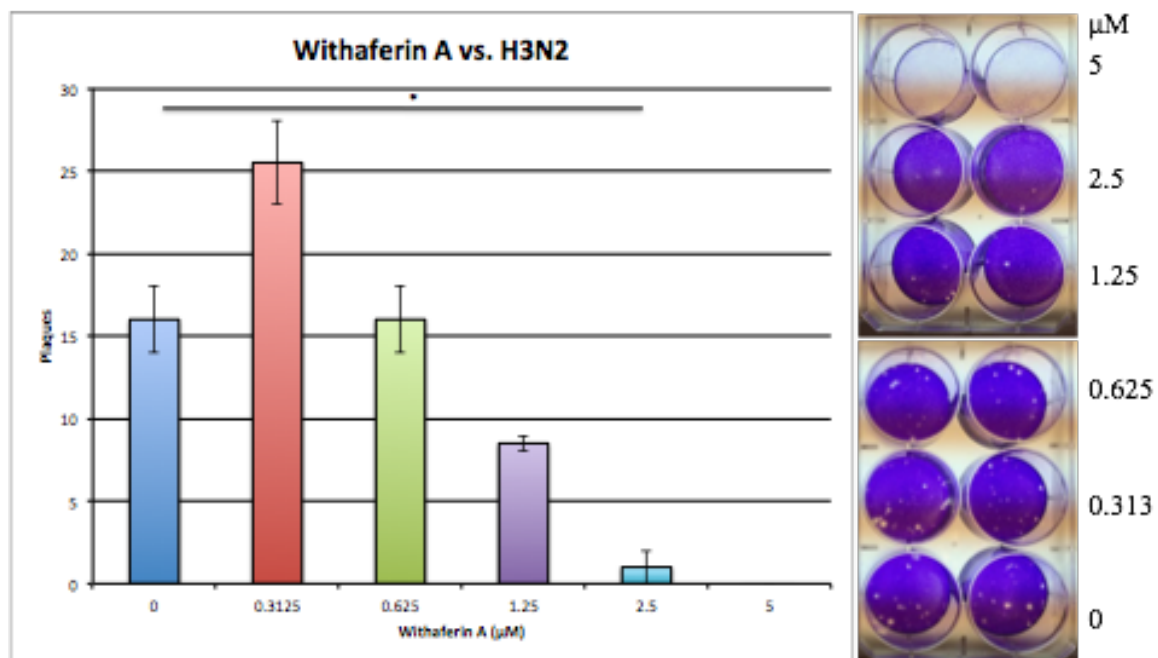


Figure 40 Withaferin A and H3N2. Withaferin A caused significant reduction of Influenza plaques at 2.5 [μM]. N = 2, * p = 0.04, error bars = SEM.

Dengue Serotype 2 (DS2) was assessed for plaque reduction assay with Withaferin A, TMN355, and T-56 (**Figure 41**). At 0.3 μM Withaferin A causes significant reduction in DS2 with $p = 0.008$ but not at 0.6 or 1.25 μM . Withaferin A is toxic to the assay cell type (VERO-E6) at 2.5 μM and above. At 2.5 μM TMN355 causes significant reduction of DS2 with $p = 0.01$. At 5 μM TMN355 has no difference in plaques. T-56 has no significant activity on DS2 at any concentrations.

Figure 41

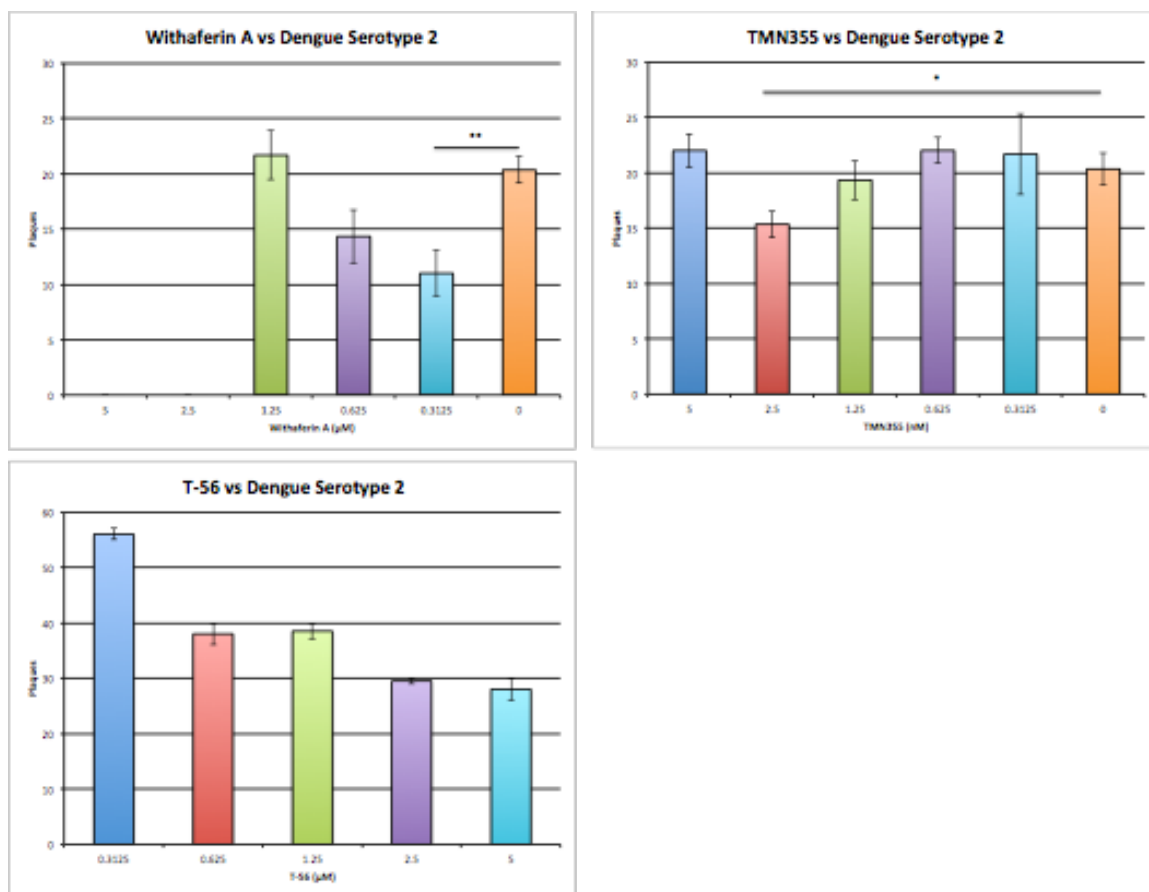


Figure 41 Dengue screening. Dengue serotype 2 was assayed for antiviral activity of Withaferin A, TMN355, and T-56 at concentrations with observed activity in the pseudoparticle screen. N = 3, *p=0.01, **p=0.008 error bars = SEM.

In separate experiments CGP 3466B maleate significantly reduced DS2 plaques (**Figure 42**). In the first at a concentration of 12.5 μM reduction with a $p = 0.008$ is observed. More plaques are present at 25 and 50 μM though still significantly lower at 50 μM with a $p = 0.01$ compared to control. In a repeat experiment significant reduction is observed at 12.5 ($p = 0.04$) and 25 μM ($p = 0.02$). At 6.25 μM $p = 0.05$; more plaques are present in 50 μM while CGP 3466B maleate is toxic at 100 μM .

Figure 42

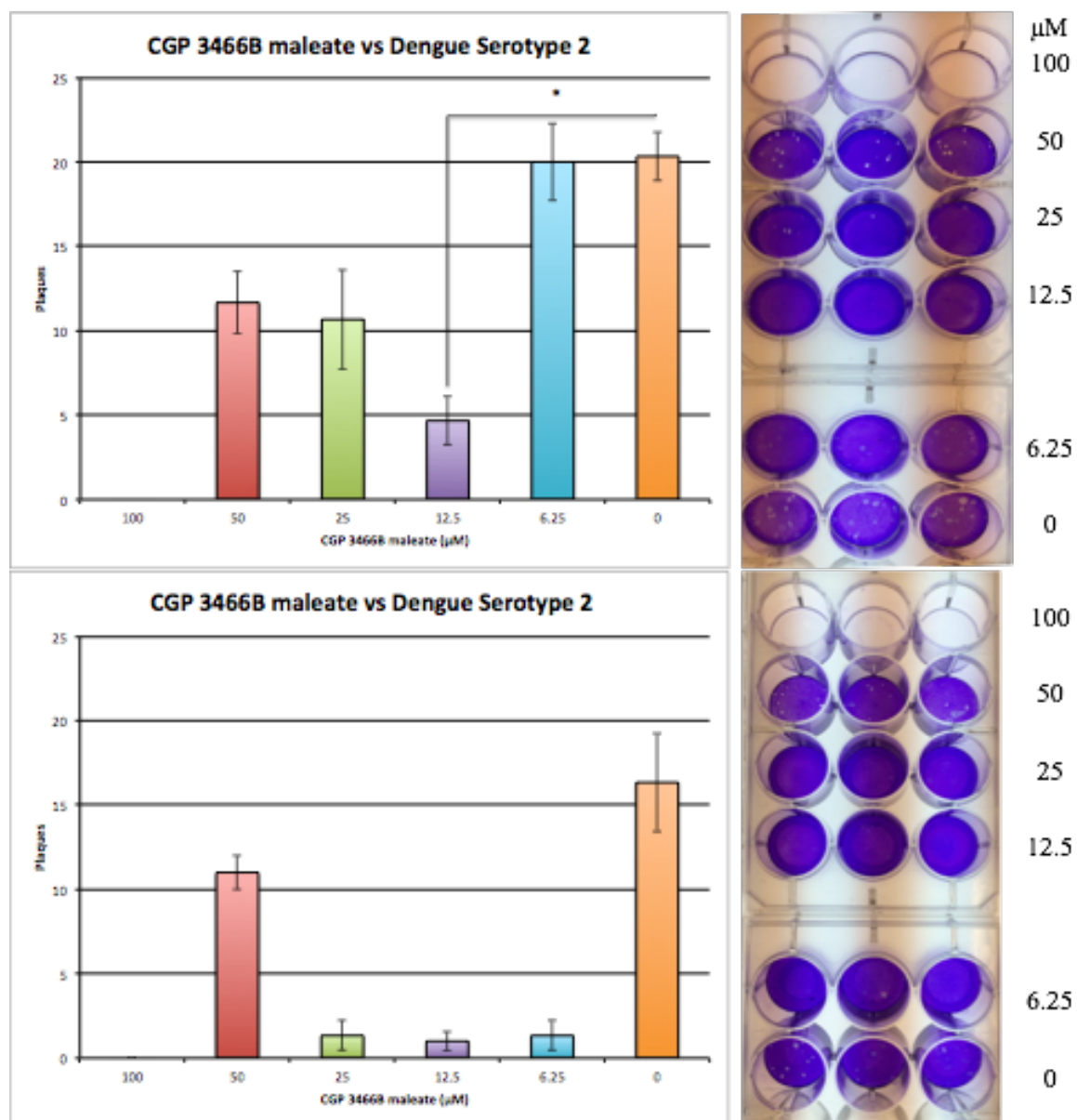


Figure 42 Dengue and CGP 3466B maleate. Dengue serotype 2 was screened twice for activity of CGP3466B maleate at concentrations yielding inhibition in the pseudoparticle screen. In two separate experiments 12.5 [μM] resulted in significant plaque inhibition. $N = 3$, $*p=0.01$, $**p = 0.008$, error bars = SEM.

In plaque forming unit reduction assay's a predefined quantity of DS2 was cultivated in the presence of complete media amended with Withaferin A or CGP 3466B maleate, harvested, diluted to 10^{-3} and assessed via plaque assay. At 0.625, 1.25, and 2.5 μM Withaferin A caused significant reduction in progeny DS2 virions compared to virus cultivated in DMSO control or complete media with $p = 0.01$ for all concentrations (**Figure 43**). At 25 and 50 μM CGP 3466B maleate caused significant reduction in progeny DS2 virions compared to controls with $p = 0.01$ at both concentrations. Both Withaferin A and CGP 3466B therefore exhibit antiviral activity against Dengue Serotype 2.

Figure 43

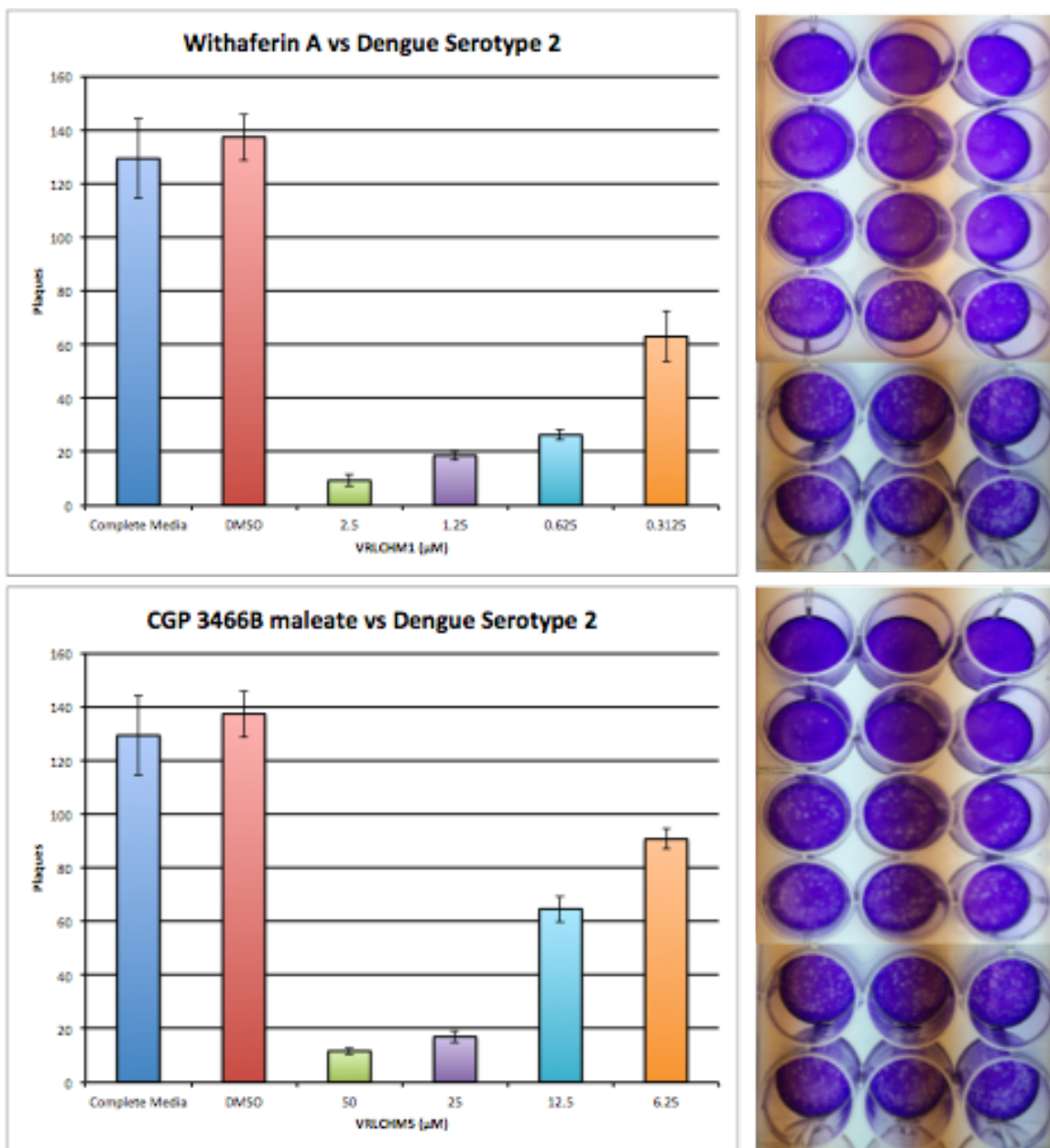


Figure 43 Dengue PFU reduction. Dengue serotype 2 was screened for activity of Withaferin A and CGP3466B maleate in PFU reduction assays where a set concentration of Dengue virus was cultivated in the presence of the indicated inhibitor and progeny virus quantified via plaque assay. Withaferin A significantly reduced the number of PFU at 0.625, 1.25, and 2.5 [μ]M while CGP3466B significantly reduced PFU at 50 and 25 [μ]M. N = 3 and error bars = SEM.

DISCUSSION

Aim 1

The nondescript febrile presentation of viral hemorrhagic fever belies the severity of terminal disease. Multiple flu-like symptoms characterize VHF[25], indifferent from multiple endemic diseases of infectious etiology[27]. Lassa fever features symptoms and mortality rates consistent across VHFs, a substantial endemic burden exceeding several hundred thousand patients, exportation[131], and is arguably the archetype VHF. For LF multiple strategies to assess the state of disease have been employed including the detection of viral antigen and monitoring immunoglobulin response[44]. Recently primary viremia has been shown to harbor prognostic potential for Ebola hemorrhagic fever[132]. Further, characterization of limited but clinically relevant metabolic markers correlates with disease outcome[44, 46]. To expand on the biomarker potential from routinely sampled clinical medium an untargeted LCMS analysis of the comprehensive composite of small molecules of the serum of LF patients was performed on a sample set representative of the global patient population triaged to the Lassa Fever Ward, Kenema Government Hospital (Tables 1 and 2, **Figure 8**). The serum metabolomes of these patients are homogenous when examined on the basis of sero-status (**Figure 9**). When samples are classified based on disease status rather than sero-status the metabolomes delineate groups more clearly (**Figures 11 and 13**). These groupings suggest that within the serum small molecule features exist at levels capable of differentiating between patients suffering febrile illness of non-Lassa etiology from those patients suffering more

severe Lassa fever stages, namely acute and terminal (**Figure 13**). The distinct grouping of FNL verse patients with antigenemia is important for efforts to translate these observations as the mix of FNL patients is clearly heterogeneous and given the validity of modern LF antigen diagnostics[25]. Loss of plasma volume in LF has been observed, supporting vascular leakage as a primary pathology of LF[39] and may be illustrated in the general pattern of loss in terminal patients where 97% (349/360) features with $p \leq 0.05$ and fold change ≥ 2.0 have lower spectral intensity than febrile non-Lassa. Differences in global metabolites suggest that virally induced physiological dysregulation is dynamic within the circulatory system of LF patients and that a population of molecules with biomarker character are embedded within the data analyzed.

One caveat to acknowledge prior deriving conclusions from the detailed dataset is the potential for coinfection and subsequent comorbidity. Sierra Leone is rife with infectious diseases causing flu-like symptoms[27], which are not largely uncharacterized in the current dataset aside from knowing the rate of antibiotic or antimalarial administration (Table 2). Therefore the grouping of FNL patients may actually represent a patient population with differential illness (more or less acute than LF). Two important points must be addressed. The first is the need for further endemic controls of healthy, non-Lassa (no active or historical exposure). Analyzing this group against the current groups will reveal a stationary endemic metabolome to help differentiate the divergence of FNL patients triaged to the LFW. Second, additional diagnostic information for LF patients will help delineate those suffering non-Lassa febrile illness (e.g. acute malaria) and those likely to be suffering acute LF. Retroactive reclassification of the current samples may offer one method to multiplex the current dataset and reclassify samples to

account for comorbidity. An additional concern is the dependence upon patients accurately recording/recalling the time of symptom onset (days since symptom onset, Table 2). Therefore, an ideal solution will be the identical analysis of non-human primate (NHP) models of LF. The disease course and manifestations recapitulates excellently in NHP systems and will allow for defined time points to be sampled, positively identifying features dysregulated during the course of LF. The confounding factor of coinfection notwithstanding, the molecules identified in this study estimate a foundational structure off which to explore and build a prognostic indicator to map the trajectory of LF.

Attempts were executed to positively confirm serum small molecules identified in this study with authenticated standards. Small molecules that were screened against authenticated standards where the observed spectral or retention time characteristic did not match significant unknowns in this LCMS study include betaine aldehyde, 1-aminocyclopropane-1-carboxylic acid, and glyceric acid (glycerate). Pure phosphoglyceric acid was undetectable with the implemented HILIC method utilized and was therefore ruled out as a feature present in the current dataset. Phosphoglyceric acid itself was not identified but rather was a downstream metabolic product in the glycine, serine, and threonine metabolic pathway. Multiple features were identified or possessed m/z 's closely matching members of this metabolic pathway including phosphohydroxypyruvic acid and 3-phospho-D-glycerate. Authenticated standards of phosphohydroxypyruvic acid and 2-phospho-D-glycerate were not readily available from reliable commercial sources, necessitating examination of downstream metabolites. The highly significant feature Unknown 2 with m/z 187.0693 (**Figure 21**) for example was postulated to be the protonated adduct of either 3-phospho-D-glycerate or 2-

phospho-D-glycerate given the absolute abundance of the feature and these metabolites role in the pentose phosphate metabolic pathway. Betaine aldehyde, another metabolite of glycine, serine, and threonine metabolism was similarly examined as the METLIN[93] entry for this molecule gives predicted m/z closely mirroring the Unknown 1 with H^+ of 102.0537 (**Figure 21**). Betaine aldehyde produced good spectra but when tolerances were sufficiently tight for the expected signal-to-noise ratio typical of the Agilent 6538 betaine aldehyde produced m/z 102.119 and therefore is not the Unknown 1. Further confirmation was that betaine aldehyde ran with an rt between 10 and 11 minutes whereas Unknown 1 had an rt of 15.95. Another small molecule with predicted H^+ m/z of 102.0550, 1-aminocyclopropane-1-carboxylic acid was analyzed from authenticated standard. As with betaine aldehyde, 1-aminocyclopropane-1-carboxylic acid was examined as a candidate for Unknown feature 1. Spectral data of 1-aminocyclopropane-1-carboxylic acid matched exceedingly close to the m/z of Unknown 1 but eluted with a retention time. This confirmed that Unknown 1 likely has a molecular formula similar to that of 1-aminocyclopropane-1-carboxylic acid of $C_4H_7NO_2$ whereas betaine aldehyde was $C_5H_{11}NO$. Due to the differences in retention time Unknown 1 is likely not 1-aminocyclopropane-1-carboxylic acid though further investigation of matrix effects are needed to conclusively rule it out and/or delineate alternate atomic structure. Liquid chromatography purification of samples the molecule is present in focused on mobile phases conditions of the retention time that Unknown 1 elutes could produce enough sample for fractionation and Nuclear Magnetic Resonance (NMR) studies. The Unknown 2 feature may represent an irregularity of the subset of FNL patients examine here rather than LF as it is not significantly different in a second sample set (not shown). Modified

LC conditions utilizing a single column to ascertain data in the second sample set may also explain the disparity. Molecules detected in both datasets feature shifted retention times uniformly eluting 2-3 minutes sooner in the second run and those with low abundance are not detected at all. As such as the protonated adduct of Unknown 1 (m/z 102.0537) is also not detected in the second sample set while the ammoniated adduct is (**Figure 23**). It is well appreciated that the bottleneck of metabolomics studies is positive identification of molecules[60] where hundreds of features may be of interest but it is not scalable to test each individual authenticated pure standard. These difficulties notwithstanding, the ability to detect, differentiate, and investigate the global metabolite profile in multiple disease states offers unique avenues of discovery with verifiable utility.

The employment of metabolomics for investigating LF will be most useful in the pursuit to identify and validate prognostic biomarkers as validated, well-performing molecular diagnostics already exist[25, 48]. However, no molecular method to monitor the trajectory of LF has been developed. A means to predict the disease course may be particularly appreciated in countries where LF is endemic and attending physicians need to allocate scarce resources to those in most dire need. Coupling prognostic biomarkers to validated ELISA detection of humoral immune responses may further refine patient populations as those febrile but singularly positive for anti-Lassa IgG are exceptionally similar to FNL patients (**Figures 9 and 10**). Alternatively patients with LV antigen form distinct groups and experience severe metabolic dysregulation (**Figure 10**), most acute and all terminal LF (**Figure 12**). Spectral intensity data of candidate prognostic features harbor differential nature of these features in the current Lassa fever dataset (**Figure 26**).

The patients whom ultimately perish with positive Lassa diagnostics form a distinct group, indicating that the input variables possess differential character (**Figure 13**).

Serum lipids were the molecular class most frequently putatively identified from LF samples. Clinically derived primary Dengue Infection also found serum lipids as a molecular class frequently dysregulated as a pathology[81]. To approximate the relationship between groups, cluster analyses were performed computing divergence of standard deviation from the mean. Complimentary computational methods following iterations of train, trim, and retrain yield refined feature lists where only those features with the most powerful statistical predictive power remain (**Figure 14**). The RF analysis ranks features on their computed predictive power, which can be used to remove poorly performing features before reanalyzing with cluster analysis (**Figure 14**). In **Figure 14A** when the load data include all 160 putative lipids dendograms indicate FNL and those surviving LF as the most distantly related groupings. Random forest analysis was executed on the entire 160 putative lipid dataset, undergoing a train, trim, and retrain iterative process. This analysis identifies the most powerful features, that when reanalyzed yield more distinct group-to-group relationships. When poorly performing lipid data are removed the survive/convalescent and FNL groups are the most related while those patients whom perish and are Ag/IgM/PCR⁺ are most related (**Figure 14C**). Cluster analysis of lipids from 5 sample groups revealed dysregulation in Terminal patients to be most unique as measured by deviation from the group median spectral intensity, forming a distinct clade (**Figures 14 and 15**). Unsurprisingly, FNL and convalescent sera were the most closely related upon cluster analysis of both lipid and peptide species (**Figure 15**) potentially an indication that convalescent patients are

returning to a state more closely to FNL which may itself more accurately reflect homeostasis. As mentioned previously, the examination of healthy endemic sera is necessary to estimate homeostasis. Only 3 samples were positive for the presence of Lassa antigen, anti-Lassa Immunoglobulin M, and viral genome but were included as they seemed to represent a transition between the most acute stages of LF and transition towards relative homeostasis. Within the lipids data, Platelet-Activating Factor (PAF), its metabolic precursors (**Figure 16**), and 9 additional glycerophosphocholine PAF-like molecules (**Figure 17**) were putatively identified or identified via manual screening for matching H^+ and Na^+ m/z and conserved rt (Table 5).

The present LF data reveal a relative measure of serum PAF/Lyso-PAF where PAF possesses ~10X more signal intensity and presumably ~10X more abundant in the patient. This ratio is important given that PAF is thought to be responsible for the inflammatory cascade. Gross inflammatory dysregulation such as the hyper-inflammatory “cytokine storm” noted in Ebola patients[133], influenza, and bacterial sepsis[134] is also mediated by PAF[135]. A pattern exists for all PAF-like molecules with uniformly decreased signal intensity in terminal patients (**Figures 16 and 17**). Platelet-activating factor and its associated metabolic products have been examined in multiple clinical virology studies. Serum PAF is found elevated while the PAF-degrading acetylhydrolase (PAF-AH) activity is diminished in Hepatitis C patient sera[136]. In the disease most closely related to LF—Dengue fever—PAF appears to be a contributing factor to vascular leakage[137] and higher expression of serum PAF-AH correlates with lower frequency of Dengue fever but not Dengue hemorrhagic fever in two ethnically distinct populations[138]. In a murine model of Dengue genetic knockout or chemical inhibition

of the platelet-activating factor receptor (PAFR) resulted in a less severe disease and increased survival in those animals deficient or inhibited for PAFR[139]. In a murine model of Arenavirus infection with Lymphocytic Choriomeningitis Virus (LCMV), Stearoyl lysophosphocholine (18:0) was identified with positive ion m/z 524.37057 having reduced signal intensity in the serum of infected animals similar to PAF C-16 in LF patients detected here with m/z 524.3659[70]. It should be noted that given the sensitivity of LCMS it is unlikely that features with the aforementioned m/z are the same but it was nonetheless worth noting given that the study of LCMV is the only other metabolomics investigation of an arenavirus. Future LCMS experiments with authenticated standards are necessary to verify the PAF species characterized here.

The observed loss of serum PAF in LF has an inverse relationship with viremia, which correlates with increased mortality[44, 132, 140]. Platelet-activating factor activates platelets, leukocytes, and monocytes/macrophages[141] where leukocytes and granulocytes are dominant cells expressing PAF receptor[142, 143]. It has previously been shown that LF patient platelet activity is depressed, particularly so for terminal patients[29, 41]. The loss of PAF in terminal patient sera mirrors this pattern, potentiating a mechanism. In terminal LF patients secreted PAF may rapidly bind its receptor or remain expressed on the surface of the synthesizing cell. Alternatively, the observed loss may be a result of Lassa virus replication in PAF producing cells, particularly endothelial cells[144, 145] resulting in dysregulation of the PAF cascade [lower secretion] as disease progresses. Vascular leakage may also explain the loss of PAF. Differential activity (or lack thereof) of PAF-AH may also account for PAF/Lyso-PAF loss. PAF-AH metabolizes PAF to Lyso-PAF and therefore monitoring of PAF-AH activity presents a

proxy measure of serum PAF. Commercially available reagents exist requiring relatively low technologies available in facilities such as the clinical lab of the LFW in KGH (Cayman Chemical No 760901). Serial assay of serum PAF, Lyso-PAF, and/or PAF-AH may present a worthwhile prognostic target with Lyso-PAF exhibiting a Sensitivity and Specificity of 97% and 92%, respectively, in this machine learning analysis (Table 6).

Hemoglobin breakdown products are significantly lower in the blood of terminal LF patients, again potentiating as a prognostic indicator. An obvious primary question is whether readily available urinalysis sticks can be adapted as a LF prognostic. The first limitation will be establishing a baseline as there is relatively little-to-no Urobilinogen normally in the urine[146]. Therefore establishing a baseline by which to serially monitor for Urobilinogen loss and thus poor prognosis, is difficult. As is the case with any metabolite identified here, a specific ion monitoring LCMS method can be developed with a pure standard and monitored as such. Though potentially available, the technology is cumbersome and difficult to reliably operate in a resource starved setting in addition to the rigorous testing necessary to validate a clinically relevant LCMS-based test. Utilization of economical, commercially available urinalysis sticks with serum prepared as above may afford a cheap, rapid prognostic requiring low technology and little training of clinic personnel. The process of evaluating the potential to adapt urinalysis sticks will require comparison of whole sera verse plasma verse metabolite preparation such as methanol extraction detailed above. Organic extraction appears to be a viable method to purify the targeted molecule and reduce background/interfering blood products. A necessary starting point is the need to quantify via standard addition LCMS experiment the actual amount of Urobilinogen in the present samples, i.e. the quantity of

Urobilinogen in 100 μ L prepared and extracted serum where urinalysis sticks detect 0.2 – 1.0 mg/dL Urobilinogen. Despite Urobilinogen's poor specificity (67%), the 94% sensitivity and aforementioned cheap, scalable, and validated measure present a strong case for further investigation as a prognostic. An advantage is further illustrated in the fact that in other infectious etiologies such as infective hepatitis and malaria urine Urobilinogen is elevated. Parallel observations in the serum in conjunction to a LF diagnostic may facilitate monitoring the trajectory of LF.

The protonated, sodiated, and potassiated adducts of the modified nucleoside 1-methylinosine are significantly elevated in the serum of terminal LF patients. Elevated abundance is notable given the dominant pattern of signal loss in terminal sera compared to febrile non-Lassa. Despite the fact that the protonated adduct is the least significant ($p = 0.05$) it is the strongest marker in machine learning analysis (Table) with a specificity of 75% and sensitivity of 93%. This molecule has previously been identified as elevated in the urine of cancer patients [147] and in the plasma of patients in renal failure whose removal possesses biomarker utility for effective hemodialysis [148]. Further, ELISA reagents directed against 1-methylinosine have been developed [149]. As with the above detailed biomarkers, 1-methylinosine may present a viable prognostic indicator of LF if utilized in parallel with LF diagnostics.

Many features are present in the dataset beyond those detailed above that merit further investigation. For example, a feature significantly elevated in terminal LF patient sera identified as L-alpha-Aspartyl-L-hydroxyproline with m/z 247.0908 closely matches γ -glutamyl-Valine with observed m/z 247.1297 and diminished in the plasma over the course of LCMV infection in mice [70]. The present study seeks a first approximation of

the dynamic and differential character of the serum metabolome of clinical LF patients compared to febrile non-LF. Small molecule biomarkers with high translational potential have been described, with an emphasis on those for which commercially available reagents already exist. Recent advances in the diagnosis of VHF[25, 48], spurred by the presence and persistence of viral diseases such as Ebola and LF, afford the clinician reliable determination of the presence of VHF agent in a febrile patient but no single or group of biomarkers have been uniformly implemented as clinical prognostics, which again, is itself not a trivial task given the rigorous validations necessary to prove clinical significance. Multiple studies have sought to elucidate prognostic biomarkers but none have taken the global view of the serum metabolome. Small molecules such as those detailed above whose absolute abundance is dysregulated as a result of viral infection and replication processes offer a phenotypic media to assay and assess the trajectory of disease.

Aim 2

It has long been appreciated that HIV-1 incorporates host factors into the extracellular virus[85] a phenomena which has been previously reviewed[150] and compiled to an open-access web-accessible database[151]. Proposed roles for cellular proteins in multiple viral species have also been reviewed[152]. Further, efforts have been focused on development of informatics tools accounting for the comprehensive characterization of all host-virus protein-protein interactions[153-155]. These tools are built off evidence that viral proteins have hundreds of cellular interacting partners, some of which are specifically or incidentally packaged in the virion. It is worth noting there can be an extreme difficulty in discriminating between genuine virion particles and

contaminants such as exosomes[156, 157]. Most studies included in the present analysis make efforts such as the analysis of VEC control material—accounting for both transfection-mediated toxicity, cellular aging, and shedding—to verify the authenticity of viral materials (**Figure 29**). The efforts detailed above have been made to account for the total host protein representation of enveloped viruses. These data have been summarized and employed to generate an additional informatics resource: the virion proteome database, a constitution of viral and *extraviral* proteins present in extracellular enveloped viruses. Aggregation of this data—especially so for accessible, high global impact viruses such as HIV or Influenza—and extrapolation of this information to medically relevant, low global impact viruses such as LV may facilitate the identification of conserved targets for antiviral therapy.

Examples abound of characterized affiliation and/or incorporation of host factors into enveloped viral particles preceding the wide implementation of LCMS-based global protein characterization. These studies have typically utilized genetic detection systems and antibodies targeting the host protein(s) identified. Epstein Barr and HIV-1 were shown to interact with or incorporate cytoskeletal components[12, 13]. Yeast two-hybrid screening facilitated early delineation of host factors interacting with viral proteins; subsequent incorporation into virions can be validated with immune-specific methodologies. These assays revealed incorporation of—amongst others—elongation factor 1-alpha[14], Vacuolar Protein Sorting 4B[15], TSG101/VPS28[16], and SUMO-1[17], into nascent HIV-1 particles. Additionally, human TSG101 was shown to be recruited by the HIV-1 Gag protein where the virus exhibits dependence on TSG101 for replication and infectivity[18, 19]. The nucleoprotein of Influenza A virus (IAV) was shown to

interact with karyopherin alpha 1 (formerly nucleoprotein interactor 1) and incorporated into infectious virus[20] while the non-structural protein 1 was shown to interact with human Staufen protein, also found in extracellular virus[21]. Hepatitis B Virus X protein was found to interact with DNA damage-binding protein (formerly XAP-1)[22] while Hepatitis C was shown to interact with and package Heterogeneous Nuclear Ribonucleoprotein K[23]. Cellular stress driven by viral infection revealed the regulation and molecular interaction of Heat Shock Protein 70 in Vaccinia virus infection[24]. Coimmunoprecipitation targeting the viral transactivation molecule p40tax of Human T-Cell Leukemia Virus revealed two host factors interacting with the viral protein, one of which was identical to heat shock protein 60-kDa[25]. Yeast two-hybrid screening followed by coimmunoprecipitation revealed the Respiratory Syncytial virus fusion glycoprotein to specifically bind Transforming protein RhoA and [26] while Rubella virus nucleocapsid was revealed to bind the mitochondrial protein p32 (now C1QBP)[27] and Rabies polymerase cofactor P protein to bind Signal Transducer and Activator of Transcription 1[28]. These interactions offer examples of viral subversion and consequent affinity for binding host proteins, shedding insight on the various pathways viruses use for replication or defense from host inflammatory response. In most (all) cases these factors have been identified as interacting with viral particles/proteins using global characterization technologies.

To augment this data with medically important viruses for which special containment is necessary, we have performed an LCMS-based proteomics investigation of Lassa virus-like particles. Multiple host proteins present in other distantly related viral species are also found packaged in Lassa VLPs. A primary observation of particular

importance is the detection of NP (~65kDa, **Figure 29**). The LV NP possesses no translocation mechanism; therefore the only means by which it is detected in an extracellular fraction is through an active export process. Through Z-driven biogenesis the coordinated translocation, incorporation, and export of NP is facilitated indicating that Lassa VLPs possess egress mechanisms of native virus[17] (**Figures 2 and 4**). It can therefore be assumed that VLPs interact with cellular partners in a manner akin to native virus (**Figure 4**) a process shown necessary for proper assembly/egress[90]. Detection of ribosomal proteins in Lassa VLPs (Table 7 & **Figure 29**) is a significant observation as it has been known or assumed for decades that organelles staining with high contrast upon electron microscopy of Arenaviruses are ribosomes but conclusive links are scarce[20]. Lassa VLPs have been visualized and appear to possess dark staining organelles but detection of ribosomal RNA species was unsuccessful[18]. Attempts were made here to characterize the catalytic activity of ribosomes affiliated with Lassa VLPs. Adaptation of an *in vitro* translation system employing a (+) Control gfp mRNA and human cell lysate resulted in no fluorescent product beyond background in VLP fractions (not shown). Multiple methods were attempted to liberate VLP-bound ribosomes including variable sonication, and low concentration detergent exposure but resulted in no gfp signal. Given the subset of ribosomal protein present in Lassa VLPs, including the lack of molecular detection of ribosomal proteins also not detected by mass spectrometry (**Figure 29**) the possibility exists that only partial ribosomes affiliate with VLPs. Given the molecular complex of a ribosome composed of RNA and protein RNase protection assays were executed as well. When exposed to a RNase possessing buffer or when enzymatic assay established with purified RNase H (NEB) recovered total RNA quantity was

approximately 10-fold higher than VEC control material (not shown). Though anecdotal these observations suggest a protection effect potentially of ribosomal proteins of ribosomal RNA (rRNA) species; genetic detection via Reverse Transcriptase Polymerase Chain Reaction or quantitative Polymerase Chain Reaction did not conclusively identify rRNA (not shown). Attempts were also made to fractionate multiple protein populations from Lassa VLPS. The comprehensive protocol generated sought to isolate soluble, membrane associated, and integral membrane proteins from VLPs using a combination of physical (sonication and centrifugation) and chemical (basic sodium carbonate exposure) means. The protocol was executed multiple times but only integral membrane proteins resulted in appreciable protein recovery sufficient for further characterization. The association of ribosomal proteins with Lassa VLPs was further investigated to delineate which viral protein mediates these interactions (**Figure 30**). Western blotting for specific ribosomal proteins from different transfections with LV genes revealed two potential viral proteins mediating interaction with host ribosomes. First, it was observed that Lassa VLPs and not VEC control material contained ribosomal proteins (**Figure 29**). The prepared supernatant of cells transfected with Lassa Z alone appear to harbor ribosomal proteins (**Figure 30A, B, E**) but not always (**Figure 30D**). The prepared supernatant of cells transfected with Lassa GP also stained for ribosomal proteins (**Figure 30A, B, C**) and the presence of GP in an HIV pseudoparticle results in the presence of ribosomal proteins where they otherwise are not (**Figure 30C**) and in the VLP enhances ribosomal staining signal (**Figure 30E**). These findings are not wholly surprising as both the GP and Z proteins possess extracellular transport signals and both are sufficient to drive VLP biogenesis[18, 158]. The increase in signal intensity of ribosomal protein staining

coupled to endoplasmic reticulum trafficking of LV GP suggests that the LV glycoprotein binds ribosomal proteins or incorporates rough endoplasmic reticulum membrane (**Figure 2**) into the virion. These molecular observations corroborate the findings of LCMS analysis of Lassa VLPs where multiple host proteins—some known, some novel—in the VLP highlight the conserved nature of host representation in viral particles.

Global proteomic analysis of enveloped viral particles has been previously performed on multiple viral species. When the observations of those viruses are combined with the observations in Lassa VLPs patterns begin to emerge. Cluster analysis across the entire dataset highlights the most common single protein incorporations (**Figure 31**, red lines). This may suggest proteins most likely to be incorporated in viral particles that are previously unexplored. For example heat shock protein 90AA1 is identified 13 unique viruses a total of 18 times (VPDB) and has been shown to be a targetable protein to inhibit viral replication[159, 160]. Heat shock protein 70 and 90 are members of the Dengue virus replication complex in vitro, potentiating these proteins as targets for antiviral therapy[161]. The various displays of virion proteome data offer unique and complementary views to compare host protein representation in diverse viral species (**Figure 31**). An alternate view of common protein incorporations is to account for protein groups via summation of gene prefixes. For example, Annexins such as Annexin A2, A5, and A1 make up 3 of the top 20 most frequently incorporated proteins but 12 different Annexin species are found in total (**Figure 32**). Therefore host proteins as a group frequently incorporated into extracellular viral particles may also shed light on diverse yet conserved viral phenomena. Whether these incorporations are incidental or specific, if modulating/altering the dynamics of host proteins has a differential effect on

viral replication the protein may act as a target for broad-spectrum antiviral development[162].

The HIV based NL4-3 pseudoparticle was identified as an efficient system to test for antiviral activity[103]. This screen (i.e. NL4-3) offers a platform that is cheap, scalable, and easy to manipulate. Furthermore, the exhaustive list of host proteins found in[150] and the known interactions of HIV with host proteins[163] affords a means to capture the largest cross section of characterized interactions to identify hits to advance to more cumbersome classical screening methods. To test the utility of the VPDB a proof-of-concept experimental screen was implemented where 5 compounds were identified as acting on a pertinent host protein with no historical record of testing as an antiviral.

The PPIA peptidylprolyl isomerase A protein was the 4th most frequent host protein identified in extracellular viral particles but inhibition of its activity by TMN355, a potent cyclophilin A active in a nM range[164] had only moderate and non dose dependent activity on PP (**Figure 34**). In a screen with Dengue virus TMN355 significantly reduced plaque formation at 2.5 nM opposed to the 0.78 nM significant inhibition observed in the PP but no plaque reduction was observed at higher concentration (5 nM) or nearest the active range in the PP (0.625 nM). Inhibition can therefore be assumed likely an artifact but TMN355 may merit screening at a concentration of an order of magnitude more (μ M range) given the favorable tolerability by HEK293/T-17, MDCK (**Figure 34**), and VERO-E6 (no toxicity, not shown).

Cofilin 1, a protein with roles in actin polymerization and depolymerization was one of the top 10 most frequently identified proteins. Cofilin is a substrate of LIM kinases and a recently identified small molecule T56-LIMKi was found to decrease

phosphorylated cofilin[165]. In the PP screen significant inhibition was observed at 0.78 μM versus no inhibitor while significantly more PP infectivity was observed at high concentrations of T-56, mirroring observed signal registered in vehicle control (**Figure 35**). There was no difference in the quantity of extracellular PP suggesting no activity on viral synthesis (**Figure 38**). Preincubation of producer cells with 0.78 μM Pifithrin- μ resulted in significant reduction in infectivity compared to PP produced in equal concentration vehicle with $p = 0.04$ (**Figure 39**). In a competent viral system T-56 caused a dose-dependent reduction of Dengue virus plaques but this was not significant when comparing the highest concentration of 5 μM (lowest plaque count) versus the lowest concentration of 0.3125 μM (highest plaque count) with a $p = 0.06$ (**Figure 41**). The lack of a no inhibitor control due to experimental error makes drawing conclusions difficult, necessitating further investigation of the T-56 in competent viral systems.

The molecular chaperone proteins known as heat shock proteins where some of the most frequently identified proteins incorporated into extracellular viral particles including Lassa VLPs. The proteins heat shock protein family A member 8 and heat shock protein family A member A1 are the 1st and 6th most frequently identified host proteins in extracellular viral particles, respectively. The small molecule 2-phenylethanesulfonamide or Pifithrin- μ inhibits the chaperone activity of both heat shock proteins[166]. In the PP screen significant reduction in infectivity was observed at concentrations of $\geq 25 \mu\text{M}$ with a $p = 0.002$ at 25 μM . Significant toxicity was observed in parallel where there was only 29% viability in producing cells cultivated at 25 μM and there was no significant reduction in infectivity at the next lowest concentration of 12.5 μM with $p = 0.63$ while there was 88% viability in producing cells (**Figure 36**).

Interestingly, at 25 μM there was equal quantity virus (**Figure 38**). Therefore Pifithrin- μ merits more in depth investigation at or around a concentration of 25 μM to understand if the significant inhibition observed at 25 μM is the result of carry-over toxicity (transfer of un-metabolized Pifithrin- μ to assay cell type where there is only 38% viability at 25 μM) or if genuine antagonism of viral infectivity is observed.

The first molecule screened was Withaferin A, a steroidal lactone which has been utilized as an antineoplastic with acceptable toxicity profile in an animal model[167]. In addition to inhibition of heat shock protein 90 Withaferin A has been shown to aggregate vimentin molecules[168], two proteins incorporated into extracellular viral particles with high frequency (**Figure 31**). Presently, heat shock protein 90AA1 and vimentin are identified in 13 and 9 distinct viral species and both proteins have been detected in HIV with multiple instances. Withaferin A had a potent and dose dependent inhibitory affect in the PP screen with a $p = 0.001$ at the lowest concentration of 0.3125 μM . The observed reduction of infectivity is mirrored in cellular toxicity in both the producer and assay cell types though at the lowest concentration HEK293/T-17 cells are 52% viable and MDCK are 78% (**Figure 33**). Confounding these results where the total quantity of extracellular PP. At the highest concentration of Withaferin A of 10 μM there was a slightly significant reduction in viral p24 detected ($p = 0.05$) compared to PP produced in the presence of no inhibitor (**Figure 38**) At any other concentration there is equal amount of extracellular PP, suggesting that despite toxicity, PP is produced and quantifiably less infectious in the presence of Withaferin A (**Figure 33**). To control for potential carry-over toxicity where un-metabolized Withaferin A was toxic to assay cells a media replenish experiment was executed. In any condition where PP was produced in the

presence of 2.5 μM Withaferin A there was significantly less infectivity than in PP produced in equal concentration vehicle ($p = 0.006$) or complete media ($p = 0.002$). When PP was concentrated and resuspended in fresh media (F.M.) or spent media from producer cells cultivated in equal concentration Withaferin A (S.M.) PP from Withaferin A maintains significantly reduced infectivity compared to PP produced in vehicle and complete media (**Figure 39**). Importantly, resuspending PP in S.M. had no detrimental effect on infectivity of vehicle or control PP, indicating that the inhibitory effect observed (**Figure 33**) is the result of Withaferin A antagonism of virion infectivity. Precincubation of producer cells with 0.3125 μM Withaferin A did not have any affect on subsequent infectivity of PP compared to vehicle at equal concentration (**Figure 39**). In competent viral systems Withaferin A showed dose-dependent inhibition of Influenza H3N2. At the highest concentration of 5 μM the molecule was toxic to the assay cell while at 2.5 μM the monolayer is intact and plaques significantly reduced compared to no inhibitor control with $p = 0.04$. More H3N2 plaques are present at 0.3125 μM Withaferin A than no inhibitor wells, possibly as a concerted effect of viral replication and toxicity. In Dengue virus Serotype 2 Withaferin A shows significant plaque reduction at 0.3125 μM with $p = 0.008$ (**Figure 41**). Interestingly, at 2 higher concentrations of Withferin A DS2 plaques increase: at 0.625 reduction is just outside of statistical significance at $p = 0.06$ and equal at 1.25 μM . At 2.5 and 5 μM Withaferin A is toxic to the assay cell type (VERO-E6). Viability screening of Withaferin A in VERO-E6 cells corroborates these observations with 53% or less viability at or above 2.5 μM . At concentrations of 1.25, 0.625, and 0.3125 μM Withaferin A VERO-E6 cells exhibit 97 to 100% viability (not shown). In plaque forming unit reduction assay where a predetermined quantity of virus

is cultivated in the presence of Withaferin A and progeny virus quantified by plaque assay Withaferin A exhibits a potent and dose depended inhibitory effect on DS2 (**Figure 42**). At 2.5, 1.25, and 0.625 μM Withaferin plaque forming unit reduction is significant all $p = 0.01$ compared to virus cultivated in complete media (**Figure 42**). At 0.3125 μM plaque forming unit reduction is evident but not significantly less compared to complete media with $p = 0.08$. Collectively these results suggest an antiviral activity of Withaferin A as it significantly reduces the infectivity of an HIV pseudoparticle (**Figure 33**) without reducing gross viral progeny (**Figure 38**) and has a significant antiviral effect on Influenza H3N2 (**Figure 40**) and Dengue virus Serotype 2 in two variations of assaying infectious viral particles (**Figures 41 and 42**). Mechanistic studies are necessary to further characterize the activity of Withaferin A. Primary studies should focus on the interplay between Heat Shock Protein 90, cellular vimentin, and Withaferin A. Due to the reduction of infectivity but not total viral particles evident by p24 ELISA Withaferin A may act as an antagonist of virion infectivity and dysregulation of the chaperone activity of Heat Shock Protein 90 merits further investigation.

The final compound screened was the maleate salt of CGP 3466B, a tricyclic small molecule that has verifiable oral bioavailability and no toxic effects in mice[169] or non-human primates[170]. An idealized inhibitory dose-response response was observed on PP infectivity with increasing concentration of CGP 3466B (**Figure 37**). At a 1.56 μM infectivity is significantly reduced ($p = 0.01$), not at 3.125 μM and significantly lower at 6.25 μM and above. The only toxic concentration in the producer cell is at 100 μM where the cells are reduced to 48% viability. Separate experiments in DS2 yielded conflicting data. In one experiment at 12.5 μM CGP 3466B maleate significantly reduced DS2

plaques with a $p = 0.008$ compared to no inhibitor controls (**Figure 42**). When the experiment was repeated significant inhibition was registered at 12.5 and 25 μM ($p = 0.04$ and 0.02 , respectively). In both cases 100 μM is toxic to the assay cell while at a concentration of 50 μM more plaques are present (**Figure 41**). In plaque forming unit reduction assay CGP 3466B significantly reduces viral progeny at 25 and 50 μM both with $p = 0.01$ versus no inhibitor control (**Figure 42**). As with Witharferin A, further mechanistic studies of the antiviral activity of CGP 3466B are necessary though the evidence suggests CGP 3466B maleate to inhibit virion formation. The parent molecule TCH346 is an inhibitor of glyceraldehyde-3-phosphate dehydrogenase, a protein whose activity has recently been shown increased in the presence of Dengue virus NS1[171] protein and Dengue virus[172]. Glyceraldehyde-3-phosphate dehydrogenase has been found to carry out multiple catalytic functions but one of the most notable is carbohydrate metabolism. Alteration of cellular metabolic activity by Dengue virus may see the increase of metabolic activity to promote viral replication and the inhibition of glyceraldehyde-3-phosphate dehydrogenase may therefore retard replication. In the VPDB glyceraldehyde-3-phosphate dehydrogenase is identified in 8 different viral species in 15 unique instances including multiple HIV annotations. The observed dose-response of CGP 3466B maleate on the HIV pseudoparticle therefore is no surprise and as with Withaferin A, suggests a conserved role for glyceraldehyde-3-phosphate dehydrogenase in viral synthesis which may be a target for antiviral intervention with broad-spectrum activity.

The ability to discover in scalable system and translate to medically relevant viruses requiring special handling an example of how ‘big data’ techniques, i.e. “-omics”

technologies could be rationally employed. High(er) throughput screening of novel antivirals: it is recognized that thousands of small molecules exist whose molecular activity is characterized at some level. Using protein lists generated from observations driven by the VPDB a group of molecules was identified as having some activity on a VPDB protein identified with high frequency and no previous screen for antiviral activity. Assaying these molecule with foreknowledge disregarding the specific activity but acknowledging that modulating of host protein activity may have a differential affect on viral synthesis was used as a rationale to screen for novel antivirals...

REFERENCES

1. Fields, B.N., D.M. Knipe, and P.M. Howley, *Fields virology*. 6th ed. Arenaviridae, ed. M.J. Buchmeier, de la Torea, J.C, Peters, C. J. . Vol. 2. 2013, Philadelphia: Lippincott-Raven Publishers. 20.
2. Cao, W., et al., *Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus*. *Science*, 1998. **282**(5396): p. 2079-81.
3. Oldstone, M.B. and K.P. Campbell, *Decoding arenavirus pathogenesis: essential roles for alpha-dystroglycan-virus interactions and the immune response*. *Virology*, 2011. **411**(2): p. 170-9.
4. Rojek, J.M., et al., *Different mechanisms of cell entry by human-pathogenic Old World and New World arenaviruses*. *J Virol*, 2008. **82**(15): p. 7677-87.
5. Jae, L.T., et al., *Virus entry. Lassa virus entry requires a trigger-induced receptor switch*. *Science*, 2014. **344**(6191): p. 1506-10.
6. Lelke, M., et al., *An N-terminal region of Lassa virus L protein plays a critical role in transcription but not replication of the virus genome*. *J Virol*, 2010. **84**(4): p. 1934-44.
7. Hulo, C., et al., *ViralZone: a knowledge resource to understand virus diversity*. *Nucleic Acids Res*, 2011. **39**(Database issue): p. D576-82.
8. Baird, N.L., J. York, and J.H. Nunberg, *Arenavirus infection induces discrete cytosolic structures for RNA replication*. *J Virol*, 2012. **86**(20): p. 11301-10.
9. Eichler, R., et al., *Lassa virus glycoprotein signal peptide displays a novel topology with an extended endoplasmic reticulum luminal region*. *J Biol Chem*, 2004. **279**(13): p. 12293-9.
10. Lenz, O., et al., *The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/SIP*. *Proc Natl Acad Sci U S A*, 2001. **98**(22): p. 12701-5.
11. Burri, D.J., et al., *Envelope glycoprotein of arenaviruses*. *Viruses*, 2012. **4**(10): p. 2162-81.
12. Schlie, K., et al., *Characterization of Lassa virus glycoprotein oligomerization and influence of cholesterol on virus replication*. *J Virol*, 2010. **84**(2): p. 983-92.
13. Hastie, K.M., et al., *Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression*. *Proc Natl Acad Sci U S A*, 2011. **108**(6): p. 2396-401.
14. Reynard, S., et al., *Exonuclease domain of the Lassa virus nucleoprotein is critical to avoid RIG-I signaling and to inhibit the innate immune response*. *J Virol*, 2014. **88**(23): p. 13923-7.
15. Russier, M., et al., *The exonuclease domain of Lassa virus nucleoprotein is involved in antigen-presenting-cell-mediated NK cell responses*. *J Virol*, 2014. **88**(23): p. 13811-20.

16. Strecker, T., et al., *Lassa virus Z protein is a matrix protein and sufficient for the release of virus-like particles [corrected]*. J Virol, 2003. **77**(19): p. 10700-5.
17. Eichler, R., et al., *Characterization of the Lassa virus matrix protein Z: electron microscopic study of virus-like particles and interaction with the nucleoprotein (NP)*. Virus Res, 2004. **100**(2): p. 249-55.
18. Branco, L.M., et al., *Lassa virus-like particles displaying all major immunological determinants as a vaccine candidate for Lassa hemorrhagic fever*. Virol J, 2010. **7**: p. 279.
19. Fehling, S.K., F. Lennartz, and T. Strecker, *Multifunctional nature of the arenavirus RING finger protein Z*. Viruses, 2012. **4**(11): p. 2973-3011.
20. Borden, K.L., et al., *Two RING finger proteins, the oncoprotein PML and the arenavirus Z protein, colocalize with the nuclear fraction of the ribosomal P proteins*. J Virol, 1998. **72**(5): p. 3819-26.
21. Monath, T.P., et al., *Lassa virus isolation from Mastomys natalensis rodents during an epidemic in Sierra Leone*. Science, 1974. **185**(4147): p. 263-5.
22. McCormick, J.B., et al., *A prospective study of the epidemiology and ecology of Lassa fever*. J Infect Dis, 1987. **155**(3): p. 437-44.
23. Demby, A.H., et al., *Lassa fever in Guinea: II. Distribution and prevalence of Lassa virus infection in small mammals*. Vector Borne Zoonotic Dis, 2001. **1**(4): p. 283-97.
24. Bausch, D.G., et al., *Lassa fever in Guinea: I. Epidemiology of human disease and clinical observations*. Vector Borne Zoonotic Dis, 2001. **1**(4): p. 269-81.
25. Shaffer, J.G., et al., *Lassa fever in post-conflict sierra leone*. PLoS Negl Trop Dis, 2014. **8**(3): p. e2748.
26. Frame, J.D., et al., *Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings*. Am J Trop Med Hyg, 1970. **19**(4): p. 670-6.
27. Boisen, M.L., et al., *Multiple circulating infections can mimic the early stages of viral hemorrhagic fevers and possible human exposure to filoviruses in Sierra Leone prior to the 2014 outbreak*. Viral Immunol, 2015. **28**(1): p. 19-31.
28. McCormick, J.B., et al., *A case-control study of the clinical diagnosis and course of Lassa fever*. J Infect Dis, 1987. **155**(3): p. 445-55.
29. Fisher-Hoch, S., et al., *Hematologic dysfunction in Lassa fever*. J Med Virol, 1988. **26**(2): p. 127-35.
30. Richmond, J.K. and D.J. Baglolle, *Lassa fever: epidemiology, clinical features, and social consequences*. BMJ, 2003. **327**(7426): p. 1271-5.
31. Knobloch, J., et al., *Clinical observations in 42 patients with Lassa fever*. Tropenmed Parasitol, 1980. **31**(4): p. 389-98.
32. Fisher-Hoch, S.P., et al., *Review of cases of nosocomial Lassa fever in Nigeria: the high price of poor medical practice*. BMJ, 1995. **311**(7009): p. 857-9.
33. Stephenson, E.H., E.W. Larson, and J.W. Dominik, *Effect of environmental factors on aerosol-induced Lassa virus infection*. J Med Virol, 1984. **14**(4): p. 295-303.
34. Mahanty, S., et al., *Cutting edge: impairment of dendritic cells and adaptive immunity by Ebola and Lassa viruses*. J Immunol, 2003. **170**(6): p. 2797-801.

35. Baize, S., et al., *Lassa virus infection of human dendritic cells and macrophages is productive but fails to activate cells*. J Immunol, 2004. **172**(5): p. 2861-9.
36. Walker, D.H., et al., *Comparative pathology of Lassa virus infection in monkeys, guinea-pigs, and Mastomys natalensis*. Bull World Health Organ, 1975. **52**(4-6): p. 523-34.
37. Jahrling, P.B., et al., *Lassa virus infection of rhesus monkeys: pathogenesis and treatment with ribavirin*. J Infect Dis, 1980. **141**(5): p. 580-9.
38. Walker, D.H., et al., *Pathologic and virologic study of fatal Lassa fever in man*. Am J Pathol, 1982. **107**(3): p. 349-56.
39. Fisher-Hoch, S.P., et al., *Physiological and immunologic disturbances associated with shock in a primate model of Lassa fever*. J Infect Dis, 1987. **155**(3): p. 465-74.
40. Lange, J.V., et al., *Kinetic study of platelets and fibrinogen in Lassa virus-infected monkeys and early pathologic events in Mopeia virus-infected monkeys*. Am J Trop Med Hyg, 1985. **34**(5): p. 999-1007.
41. Cummins, D., et al., *A plasma inhibitor of platelet aggregation in patients with Lassa fever*. Br J Haematol, 1989. **72**(4): p. 543-8.
42. Cummins, D., et al., *A plasma inhibitor of platelet aggregation in patients with Argentine hemorrhagic fever*. Am J Trop Med Hyg, 1990. **42**(5): p. 470-5.
43. Biomarkers Definitions Working, G., *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework*. Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.
44. Branco, L.M., et al., *Emerging trends in Lassa fever: redefining the role of immunoglobulin M and inflammation in diagnosing acute infection*. Virol J, 2011. **8**: p. 478.
45. McCormick, J.B., et al., *Lassa fever. Effective therapy with ribavirin*. N Engl J Med, 1986. **314**(1): p. 20-6.
46. Grove, J.N., et al., *Capacity building permitting comprehensive monitoring of a severe case of Lassa hemorrhagic fever in Sierra Leone with a positive outcome: case report*. Virol J, 2011. **8**: p. 314.
47. Sewlall, N.H., et al., *Clinical features and patient management of Lujo hemorrhagic fever*. PLoS Negl Trop Dis, 2014. **8**(11): p. e3233.
48. Boisen, M.L., et al., *Development of Prototype Filovirus Recombinant Antigen Immunoassays*. J Infect Dis, 2015. **212** Suppl 2: p. S359-67.
49. McElroy, A.K., et al., *Ebola hemorrhagic Fever: novel biomarker correlates of clinical outcome*. J Infect Dis, 2014. **210**(4): p. 558-66.
50. Hangartner, L., R.M. Zinkernagel, and H. Hengartner, *Antiviral antibody responses: the two extremes of a wide spectrum*. Nat Rev Immunol, 2006. **6**(3): p. 231-43.
51. Baize, S., et al., *Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients*. Nat Med, 1999. **5**(4): p. 423-6.
52. Hutchinson, K.L. and P.E. Rollin, *Cytokine and chemokine expression in humans infected with Sudan Ebola virus*. J Infect Dis, 2007. **196** Suppl 2: p. S357-63.

53. Villinger, F., et al., *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.* J Infect Dis, 1999. **179 Suppl 1**: p. S188-91.
54. McElroy, A.K., et al., *Biomarker correlates of survival in pediatric patients with Ebola virus disease.* Emerg Infect Dis, 2014. **20**(10): p. 1683-90.
55. Wang, W., et al., *Quantification of circulating D-dimer by peptide immunoaffinity enrichment and tandem mass spectrometry.* Anal Chem, 2012. **84**(15): p. 6891-8.
56. Mairuhu, A.T., et al., *Is clinical outcome of dengue-virus infections influenced by coagulation and fibrinolysis? A critical review of the evidence.* Lancet Infect Dis, 2003. **3**(1): p. 33-41.
57. Davie, E.W., K. Fujikawa, and W. Kisiel, *The coagulation cascade: initiation, maintenance, and regulation.* Biochemistry, 1991. **30**(43): p. 10363-70.
58. Patti, G.J., O. Yanes, and G. Siuzdak, *Innovation: Metabolomics: the apogee of the omics trilogy.* Nat Rev Mol Cell Biol, 2012. **13**(4): p. 263-9.
59. Nordstrom, A., et al., *Nonlinear data alignment for UPLC-MS and HPLC-MS based metabolomics: quantitative analysis of endogenous and exogenous metabolites in human serum.* Anal Chem, 2006. **78**(10): p. 3289-95.
60. Tautenhahn, R., et al., *An accelerated workflow for untargeted metabolomics using the METLIN database.* Nat Biotechnol, 2012. **30**(9): p. 826-8.
61. Tautenhahn, R., et al., *XCMS Online: a web-based platform to process untargeted metabolomic data.* Anal Chem, 2012. **84**(11): p. 5035-9.
62. Gowda, H., et al., *Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses.* Anal Chem, 2014. **86**(14): p. 6931-9.
63. Sumner, L.W., P. Mendes, and R.A. Dixon, *Plant metabolomics: large-scale phytochemistry in the functional genomics era.* Phytochemistry, 2003. **62**(6): p. 817-36.
64. Panopoulos, A.D., et al., *The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming.* Cell Res, 2012. **22**(1): p. 168-77.
65. Kleinstreuer, N.C., et al., *Identifying developmental toxicity pathways for a subset of ToxCast chemicals using human embryonic stem cells and metabolomics.* Toxicol Appl Pharmacol, 2011. **257**(1): p. 111-21.
66. Denery, J.R., et al., *Metabolomics-based discovery of diagnostic biomarkers for onchocerciasis.* PLoS Negl Trop Dis, 2010. **4**(10).
67. Globisch, D., et al., *Onchocerca volvulus-neurotransmitter tyramine is a biomarker for river blindness.* Proc Natl Acad Sci U S A, 2013. **110**(11): p. 4218-23.
68. Wikoff, W.R., et al., *Metabolomic analysis of the cerebrospinal fluid reveals changes in phospholipase expression in the CNS of SIV-infected macaques.* J Clin Invest, 2008. **118**(7): p. 2661-9.
69. Pendyala, G., et al., *Cerebrospinal fluid proteomics reveals potential pathogenic changes in the brains of SIV-infected monkeys.* J Proteome Res, 2009. **8**(5): p. 2253-60.
70. Wikoff, W.R., et al., *Response and recovery in the plasma metabolome tracks the acute LCMV-induced immune response.* J Proteome Res, 2009. **8**(7): p. 3578-87.

71. Wilcken, B., et al., *Screening newborns for inborn errors of metabolism by tandem mass spectrometry*. N Engl J Med, 2003. **348**(23): p. 2304-12.
72. Garg, U. and M. Dasouki, *Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: clinical and laboratory aspects*. Clin Biochem, 2006. **39**(4): p. 315-32.
73. Centers for Disease, C. and Prevention, *Human rabies--Indiana and California, 2006*. MMWR Morb Mortal Wkly Rep, 2007. **56**(15): p. 361-5.
74. Willoughby, R.E., Jr., et al., *Survival after treatment of rabies with induction of coma*. N Engl J Med, 2005. **352**(24): p. 2508-14.
75. McDermid, R.C., et al., *Human rabies encephalitis following bat exposure: failure of therapeutic coma*. CMAJ, 2008. **178**(5): p. 557-61.
76. Aramburo, A., et al., *Failure of the Milwaukee protocol in a child with rabies*. Clin Infect Dis, 2011. **53**(6): p. 572-4.
77. O'Sullivan, A., et al., *Metabolomics of cerebrospinal fluid from humans treated for rabies*. J Proteome Res, 2013. **12**(1): p. 481-90.
78. Schutsky, K., et al., *Limited brain metabolism changes differentiate between the progression and clearance of rabies virus*. PLoS One, 2014. **9**(4): p. e87180.
79. Halstead, S.B., *Pathogenesis of dengue: challenges to molecular biology*. Science, 1988. **239**(4839): p. 476-81.
80. Monath, T.P., *Dengue: the risk to developed and developing countries*. Proc Natl Acad Sci U S A, 1994. **91**(7): p. 2395-400.
81. Cui, L., et al., *Serum metabolome and lipidome changes in adult patients with primary dengue infection*. PLoS Negl Trop Dis, 2013. **7**(8): p. e2373.
82. Voge, N.V., et al., *Metabolomics-Based Discovery of Small Molecule Biomarkers in Serum Associated with Dengue Virus Infections and Disease Outcomes*. PLoS Negl Trop Dis, 2016. **10**(2): p. e0004449.
83. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. Nature, 2003. **422**(6928): p. 198-207.
84. Weston, A.D. and L. Hood, *Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine*. J Proteome Res, 2004. **3**(2): p. 179-96.
85. Tremblay, M.J., J.F. Fortin, and R. Cantin, *The acquisition of host-encoded proteins by nascent HIV-1*. Immunol Today, 1998. **19**(8): p. 346-51.
86. Thomas, J.J., R. Bakhtiar, and G. Siuzdak, *Mass spectrometry in viral proteomics*. Acc Chem Res, 2000. **33**(3): p. 179-87.
87. Chosewood, L.C., et al., *Biosafety in microbiological and biomedical laboratories*. 5th ed. HHS publication. 2009, Washington, D.C.: U.S. Dept. of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institutes of Health. xxii, 415 p.
88. Grant-Klein, R.J., L.A. Altamura, and C.S. Schmaljohn, *Progress in recombinant DNA-derived vaccines for Lassa virus and filoviruses*. Virus Res, 2011. **162**(1-2): p. 148-61.
89. Urata, S., et al., *Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of Marburg virus-like particles induced by VP40, NP, and GP*. J Virol, 2007. **81**(9): p. 4895-9.

90. Urata, S., et al., *Cellular factors required for Lassa virus budding*. J Virol, 2006. **80**(8): p. 4191-5.
91. Spurgers, K.B., et al., *Identification of essential filovirion-associated host factors by serial proteomic analysis and RNAi screen*. Mol Cell Proteomics, 2010. **9**(12): p. 2690-703.
92. Want, E.J., et al., *Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry*. Anal Chem, 2006. **78**(3): p. 743-52.
93. Smith, C.A., et al., *METLIN: a metabolite mass spectral database*. Ther Drug Monit, 2005. **27**(6): p. 747-51.
94. Ogata, H., et al., *KEGG: Kyoto Encyclopedia of Genes and Genomes*. Nucleic Acids Res, 1999. **27**(1): p. 29-34.
95. Wishart, D.S., et al., *HMDB 3.0--The Human Metabolome Database in 2013*. Nucleic Acids Res, 2013. **41**(Database issue): p. D801-7.
96. Ihaka, R. and R. Gentleman, *R: a language for data analysis and graphics*. Journal of computational and graphical statistics, 1996. **5**(3): p. 299-314.
97. Motulsky, H., *Analyzing data with GraphPad prism*. 1999: GraphPad Software Incorporated.
98. Breiman, L., *Random forests*. Machine learning, 2001. **45**(1): p. 5-32.
99. Lê, S., J. Josse, and F. Husson, *FactoMineR: an R package for multivariate analysis*. Journal of statistical software, 2008. **25**(1): p. 1-18.
100. Searle, B.C., *Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies*. Proteomics, 2010. **10**(6): p. 1265-9.
101. Letunic, I. and P. Bork, *Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation*. Bioinformatics, 2007. **23**(1): p. 127-8.
102. Binns, D., et al., *QuickGO: a web-based tool for Gene Ontology searching*. Bioinformatics, 2009. **25**(22): p. 3045-6.
103. Zhang, H., et al., *Novel single-cell-level phenotypic assay for residual drug susceptibility and reduced replication capacity of drug-resistant human immunodeficiency virus type 1*. J Virol, 2004. **78**(4): p. 1718-29.
104. Connor, R.I., et al., *Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes*. Virology, 1995. **206**(2): p. 935-44.
105. Hamid, R., et al., *Comparison of alamar blue and MTT assays for high throughput screening*. Toxicol In Vitro, 2004. **18**(5): p. 703-10.
106. Denard, J., et al., *Quantitative proteomic analysis of lentiviral vectors using 2-DE*. Proteomics, 2009. **9**(14): p. 3666-76.
107. Bregnard, C., et al., *Comparative proteomic analysis of HIV-1 particles reveals a role for Ezrin and EHD4 in the Nef-dependent increase of virus infectivity*. J Virol, 2013. **87**(7): p. 3729-40.
108. Santos, S., et al., *Virus-producing cells determine the host protein profiles of HIV-1 virion cores*. Retrovirology, 2012. **9**: p. 65.
109. Saphire, A.C., P.A. Gallay, and S.J. Bark, *Proteomic analysis of human immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively distinguishes specific incorporated host proteins*. J Proteome Res, 2006. **5**(3): p. 530-8.

110. Ott, D.E., et al., *Actin-binding cellular proteins inside human immunodeficiency virus type 1*. *Virology*, 2000. **266**(1): p. 42-51.
111. Chertova, E., et al., *Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages*. *J Virol*, 2006. **80**(18): p. 9039-52.
112. Segura, M.M., et al., *Identification of host proteins associated with retroviral vector particles by proteomic analysis of highly purified vector preparations*. *J Virol*, 2008. **82**(3): p. 1107-17.
113. Chung, C.S., et al., *Vaccinia virus proteome: identification of proteins in vaccinia virus intracellular mature virion particles*. *J Virol*, 2006. **80**(5): p. 2127-40.
114. Varnum, S.M., et al., *Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome*. *J Virol*, 2004. **78**(20): p. 10960-6.
115. Loret, S., G. Guay, and R. Lippe, *Comprehensive characterization of extracellular herpes simplex virus type 1 virions*. *J Virol*, 2008. **82**(17): p. 8605-18.
116. Johannsen, E., et al., *Proteins of purified Epstein-Barr virus*. *Proc Natl Acad Sci U S A*, 2004. **101**(46): p. 16286-91.
117. Bechtel, J.T., R.C. Winant, and D. Ganem, *Host and viral proteins in the virion of Kaposi's sarcoma-associated herpesvirus*. *J Virol*, 2005. **79**(8): p. 4952-64.
118. Zhu, F.X., et al., *Virion proteins of Kaposi's sarcoma-associated herpesvirus*. *J Virol*, 2005. **79**(2): p. 800-11.
119. Bortz, E., et al., *Identification of proteins associated with murine gammaherpesvirus 68 virions*. *J Virol*, 2003. **77**(24): p. 13425-32.
120. Vidick, S., et al., *Proteomic characterization of murid herpesvirus 4 extracellular virions*. *PLoS One*, 2013. **8**(12): p. e83842.
121. Kramer, T., et al., *Proteomic characterization of pseudorabies virus extracellular virions*. *J Virol*, 2011. **85**(13): p. 6427-41.
122. Zhang, C., et al., *Profiling of cellular proteins in porcine reproductive and respiratory syndrome virus virions by proteomics analysis*. *Virol J*, 2010. **7**: p. 242.
123. Kang, S.M., et al., *Proteomic profiling of cellular proteins interacting with the hepatitis C virus core protein*. *Proteomics*, 2005. **5**(8): p. 2227-37.
124. Nuss, J.E., et al., *Multi-faceted proteomic characterization of host protein complement of Rift Valley fever virus virions and identification of specific heat shock proteins, including HSP90, as important viral host factors*. *PLoS One*, 2014. **9**(5): p. e93483.
125. Shaw, M.L., et al., *Cellular proteins in influenza virus particles*. *PLoS Pathog*, 2008. **4**(6): p. e1000085.
126. Ren, X., et al., *Proteomic analysis of purified Newcastle disease virus particles*. *Proteome Sci*, 2012. **10**(1): p. 32.
127. Radhakrishnan, A., et al., *Protein analysis of purified respiratory syncytial virus particles reveals an important role for heat shock protein 90 in virus particle assembly*. *Mol Cell Proteomics*, 2010. **9**(9): p. 1829-48.
128. Moerdyk-Schauwecker, M., S.I. Hwang, and V.Z. Grdzlishvili, *Analysis of virion associated host proteins in vesicular stomatitis virus using a proteomics approach*. *Virol J*, 2009. **6**: p. 166.

129. Ashburner, M., et al., *Gene ontology: tool for the unification of biology*. *The Gene Ontology Consortium*. Nat Genet, 2000. **25**(1): p. 25-9.
130. Gene Ontology, C., *Gene Ontology Consortium: going forward*. Nucleic Acids Res, 2015. **43**(Database issue): p. D1049-56.
131. Holmes, G.P., et al., *Lassa fever in the United States. Investigation of a case and new guidelines for management*. N Engl J Med, 1990. **323**(16): p. 1120-3.
132. Crowe, S.J., et al., *Prognostic Indicators for Ebola Patient Survival*. Emerg Infect Dis, 2016. **22**(2): p. 217-23.
133. Wauquier, N., et al., *Human fatal zaire ebola virus infection is associated with an aberrant innate immunity and with massive lymphocyte apoptosis*. PLoS Negl Trop Dis, 2010. **4**(10).
134. London, N.R., et al., *Targeting Robo4-dependent Slit signaling to survive the cytokine storm in sepsis and influenza*. Sci Transl Med, 2010. **2**(23): p. 23ra19.
135. Prescott, S.M., et al., *Platelet-activating factor and related lipid mediators*. Annu Rev Biochem, 2000. **69**: p. 419-45.
136. Caini, P., et al., *Modifications of plasma platelet-activating factor (PAF)-acetylhydrolase/PAF system activity in patients with chronic hepatitis C virus infection*. J Viral Hepat, 2007. **14**(1): p. 22-8.
137. Jeewandara, C., et al., *Platelet activating factor contributes to vascular leak in acute dengue infection*. PLoS Negl Trop Dis, 2015. **9**(2): p. e0003459.
138. Restrepo, B.N., et al., *[Serum platelet-activating factor acetylhydrolase activity in dengue patients of African or mestizo descendency]*. Biomedica, 2011. **31**(4): p. 599-607.
139. Souza, D.G., et al., *Essential role of platelet-activating factor receptor in the pathogenesis of Dengue virus infection*. Proc Natl Acad Sci U S A, 2009. **106**(33): p. 14138-43.
140. Johnson, K.M., et al., *Clinical virology of Lassa fever in hospitalized patients*. J Infect Dis, 1987. **155**(3): p. 456-64.
141. Prescott, S.M., G.A. Zimmerman, and T.M. McIntyre, *Platelet-activating factor*. J Biol Chem, 1990. **265**(29): p. 17381-4.
142. Nakamura, M., et al., *Molecular cloning and expression of platelet-activating factor receptor from human leukocytes*. J Biol Chem, 1991. **266**(30): p. 20400-5.
143. Ye, R.D., et al., *Characterization of a human cDNA that encodes a functional receptor for platelet activating factor*. Biochem Biophys Res Commun, 1991. **180**(1): p. 105-11.
144. Lukashevich, I.S., et al., *Lassa and Mopeia virus replication in human monocytes/macrophages and in endothelial cells: different effects on IL-8 and TNF-alpha gene expression*. J Med Virol, 1999. **59**(4): p. 552-60.
145. Venable, M.E., et al., *Platelet-activating factor: a phospholipid autacoid with diverse actions*. J Lipid Res, 1993. **34**(5): p. 691-702.
146. Simerville, J.A., W.C. Maxted, and J.J. Pahira, *Urinalysis: a comprehensive review*. Am Fam Physician, 2005. **71**(6): p. 1153-62.
147. Langridge, J.I., et al., *Gas chromatography/mass spectrometric analysis of urinary nucleosides in cancer patients; potential of modified nucleosides as tumour markers*. Rapid Commun Mass Spectrom, 1993. **7**(6): p. 427-34.

148. Sato, E., et al., *Metabolomic analysis of human plasma from haemodialysis patients*. Eur J Clin Invest, 2011. **41**(3): p. 241-55.
149. D'Ambrosio, S.M., R.E. Gibson-D'Ambrosio, and R.W. Trewyn, *An enzyme-linked immunosorbent assay (ELISA) for the detection and quantitation of the tumor marker 1-methylinosine in human urine*. Clin Chim Acta, 1991. **199**(2): p. 119-28.
150. Ott, D.E., *Cellular proteins detected in HIV-1*. Rev Med Virol, 2008. **18**(3): p. 159-75.
151. Ott, D.E. *AIDS and Cancer Virus Program*. Available from: <http://ncifrederick.cancer.gov/research/avp/>.
152. Cantin, R., S. Methot, and M.J. Tremblay, *Plunder and stowaways: incorporation of cellular proteins by enveloped viruses*. J Virol, 2005. **79**(11): p. 6577-87.
153. Calderone, A., L. Licata, and G. Cesareni, *VirusMentha: a new resource for virus-host protein interactions*. Nucleic Acids Res, 2015. **43**(Database issue): p. D588-92.
154. Navratil, V., et al., *VirHostNet: a knowledge base for the management and the analysis of proteome-wide virus-host interaction networks*. Nucleic Acids Res, 2009. **37**(Database issue): p. D661-8.
155. Yu, X., et al., *Exploration of panviral proteome: high-throughput cloning and functional implications in virus-host interactions*. Theranostics, 2014. **4**(8): p. 808-22.
156. Cantin, R., et al., *Discrimination between exosomes and HIV-1: purification of both vesicles from cell-free supernatants*. J Immunol Methods, 2008. **338**(1-2): p. 21-30.
157. Leroy, B., et al., *Structural Proteomics of Herpesviruses*. Viruses, 2016. **8**(2).
158. Schlie, K., et al., *Viral protein determinants of Lassa virus entry and release from polarized epithelial cells*. J Virol, 2010. **84**(7): p. 3178-88.
159. Connor, J.H., et al., *Antiviral activity and RNA polymerase degradation following Hsp90 inhibition in a range of negative strand viruses*. Virology, 2007. **362**(1): p. 109-19.
160. Smith, D.R., et al., *Inhibition of heat-shock protein 90 reduces Ebola virus replication*. Antiviral Res, 2010. **87**(2): p. 187-94.
161. Reyes-Del Valle, J., et al., *Heat shock protein 90 and heat shock protein 70 are components of dengue virus receptor complex in human cells*. J Virol, 2005. **79**(8): p. 4557-67.
162. de Chasseay, B., et al., *Virus-host interactomics: new insights and opportunities for antiviral drug discovery*. Genome Med, 2014. **6**(11): p. 115.
163. Fu, W., et al., *Human immunodeficiency virus type 1, human protein interaction database at NCBI*. Nucleic Acids Res, 2009. **37**(Database issue): p. D417-22.
164. Ni, S., et al., *Discovering potent small molecule inhibitors of cyclophilin A using de novo drug design approach*. J Med Chem, 2009. **52**(17): p. 5295-8.
165. Rak, R., et al., *Novel LIMK2 Inhibitor Blocks Panc-1 Tumor Growth in a mouse xenograft model*. Oncoscience, 2014. **1**(1): p. 39-48.
166. Schlecht, R., et al., *Functional analysis of Hsp70 inhibitors*. PLoS One, 2013. **8**(11): p. e78443.

167. Yu, Y., et al., *Withaferin A targets heat shock protein 90 in pancreatic cancer cells*. *Biochem Pharmacol*, 2010. **79**(4): p. 542-51.
168. Bargagna-Mohan, P., et al., *The tumor inhibitor and antiangiogenic agent withaferin A targets the intermediate filament protein vimentin*. *Chem Biol*, 2007. **14**(6): p. 623-34.
169. Sagot, Y., et al., *An orally active anti-apoptotic molecule (CGP 3466B) preserves mitochondria and enhances survival in an animal model of motoneuron disease*. *Br J Pharmacol*, 2000. **131**(4): p. 721-8.
170. Andringa, G., et al., *TCH346 prevents motor symptoms and loss of striatal FDOPA uptake in bilaterally MPTP-treated primates*. *Neurobiol Dis*, 2003. **14**(2): p. 205-17.
171. Cervantes-Salazar, M., et al., *Dengue virus NS1 protein interacts with the ribosomal protein RPL18: this interaction is required for viral translation and replication in Huh-7 cells*. *Virology*, 2015. **484**: p. 113-26.
172. Allonso, D., et al., *Dengue Virus NS1 Protein Modulates Cellular Energy Metabolism by Increasing Glyceraldehyde-3-Phosphate Dehydrogenase Activity*. *J Virol*, 2015. **89**(23): p. 11871-83.